

VCAP-DCA Study Guide v1.0

Ed Grigson, April 2011

1 Contents

1	STORAGE	1
1.1	IMPLEMENT AND MANAGE COMPLEX STORAGE SOLUTIONS	1
1.2	MANAGE STORAGE CAPACITY IN A VSPHERE ENVIRONMENT	11
1.3	CONFIGURE AND MANAGE COMPLEX MULTI-PATHING AND PSA PLUGINS	17
2	NETWORK	24
2.1	IMPLEMENT AND MANAGE COMPLEX VIRTUAL NETWORKS	24
2.2	CONFIGURE AND MAINTAIN VLANS, PVLANS AND VLAN SETTINGS	30
2.3	DEPLOY AND MAINTAIN SCALABLE VIRTUAL NETWORKING	33
2.4	ADMINISTER VNETWORK DISTRIBUTED SWITCH SETTINGS	36
3	DEPLOY DRS CLUSTERS AND MANAGE PERFORMANCE	39
3.1	TUNE AND OPTIMIZE VSPHERE PERFORMANCE	39
3.2	OPTIMIZE VIRTUAL MACHINE RESOURCES	48
3.3	IMPLEMENT AND MAINTAIN COMPLEX DRS SOLUTIONS	52
3.4	PERFORM CAPACITY PLANNING IN A VSPHERE ENVIRONMENT	59
3.5	UTILIZE ADVANCED VSPHERE PERFORMANCE MONITORING TOOLS	61
4	BUSINESS CONTINUITY	64
4.1	IMPLEMENT AND MAINTAIN COMPLEX VMWARE HA SOLUTIONS	64
4.2	DEPLOY AND TEST VMWARE FT	69
4.3	CONFIGURE A VSPHERE ENVIRONMENT TO SUPPORT MSCS CLUSTERING	74
4.4	DEPLOY AND MAINTAIN VCENTER SERVER HEARTBEAT	79
5	PERFORM OPERATIONS MAINTENANCE	88
5.1	IMPLEMENT AND MAINTAIN HOST PROFILES	88
5.2	DEPLOY AND MANAGE COMPLEX UPDATE MANAGER ENVIRONMENTS	93
6	PERFORM ADVANCED TROUBLESHOOTING	99
6.1	CONFIGURE, MANAGE, AND ANALYSE VSPHERE LOG FILES	99
6.2	TROUBLESHOOT CPU AND MEMORY PERFORMANCE	104
6.3	TROUBLESHOOT NETWORK PERFORMANCE AND CONNECTIVITY	106
6.4	TROUBLESHOOT STORAGE PERFORMANCE AND CONNECTIVITY	113
6.5	TROUBLESHOOT VCENTER SERVER AND ESX/ESXI HOST MANAGEMENT	117
7	SECURE A VSPHERE ENVIRONMENT	120
7.1	SECURE ESX/ESXI HOSTS	120
7.2	CONFIGURE AND MAINTAIN THE ESX FIREWALL	127
7.3	DEPLOY AND ADMINISTER VSHIELD ZONES	133
8	SCRIPTING AND AUTOMATION	143
8.1	EXECUTE VMWARE CMDLETS AND CUSTOMIZE SCRIPTS USING POWERCLI	143
8.2	ADMINISTER VCENTER ORCHESTRATOR	150
8.3	ADMINISTER VSPHERE USING THE VMA	155
9	ADVANCED INSTALLATIONS	161

9.1	INSTALL ESX SERVER WITH CUSTOM SETTINGS.....	161
9.2	PLAN AND EXECUTE SCRIPTED INSTALLATIONS	167
9.3	CONFIGURE VCENTRE SERVER LINKED MODE	173
10	SCENARIO QUESTIONS	ERROR! BOOKMARK NOT DEFINED.
11	APPENDIX B - ADVANCED PARAMETERS YOU MIGHT HAVE TO RECALL.....	178

1 Storage

1.1 Implement and Manage complex storage solutions

Knowledge

- Identify RAID levels
- Identify supported HBA types
- Identify virtual disk format types

Skills and Abilities

- Determine use cases for and configure VMware DirectPath I/O
- Determine requirements for and configure NPIV
- Determine appropriate RAID level for various Virtual Machine workloads
- Apply VMware storage best practices
- Understand use cases for Raw Device Mapping
- Configure vCenter Server storage filters
- Understand and apply VMFS resignaturing
- Understand and apply LUN masking using PSA-related commands
- Analyze I/O workloads to determine storage performance requirements

Tools & learning resources

- Product Documentation
 - [Fibre Channel SAN Configuration Guide](#)
 - [iSCSI SAN Configuration Guide](#)
 - [ESX Configuration Guide](#)
 - [ESXi Configuration Guide](#)
 - [vSphere Command-Line Interface Installation and Scripting Guide](#)
 - [I/O Compatibility Guide](#)
- vSphere CLI
 - vscsiStats, vicfg-*, vifs, vmkfstools, esxstop/resxstop
- [Storage Best Practices for Scaling Virtualisation Deployments](#) (TA2509, VMworld 2009)
- [Best Practices for Managing and Monitoring Storage](#) (VM3566, VMworld '09)
- [Storage Best Practices and Performance Tuning](#) (TA8065, VMworld 2010, subscription required)
- [Analyse I/O workloads](#) (at vSpecialist's blog)
- [Sean Crookston's study notes](#)

Storage is an area where you can never know too much. For many infrastructures storage is the most likely cause of performance issues and a source of complexity and misconfiguration – especially given that many VI admins come from a server background (not storage) due to VMware's server consolidation roots.

1.1.1 Identify RAID levels

Common RAID types: 0, 1, 5, 6, 10. [Wikipedia do a good summary](#) of the basic RAID types if you're not familiar with them.

The impact of RAID types will vary depending on your storage vendor and how they implement RAID. Netapp (which I'm most familiar with) using a proprietary RAID-DP which is [like RAID-6 but without the performance penalties](#) (so Netapp say).

Scott Lowe has [a good article about RAID in storage arrays](#), as does [Josh Townsend over at VMtoday](#).

1.1.2 Supported HBA types

The best (only!) place to look for real world info is [VMware's HCL](#) (which is now an online, searchable repository). Essentially it comes down to Fibre Channel or iSCSI HBAs. You should not mix HBAs from different vendors in a single server. It can work but isn't officially supported.

Remember you can have a maximum of 8 HBAs or 16 HBA ports per ESX/ESXi server.

This is a slightly odd exam topic – presumably we won't be buying HBAs as part of the exam so what's there to know? 😊

1.1.3 Identify virtual disk format types

Virtual disk (VMDK) format types:

- Eagerzeroedthick
- Zeroedthick (default)
- Thick
- Thin

Three factors primarily determine the disk format;

- Initial disk size
- Blanking underlying blocks during initial creation
- Blanking underlying blocks when deleting data in the virtual disk (reclamation)

The differences stem from whether the physical files are 'zeroed' or not (ie where there is no data in the 'virtual' disk what in the underlying VMDK?). Several features (such as FT and MSCS) require an 'eagerzeroedthick' disk. Check out this [great diagram](#) (courtesy of Steve Foskett) which shows the differences.

The other possible type is an RDM which itself can have two possible types;

- RDM (virtual) – enables snapshots, vMotion but masks some physical features
- RDM (physical) – required for MSCS clustering and some SAN applications

1.1.4 DirectPath I/O

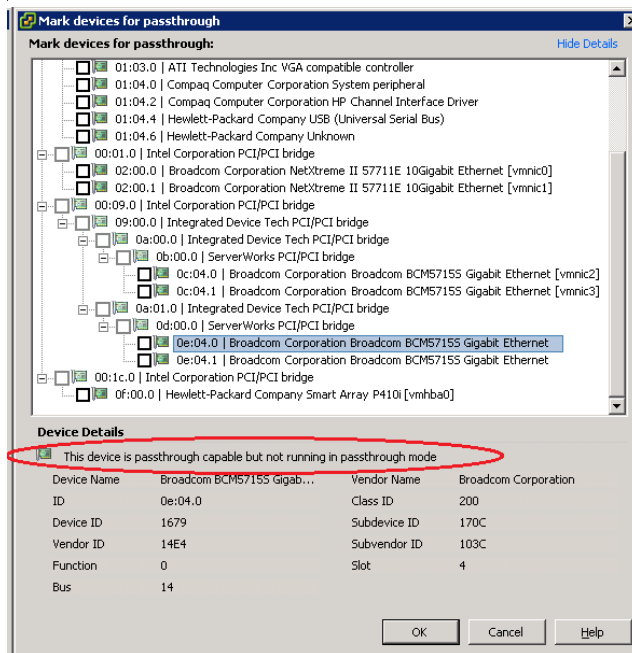
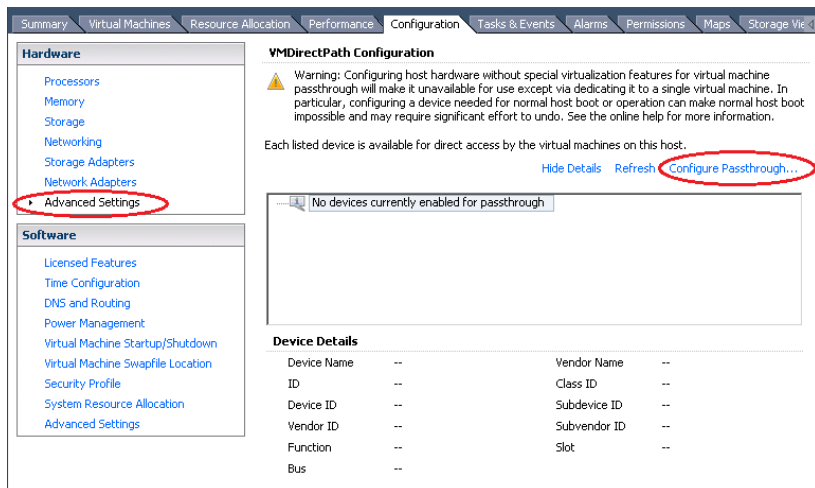
Lets a VM bypass the virtualisation layer and speak directly to a PCI device. Benefits are reduced CPU on the host, and potentially slightly higher I/O to a VM when presenting a 10GB NIC, alternatively you could present a physical USB device directly to a VM ([see this example at Petri.nl](#), link courtesy of [Sean Crookston's study notes](#))

Requirements

- Intel Nehalem only (experimental support for AMD)
- Very limited device support (10GB Ethernet cards, and only a few). As usual the list of devices which work will be much larger than the officially certified HCL (the quad port card for my HP BL460G6 worked fine as do USB devices)
- Once a device is used for passthrough it's NOT available to the host and therefore other VMs

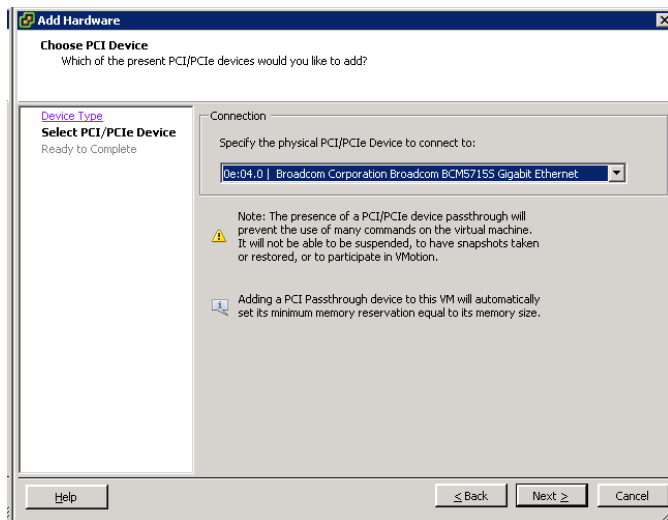
Configuring the host (step 1 of 2)

1. Configure PCI device at host level (Configuration -> Advanced Settings under Hardware). Click 'Configure Passthrough' and select a device from the list.
NOTE: If the host doesn't support DirectPath a message to that effect will be shown in the display window.
2. Reboot the host for the changes to take effect.



Configuring the VM (step 2 of 2)

1. Edit the VM settings and add a PCI Device.
NOTE: The VM must be powered off.
2. Select the passthrough device from the list (which only shows enabled devices). There is a warning that enabling this device will limit features such as snapshots and vMotion).



If you want in-depth information about VMDirectPath read this [VMware whitepaper](#).

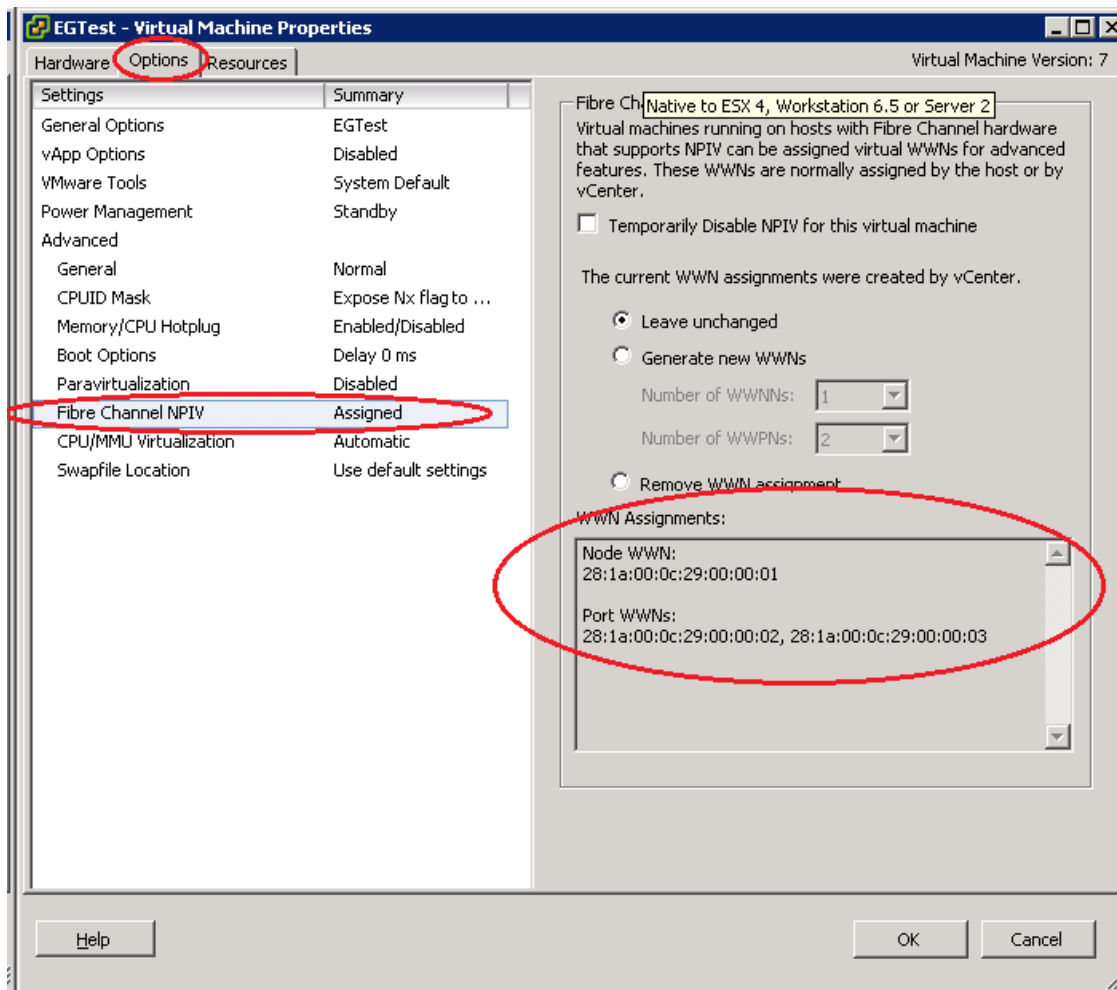
To check: does enabling DirectPath I/O on a VM set a memory reservation? *P62 ESXi configuration guide*.

1.1.5 NPIV

Stands for N-Port ID Virtualisation. This allows a single HBA adaptor port (provided it supports NPIV) to register multiple WWPN's with the SAN fabric, rather than the single address normally registered. You can then present one of these WWPN's directly to a VM, thus allowing you to zone storage to a specific VM rather than a host (which is normally the only option). Read more in [Scott Lowe's blogpost](#), [Jason Boche's \(in depth\) blogpost](#), [Simon Long's post](#), and [Nick Triantos' summary](#). They left me wondering what the real world benefit is to VI admins!

To use NPIV;

1. In the VM properties, to go Options -> NPIV.
NOTE: These options will only be enabled if the VM has an RDM attached. Even if enabled it does not guarantee that the HBA/switches support NPIV.
2. For a new VM, click 'Generate new WWNs'
3. For an existing VM (which is already NPIV enabled) click either;
 - a. 'Generate WWNs' to change the WWN assigned to this VM
 - b. Temporarily Disable WWN
 - c. Remove WWN
4. You'll also have to add the newly generated WWPN's to your SAN zones and storage array masking (Initiator groups in the case of Netapp).



NPIV Requirements

- HBAs and SAN switches must support NPIV.
- NPIV only works with RDM disks
- svMotion on an NPIV enabled VM is not allowed (although vMotion is)

1.1.6 RDM

Joep Piscaer has written up [a good summary of RDMs](#), and from that article –“RDM’s gives you some of the advantages of direct access to a physical device while keeping some advantages of a virtual disk in VMFS. As a result, they merge VMFS manageability with raw device access”.

Use cases include;

- Various types of clustering including MSCS (see section 4.2) and Oracle OCFS/ASM
- NPIV
- Anytime you want to use underlying storage array features (such as snapshots). Some SAN management software needs direct access to the underlying storage such as Netapp’s SnapManager suite for Exchange and SQL.

Two possible modes

- Virtual compatibility
- Physical compatibility

Created;

- like any other VMDK through the VI client, then select RDM and choose mode
- using `vmkfstools -z` or `vmkfstools -r` (see section 1.2 for details)
- requires block storage (FC or iSCSI)

NOTE: When cloning a VM with RDM's (in virtual compatibility mode) they will be converted to VMDKs. Cloning a VM with an RDM (in physical compatibility mode) is not supported.

1.1.7 Storage Filters

Storage filters are used to adjust default vCenter behaviour when scanning storage. See this post [about storage filters at Duncan Epping's site](#).

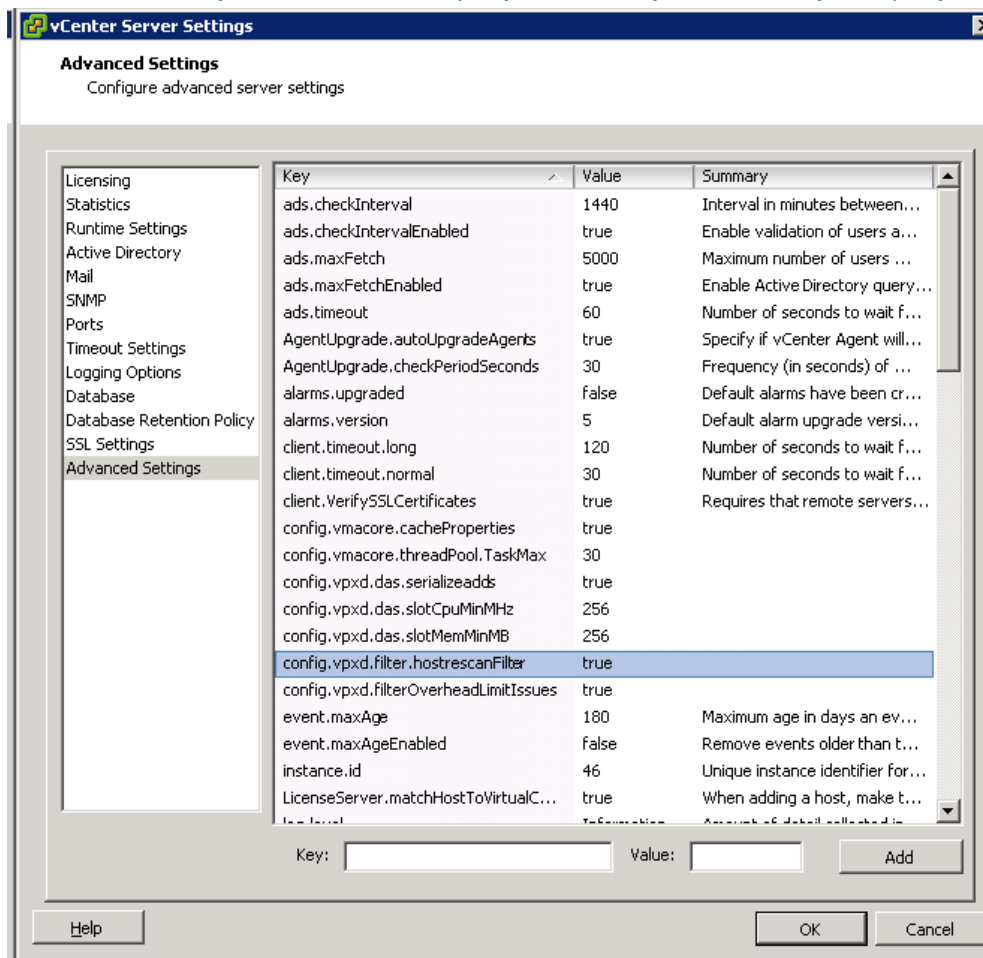
There are four filters (all of which are enabled by default);

1. Host rescan (config.vpxd.filter.hostrescanFilter)
2. RDM filter (config.vpxd.filter.rdmFilter)
3. VMFS (config.vpxd.filter.vmfsFilter)
4. Same hosts and transport (config.vpxd.filter.SameHostAndTransportsFilter)

Configuring storage filters is done in vCenter (not per host);

1. Go to Administration -> vCenter Settings -> Advanced Settings
2. Add a key for the filter you want to enable and set the key to FALSE or TRUE.

NOTE: All filters are enabled by default (value if TRUE) even if not specifically listed.



Turning off the 'Host Rescan' filter does NOT stop newly created LUNs being automatically scanned for – it simply stops each host automatically scanning when newly created VMFS Datastores are added on another host. This is useful when you're adding a large number of VMFS Datastores in one go (200 via PowerCLI for example) and you want to complete the addition before rescanning all hosts in a cluster (otherwise each host could perform 200 rescans). See p50 of the FC SAN Configuration Guide.

One occasion where the VMFS filter might be useful is extending a VMFS volume. With vSphere this is now supported but I've had intermittent success when expanding a LUN presented by a Netapp array. The LUN (and underlying volume) has been resized OK but when I try to extend the VMFS no valid LUNs are presented. Next time this happens I can try turning off the storage filters (VMFS in particular) and see if maybe the new space isn't visible to all hosts that share the VMFS Datastore.

1.1.8 VMFS Resignaturing

LUN Resignaturing is used when you present a copy of a LUN to an ESX host, typically created via a storage array snapshot. Been around since VI3 but ease of use has increased since.

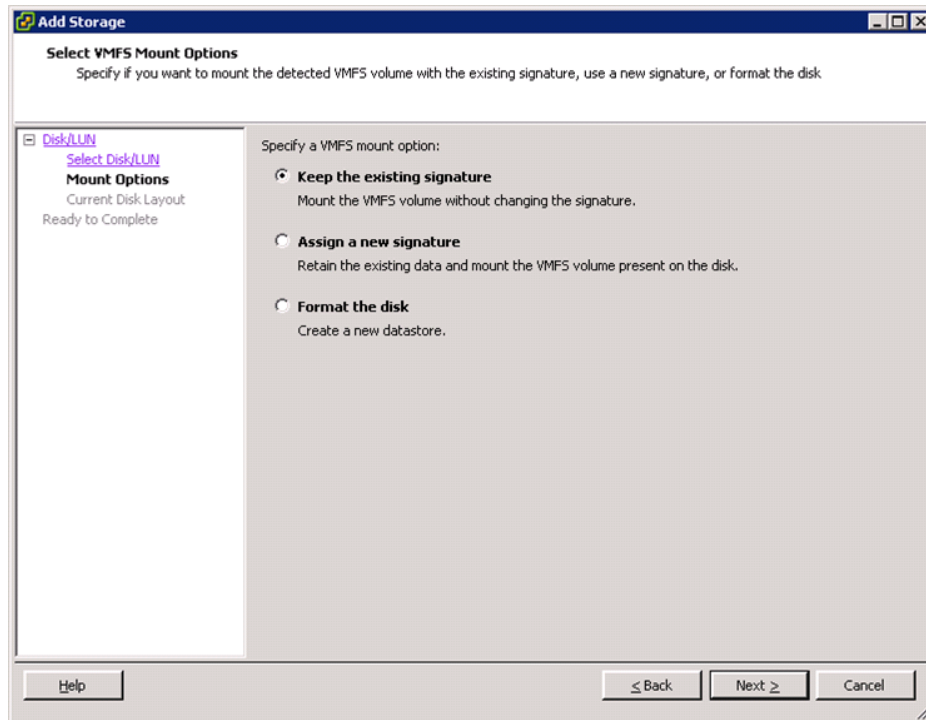
NOTE: This doesn't apply to NFS datastores as they don't embed a UUID in the metadata.

Resignaturing a LUN copy using the VI Client;

1. Click Add Storage on a host and select the LUN copy.
2. On the next screen choose either;
 - a. Keep existing signature. This can only be done if the original VMFS Datastore is offline or unavailable to this host (you're trying to mount a mirrored volume at a DR site for example).

NOTE: If you try and the other VMFS Datastore is accessible you'll get an error stating that the host configuration change was not possible and the new datastore won't be mounted.
 - b. Assign a new signature (data is retained). This is persistent and irreversible.

- c. Format the disk. This assigns a new signature but any existing data IS LOST.



Resignaturing a LUN copy using the command line (use vicfg-volume from RCLI or vMA);

1. 'esxcfg-volume -l' to see a list of copied volumes
2. Choose either;
 - a. 'esxcfg-volume -r <previous VMFS label | UUID>' to resignature the volume
 - b. 'esxcfg-volume -M <previous VMFS label | UUID>' to mount the volume without resignaturing (use lower case m for temporary mount rather than persistent).

Full details for these operations can be followed in [VMwareKB1011387](https://kb.vmware.com/s/article/1011387)

1.1.9 LUN Masking

This can be done in various ways (Netapp implement LUN Masking through the use of Initiator Groups) but for the VCAP-DCA they're referring to PSA rules. For a good overview of both why you might want to do this at the hypervisor layer (and how) see this [blogpost at Punching Clouds](#).

You can mask;

- Complete storage array
- One or more LUNs
- Specific paths to a LUN

To mask a particular LUN (for example);

1. Get a list of existing claim rules (so you can get a free ID for your new rule);
esxcli corestorage claimrule list
NOTE: You can skip this if you use -u to autoassign a new ID
2. Create a new rule to mask the LUN;
esxcli corestorage claimrule add --rule <number> -t location -A <hba_adapter> -C <channel> -T <target> -L <lun> -P MASK_PATH

For example;

```
esxcli corestorage claimrule add --rule 120 -t location -A vmhba1 -C 0 -T 0 -L 20 -P  
MASK_PATH
```

3. Load this new rule to make MASK_PATH module the owner;

```
esxcli corestorage claimrule load
```
4. Unclaim existing paths to the masked LUN. Unclaiming disassociates the paths from a PSA plugin. These paths are currently owned by NMP. You need to dissociate them from NMP and have them associated to MASK_PATH;

```
esxcli corestorage claiming reclaim -t location -A vmhba1 -C 0 -T 0 -L 20
```

NOTE: You cant reclaim a path that's active.
5. Run the claim rules

```
esxcli corestorage claimrule run
```
6. Verify that the LUN/datastore is no longer visible to the host.

```
esxconfig-scsidevs --vmfs
```

This is a pretty convoluted procedure which I hope I don't have to remember in the exam! [VMwareKB1009449](#) describes this process in detail and it's also documented on p82 of the FC SAN Configuration Guide and p96 of the Command Line Interface Installation and Reference Guide (both of which should be available during the exam).

To determine if you have any masked LUNs;

1. List the visible paths on the ESX host and look for any entries containing MASK_PATH

```
esxconfig-mpath -L | grep MASK_PATH
```

Obviously if you want to mask a LUN from all the hosts in a cluster you'd have to run these commands on every host. William Lam's done [a useful post about automating esxcli](#). There's also an [online reference to the esxcli command line](#).

NOTE: It's recommended that you follow the procedure above EVERYTIME you remove a LUN from a host! So if you have 16 hosts in a cluster and you want to delete one empty LUN you have to

1. Mask the LUN at the VMkernel layer on each of the hosts
2. Unpresent the LUN on your storage array
3. Cleardown the masking rules you created in step 1 (again for every host)

Seem unbelievable? Read [VMwareKB1015084](#)...

1.1.10 Analyse I/O workloads

There are numerous tools to analyse I/O - IOMeter, vSCSIStats, Perfmon, esxtop etc . Things to measure;

- IOPs
- Throughput
- Bandwidth
- Latency
- Workload pattern – percentage reads vs writes, random vs sequential, packet sizes (small vs large)

- IO adjacency (this is where vscsiStats comes in)

You can read more

There's a useful [post at vSpecialist.co.uk](#), plus some [good PowerCLI scripts to collect metrics via esxtop \(courtesy of Clinton Kitson\)](#)

VMware publish plenty of resources to help people virtualise existing apps, and that includes the I/O requirements (links courtesy of [Scott Lowe](#));

- [Exchange](#)
- [SQL Server](#)
- [Oracle](#)

1.1.11 Storage best practices

How long is a piece of string?

- [VMware white paper on storage best practices](#)
- [Storage Best Practices for Scaling Virtualisation Deployments](#) (TA2509, VMworld 2009)
- [Best Practices for Managing and Monitoring Storage](#) (VM3566, VMworld '09)
- [Storage Best Practices and Performance Tuning](#) (TA8065, VMworld 2010, subscription required at time of writing)
- [Netapp and vSphere storage best practices](#) (including filesystem alignment)
- Read Scott Lowe's Mastering VMware vSphere 4 book, chapter 6 (where he specifically talks about storage best practices)
- Stephen Fosket's [presentation about thin provisioning](#)

1.2 Manage storage capacity in a vSphere environment

Knowledge

- Identify storage provisioning methods
- Identify available storage monitoring tools, metrics and alarms

Skills and Abilities

- Apply space utilization data to manage storage resources
- Provision and manage storage resources according to Virtual Machine requirements
- Understand interactions between virtual storage provisioning and physical storage provisioning
- Apply VMware storage best practices
- Configure datastore alarms
- Analyze datastore alarms and errors to determine space availability

Tools & learning resources

- Product Documentation
 - [vSphere Datacenter Administration Guide](#)
 - [Fibre Channel SAN Configuration Guide](#)
 - [iSCSI SAN Configuration Guide](#)
 - [vSphere Command-Line Interface Installation and Scripting Guide](#)
- vSphere Client
- vSphere CLI
 - vmkfstools
- [Marc Polo's study guide notes](#)

Managing storage capacity is another potentially huge topic, even for a midsized company. The storage management functionality within vSphere is fairly comprehensive.

1.2.1 Storage provisioning methods

There are three main protocols you can use to provision storage;

- Fibre channel
 - Block protocol
 - Uses multipathing (PSA framework)
 - Configured via vicfg-mpath, vicfg-scsidevs
- iSCSI
 - block protocol
 - Uses multipathing (PSA framework)
 - hardware or software (boot from SAN is h/w initiator only)
 - configured via vicfg-iscsi, esxcfg-swiscsi and esxcfg-hwiscsi, vicfg-mpath, esxcli
- NFS
 - File level (not block)
 - No multipathing (uses underlying Ethernet network resilience)
 - Thin by default
 - no RDM, MSCS
 - configured via vicfg-nas

I won't go into much detail on each, just make sure you're happy provisioning storage for each protocol both in the VI client and the CLI.

Know the various options for provisioning storage;

- VI client. Can be used to create/extend/delete all types of storage. VMFS volumes created via the VI client are automatically aligned.
- CLI – vmkfstools.
 - NOTE: When creating a VMFS datastore via CLI you need to align it. Check VMFS alignment using 'fdisk -lu'. Read more in [Duncan Epping's blogpost](#).
- PowerCLI. Managing storage with PowerCLI - [VMwareKB1028368](#)
- Vendor plugins (Netapp RCU for example). I'm not going to cover this here as I doubt the VCAP-DCA exam environment will include (or assume any knowledge of) these!

NOTE: With vSphere you can now extend an existing VMFS partition (assuming you have contiguous space in the underlying LUN).

When provisioning storage there are various considerations;

- Thin vs thick
- Extents vs true extension
- Local vs FC/iSCSI vs NFS
- VMFS vs RDM

Using vmkfstools

The most useful vmkfstools commands;

- `vmkfstools -c 10GB <path to VMDK>`
Create a 10GB VMDK (defaults to zeroedthick with a BUSLOGIC adapter).
- `vmkfstools -c 10g -d eagerzeroedthick -a lsilogic /vmfs/volumes/zcglabsvr7/local/test.vmdk`
Create a 10GB VMDK in eagerzeroedthick format with an LSILOGIC adapter
- `vmkfstools -X 10g <path to VMDK>`
Extend the virtual disk by 10GB
- `vmkfstools -i <path to VMDK>`
Inflate the virtual disk, defaults to eagerzeroedthick. (thin to thick provisioning)
- `vmkfstools -r /vmfs/devices/disks/naa.600c0ff000d5c3830473904c01000000 myrdm.vmdk`
`vmkfstools -z /vmfs/devices/disks/naa.600c0ff000d5c3830473904c01000000 myrdmp.vmdk`
The top command creates an RDM in virtual compatibility mode
The bottom command creates an RDM in physical compatibility mode
- `vmkfstools -D <path to VMDK>`
Check the format of a VMDK to determine if it's eagerzeroedthick (required for FT and MS clustering). If the 'tbz' value is zero the disk is eagerzeroedthick. You can read a full description for this process in [VMwareKB1011170](#)

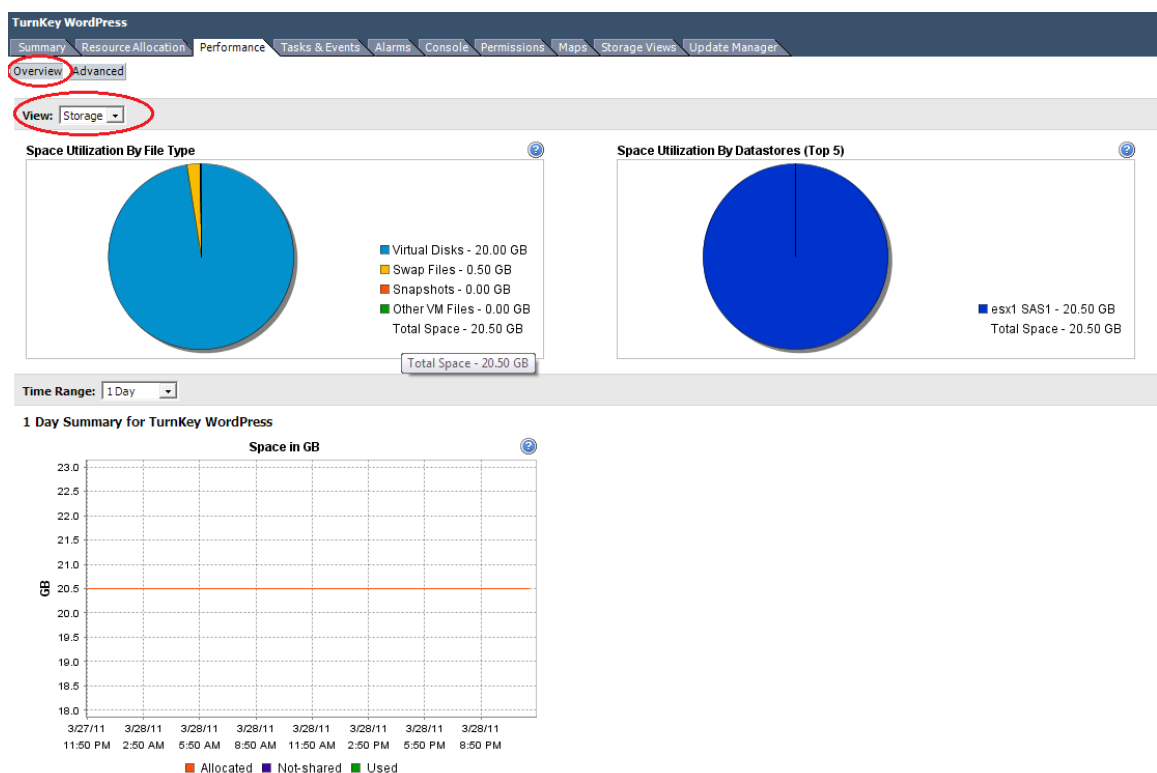
```
/vmfs/volumes/4c90802a-eb1b5cbf-d8d9-001b78373634/W2k3 r2 x32 # vmkfstools -D ./W2k3\ r2\ x32.vmdk
Lock [type 10c00001 offset 51613696 v 86, hb offset 3756032
gen 47, mode 0, owner 00000000-00000000-0000-000000000000 mtime 4477]
Addr <4, 64, 34>, gen 36, links 1, type reg, flags 0, uid 0, gid 0, mode 100600
len 524, nb 1, tbz 0, cow 0, zla 2, bs 65536
/vmfs/volumes/4c90802a-eb1b5cbf-d8d9-001b78373634/W2k3 r2 x32 #
```

[Another very useful VMwareKB article about vmkfstools](#)

1.2.2 Storage monitoring tools, metrics and alarms

Native storage monitoring GUI tools include datastore alarms, the 'Storage Views' plugin (new to vSphere), vCenter performance charts and the datastore inventory view.

- The datastore view lets you see which hosts have a given datastore mounted along with free space, track events and tasks per datastore (for auditing for example), along with permissions per datastore.
- The Storage Views tab shows more detailed information about capacity, pathing status, storage maps, snapshot sizes (very useful) etc
- vCenter performance charts help you analyse bandwidth, latency, I/Os and more. Using vCenter Charts is covered in section 3.4, Perform Capacity Planning in a vSphere environment.



NOTE: Most array vendors provide tools which are far more sophisticated than the native vSphere tools for monitoring capacity, performance and vendor specific functionality (such as dedupe) although they're rarely free!

Creating datastore alarms

When creating an alarm you can monitor either 'events' or 'condition changes' on the object in question, in this case datastores.

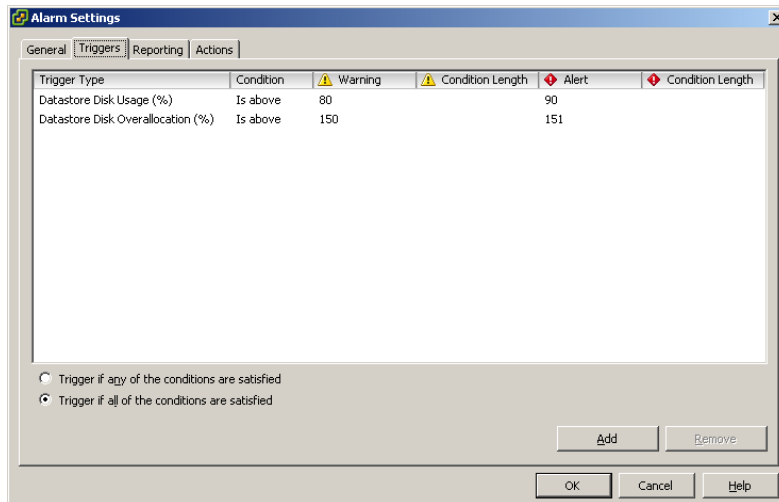
When configured for event monitoring you can monitor;

- VMFS created/extending/deleting a new datastore
- NAS created/extending/deleting a new datastore
- File or directory copied/deleted/moved

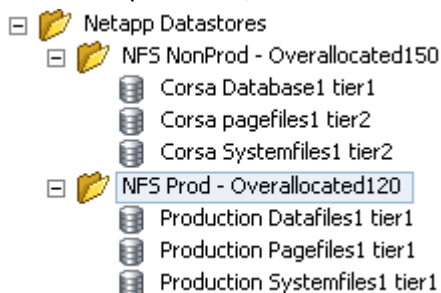
When configured for 'state' changes you can monitor;

- Disk % full
- Disk % over allocated
- Datastore to All Hosts

While these seem straightforward there are complications. If you're using NFS and array level deduplication for example the 'savings' from dedupe will be reflected in vCenter. If you set an alarm only to monitor '%Disk full' then you may find you've massively overprovisioned the volume (NFS is also thin provisioned by default) before you reach a capacity alert. The solution is to monitor both '%Disk overallocation' *and* '%Disk full' and only generate an alert if both conditions are met.



A solution I've worked with is to create multiple 'service levels' by grouping your datastores into separate folders and then setting different alarms on each folder. For instance I want an alert if a production datastore is overprovision by more than 120% but I'm happy for this threshold to be 150% for nonprod VMs;



There are also storage capacity related alarms on the VM objects;

- VM Total size on disk
- VM snapshot size
- VM Disk usage (kbps)

1.2.3 Understand interactions between virtual storage and physical storage

Wide open topic, which will also vary depending on the storage vendor. Always remember that a 'simple' change can have unexpected implications;

- Increasing capacity (extending a LUN using an extent) could improve (or decrease) performance depending on the underlying storage array (for example the new LUN is on few

slow spindles). Your requirement to increase capacity might also impact availability – if a second LUN is provisioned and added as an extent you now have two LUNs to manage (and a failure of either means the datastore goes offline).

- Consider queues - disk queues, HBA queues, LUN queues etc. [Read more](#).
- Consider SCSI reservations (number of VMs per VMFS). [VMwareKB1005009](#). See section 6.4 for information about troubleshooting SCSI reservations. [Read more here](#)
- If you run a disk defragment (or a full disk format) in the guest OS it will inflate the underlying thin provisioned VMDK to maximum size.
- Thin or thick may have performance implications - if the disk is eagerzeroedthick then the blocks in the underlying storage needs to be zeroed out, resulting in provisioning taking longer. NOTE: There is very little performance impact from using thin disks under typical circumstances. ([VMware whitepaper on thin provisioning performance](#)).
- An svMotion could be moving a VM from a tier 1 datastore (high performance storage) to a datastore on tier 2 (slower) storage. A well planned naming scheme should make you aware of these implications.
- Doing a vMotion can't impact storage right? The VMDK's are shared and don't move. But what if you're virtual swap files are on local disk? Ah...
- Create one VM in the datastore won't necessarily consume the same amount of storage on the underlying array and might impact DR– dedupe, snap mirror, snap reserve all impacted
- VAAI – vSphere 4.1 has introduced 'array offloading' which affects interaction between virtual and physical storage, but today the VCAP-DCA lab is built on vSphere 4.0.
- You can change disk format while doing a svMotion (from thick to thin for example). These operations have an impact on your storage array so consider doing mass migrations out of hours (or whenever your array is less busy).

1.2.4 Apply VMware Storage best practices

Section 6.4 (Troubleshooting Storage) is where the blueprint lists 'recall vSphere maximums' but it makes more sense to cover them here as they impact storage capacity planning. The relevant limits;

- 255 LUNs per host
- 32 paths per LUN
- 1024 paths per host
- 256 VMs per VMFS volume
- 2TB -512 bytes max per VMFS extent
- 32 extents per VMFS (for a max 64TB volume)
- 8 NFS datastores (default, can be increased to 64)
- 8 HBAs per host
- 16 HBA ports per host

Reference the full list; [vSphere Configuration Maximums](#)

Provision and manage storage resources according to Virtual Machine requirements

- When sizing VMFS volumes ensure you account for snapshots, VM swapfiles etc as well as VMDK disk sizes.
- When sizing LUNs there are two primary options;

- Predictive sizing. Understand the IO requirements of your apps first, then create multiple VMFS volumes with different storage characteristics. Place VMs in the appropriate LUN.
- Adaptive sizing. Create a few large LUNs and start placing VMs in them. Check performance and create new LUNs when performance is impacted.
- Also consider that different VM workloads may need different performance characteristics (ie underlying RAID levels) which can affect LUN size and layout.
- Use disk shares (and SIOC with vSphere 4.1) to control storage contention. With disk shares the contention is regulated per host, with SIOC it's regulated per cluster.

1.3 Configure and manage complex multi-pathing and PSA plugins

Knowledge

- Explain the Pluggable Storage Architecture (PSA) layout

Skills and Abilities

- Install and Configure PSA plug-ins
- Understand different multipathing policy functionalities
- Perform command line configuration of multipathing options
- Change a multipath policy
- Configure Software iSCSI port binding

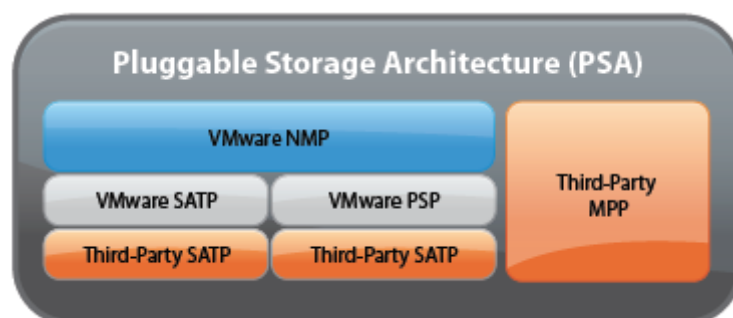
Tools & learning resources

- Product Documentation
 - [vSphere Command-Line Interface Installation and Scripting Guide](#)
 - [ESX Configuration Guide](#)
 - [ESXi Configuration Guide](#)
 - [Fibre Channel SAN Configuration Guide](#)
 - [iSCSI SAN Configuration Guide](#)
- vSphere Client
- vSphere CLI
 - esxcli
- VMware KB articles
 - [Understanding the storage path failover sequence in VMware ESX 4.x](#)
 - [VMworld 2009 - iSCSI Storage Performance Enhancement in vSphere 4](#)

This section overlaps with objectives 1.1 (Advanced storage management) and 1.2 (Storage capacity) but covers the multipathing functionality in more detail.

1.3.1 Understanding the PSA layout

The PSA layout is well documented [here](#), [here](#). The PSA architecture is for block level protocols (FC and iSCSI) - it isn't used for NFS.



Terminology;

- MPP = one or more SATP + one or more PSP
- NMP = native multipathing plugin
- SATP = traffic cop
- PSP = driver

There are four possible pathing policies;

- MRU = Most Recently Used. Typically used with active/passive (low end) arrays.
- Fixed = The path is fixed, with a 'preferred path'. On failover the alternative paths are used, but when the original path is restored it again becomes the active path.
- Fixed_AP = new to vSphere 4.1. This enhances the 'Fixed' pathing policy to make it applicable to active/passive arrays and ALUA capable arrays. If no user preferred path is set it will use its knowledge of optimised paths to set preferred paths.
- RR = Round Robin

One way to think of ALUA is as a form of 'auto negotiate'. The array communicates with the ESX host and lets it know the available path to use for each LUN, and in particular which is optimal. ALUA tends to be offered on midrange arrays which are typically asymmetric active/active rather than symmetric active/active (which tend to be even more expensive). Determining whether an array is 'true' active/active is not as simple as you might think! Read [Frank Denneman's excellent blogpost](#) on the subject. Our Netapp 3000 series arrays are asymmetric active/active rather than 'true' active/active.

1.3.2 Install and configure a PSA plugin

There are three possible scenarios where you need to install/configure a PSA plugin;

1. Third party MPP - add/configure a vendor supplied plugin
2. SATPs - configure the claim rules
3. PSPs - set the default PSP for a SATP

Installing/configuring a new MPP

Installing a third party MPP (to supplement the NMP) is either done through the command line (by using `esxupdate` or `vihostupdate` with a vendor supplied bundle) or Update Manager, depending on the vendor's support. Powerpath by EMC is one of the most well-known third party MPPs - instructions for installing it can be found in [VMwareKB1018740](#) or in this [whitepaper on RTFM](#).

After installing the new MPP you may need to configure claim rules which determine which MPP is used - the default NMP (including `MASK_PATH`) or your newly installed third party MPP.

To see which MPPs are currently used;

```
esxcli corestorage claimrule list
```

To change the owner of any given LUN (for example);

```
esxcli corestorage claimrule add --rule 110 --type location -A vmhba33 -C 0 -T 0 -L 4 -P <MPPname>
esxcli corestorage claimrule load
esxcli corestorage claiming unclaim --type location -A vmhba33 -C 0 -T 0 -L 4
esxcli corestorage claimrule run
```

To check the new settings;

```
esxcli corestorage claimrule list
```

Configuring an SATP

You can configure claim rules to determine which SATP claims which paths. This allows you to configure a new storage array from MyVendor (for example) so that any devices belonging to that storage array are automatically handled by a SATP of your choice.

To see a list of available SATPs and the default PSP for each;

```
esxcli nmp satp list (OR esxcfg-mpath -G to just see the SATPs)
```

To set criteria which determine how a SATP claims paths;

```
esxcli nmp satp listrules
```

```
esxcli nmp satp addrule --vendor="Netapp" --model="3240AE" --satp=VMW_SATP_ALUA
```

NOTE: This is different to the claimrules at the MPP level (esxcli corestorage claimrule)

Configuring the default PSP for a SATP

To change the default PSP;

```
esxcli nmp satp setdefaultpsp --satp VMW_SATP_ALUA --psp VMW_RR
```

NOTE: We use Netapp 3000 series arrays and they use the generic VMW_SATP_ALUA SATP. This has MRU as its default PSP even though RR is the 'best practice' recommendation for this storage array. If we change the default PSP however and then introduce a different array to our environment (which also uses the same SATP) it could get an unsuitable policy.

There is already a claimrule for the Vendor code NETAPP (esxcli nmp satp listrules).

[Duncan Epping's blogpost](#)

[Good post from DeinosCloud](#)

1.3.3 Administering path policy (GUI)

Viewing the multipathing properties (GUI)

From the VI client go to Configuration -> Storage, select the Datastore and then Properties -> Manage Paths. This will show the SATP and PSP in use;

The screenshot shows the vSphere Client interface. On the left, the 'Datastores' view is open, showing a table of datastores. The 'EG Test Datastore' is selected. Below the table, the 'Datastore Details' section shows the datastore's location, capacity, and usage. The 'Path Selection' dropdown is highlighted with a red circle and set to 'Most Recently Used'. The 'Properties' section shows the datastore name and volume label. On the right, the 'Ed's test LUN Manage Paths' dialog is open, showing the 'Path Selection' dropdown set to 'Most Recently Used (VMware)' and the 'Storage Array Type' set to 'VMW_SATP_ALUA'. The 'Paths' table shows three active paths for the LUN.

Runtime Name	Target	LUN	Status	Preferred
vmhba1:CO:T0:L60	50-0a:09:80:87:29:ef:26 50-0a:09:83:97:29:ef:26	60	Active (I/O)	
vmhba1:CO:T1:L60	50-0a:09:80:87:29:ef:26 50-0a:09:81:87:29:ef:26	60	Active	
vmhba2:CO:T0:L60	50-0a:09:80:87:29:ef:26 50-0a:09:83:87:29:ef:26	60	Active	
vmhba2:CO:T1:L60	50-0a:09:80:87:29:ef:26 50-0a:09:81:97:29:ef:26	60	Active	

Often you'll have access to vendor tools as well as the VI client but this doesn't always help – the Netapp plugin for example doesn't show the multipathing policy for a LUN although it does show whether ALUA is enabled or not;

Details

Storage Controller: **zcgprsan1n1** Partner: **zcgprsan1n2**

LUN	
Name:	naa.60a98000486e5874644a625058724d70
LUN Pathname:	/vol/v_test_lun_EG/lun
Serial Number:	HnXtdJbPXrMp
Status:	Online
Space Reservation:	Enabled
LUN Type:	vmware
Protocol:	fcp
IGroup:	VMWare-Chasis1_7-ALUA : 60 (Type: vmware)
Portset:	
ALUA Capable:	Enabled
Deduplication (Advanced Single Instance Storage)	
State:	Disabled
Status:	N/A
Type:	Disabled
Space Savings:	N/A (0.00B)
Last Start Time:	
Last End Time:	
Schedule:	
Space Shared:	0.00B

Capacity	
Datastore Usage (4%)	
LUN Usage (0%)	
Volume Usage (80%)	
Aggregate Usage (73%)	
Volume	
Name:	v_test_lun_EG
Status:	online
Type:	flex
Guarantee:	volume
Aggregate:	ag_tier1_02
Snapshot Reserve:	0%
Autogrow Increment:	Disabled
Autogrow Max Size:	Disabled
Snapshot Autodelete:	off, volume
Fractional Reserve:	100%
View Mapped Hosts...	

Changing the path policy (GUI)

1. Go to 'Manage Paths' and set the correct type.
2. Should be set consistently for every host with access to the LUN.
3. Use ESXTOP to monitor traffic on each HBA to ensure the paths are working as expected.

1.3.4 Administering path policy (CLI)

In general the CLI is more useful than the VI client because it's common to have multiple LUNs connected to multiple hosts (in an HA/DRS cluster for example). Rather than manually adjusting lots of paths the CLI can be scripted.

Changing a specific path policy (CLI)

Setting the path policy for a LUN, path or adaptor is very similar to viewing it;

```
[root@host ~]# esxcli nmp device setpolicy --device naa.xxx --psp VMW_PSP_RR true
[root@host ~]#
```

NOTE: esxcli doesn't work as you'd expect with vi-fastpass. You'll still need to specify `--server` as discussed on this [VMware community thread](#).

NOTE: With the release of vSphere 4.1.1 you can now set this using PowerCLI. See this [post from Arnim Van Lieshout](#) for details.

Viewing the multipathing policy (CLI)

To list all LUNs and see their multipathing policy;

- `esxcli nmp device list`
or from the vMA
`esxcli --server <myhost> ---username root --password <mypw> nmp device list`

```
naa.60a98000486e5874644a625170495054
Device Display Name: Ed's test LUN clone
```

```

Storage Array Type: VMW_SATP_ALUA
Storage Array Type Device Config:
{implicit_support=on;explicit_support=off;explicit_allow=on;alua_followover=on;{TPG_id=0,TPG_state=AO}{TPG_id=1,TPG_state=ANO}}
Path Selection Policy: VMW_PSP_MRU
Path Selection Policy Device Config: Current Path=vmhba1:C0:T0:L62
Working Paths: vmhba1:C0:T0:L62

naa.60a98000486e58756b34575976324868
Device Display Name: Systemfiles Tier 1
Storage Array Type: VMW_SATP_ALUA
Storage Array Type Device Config:
{implicit_support=on;explicit_support=off;explicit_allow=on;alua_followover=on;{TPG_id=2,TPG_state=AO}{TPG_id=3,TPG_state=ANO}}
Path Selection Policy: VMW_PSP_RR
Path Selection Policy Device Config: {policy=rr,iops=1000,bytes=10485760,useANO=0,lastPathIndex=1:NumIOsPending=0,numBytesPending=0}
Working Paths: vmhba2:C0:T0:L68, vmhba1:C0:T1:L68

```

NOTE: You get a similar output from 'vicfg-mpath -l' but that doesn't show the PSP in use or the working path.

[Good walkthrough from Jason Boche](#)

To check a particular datastore to ensure its multipathing is correct;

1. `vicfg-scsidevs --vmfs` (to get the naa ID for the datastore in question);

```

[root@zcgprvma01 ~][zcgprvsh45.mfl.co.uk]# vicfg-scsidevs -m
naa.60a98000486e5874644a625058724d70:1 /vmfs/devices/disks/naa.60a98000486e5874644a625058724d70:1 4d710389-2f1b6980-1322-001e0bee027e 0 EG Test Datastore
naa.60a98000486e5874644a5072466a772f:1 /vmfs/devices/disks/naa.60a98000486e5874644a5072466a772f:1 4a327518-83bb8fe8-6091-001e0beeb0f4 0 Another datastore

```

2. `esxcli nmp device list --device naa.60a98000486e5874644a625058724d70` (obtained from step 1);

```

[root@zcgprvma01 ~][zcgprvsh45.mfl.co.uk]# esxcli --server zcgprvsh45.mfl.co.uk --username root nmp device list -d
naa.60a98000486e5874644a625058724d70
naa.60a98000486e5874644a625058724d70
Device Display Name: Ed's test LUN
Storage Array Type: VMW_SATP_ALUA
Storage Array Type Device Config:
{implicit_support=on;explicit_support=off;explicit_allow=on;alua_followover=on;{TPG_id=1,TPG_state=ANO}{TPG_id=0,TPG_state=AO}}
Path Selection Policy: VMW_PSP_MRU
Path Selection Policy Device Config: Current Path=vmhba1:C0:T0:L60
Working Paths: vmhba1:C0:T0:L60

```

Configuring the Round Robin load balancing algorithm

When you configure Round Robin there are a set of default parameters;

- Use non-optimal paths: NO
- IOps per path: 1000 (ie swap to another path after 1000 IOps)

Some vendors recommend changing the IOps value to 1 instead of 1000 which you can do like so;


```
esxcli nmp roundrobin setconfig --device naa.xxx
```

Food for thought - an interesting [post from Duncan Epping about RR best practices](#).

Summary of useful commands

esxcli corestorage claimrule	Claimrules to determine which MPP gets used. Not to be confused with the <i>esxcli nmp satp</i> claimrules which determine which SATP (within a given MPP) is used. Typically used with MASK_PATH.
Esxcli nmp claimrule add	Add rules to determine which SATP (and PSP) is used for a given device/vendor/path
esxcli nmp device setpolicy	Configure the path policy for a given device, path, or adaptor
esxcli nmp satp setdefaultpsp	Configure the default pathing policy (such as MRU, RR or Fixed) for a SATP
esxcli nmp psp setconfig getconfig	Specific configuration parameters for each path policy. Typically used with RR algorithm (IOps, or bytes)
esxcfg-scsidevs --vmfs [--device naa.xxx]	Easy way to match a VMFS datastore with it's associated naa.xxx ID

1.3.5 Software iSCSI port binding

Port binding is the process which enables multipathing for iSCSI, a feature new to vSphere. By creating multiple vmKernel ports you can bind each to a separate pNIC and then associate them with the software iSCSI initiator, creating multiple paths. Check out a good post [here](#), more [here](#), plus the usual [Duncan Epping post](#). You can also watch this [video from VMware](#).

If you want to understand the theory behind iSCSI, check this [great multivendor post](#).

The iSCSI SAN Configuration Guide covers this in chapter 2 and the vSphere Command Line Interface Installation and Reference Guide briefly covers the syntax on page 90.

NOTE: This process can also be used for binding a hardware dependent iSCSI adaptor to its associated vmKernel ports. A hardware dependent iSCSI adaptor is simply a NIC with iSCSI offload capability as opposed to a fully-fledged iSCSI HBA.

Process summary;

1. Create multiple vmKernel ports
2. Configure networking so that each vmKernel has a dedicated pNIC
3. Bind each vmKernel port to the software iSCSI initiator
4. Rescan storage to recognise new paths

Create multiple vmKernel ports

Simply create an additional vmKernel port (with a different IP on the same subnet as the other vmKernel ports). You can use separate vSwitches (which enforces separate pNICs) or use a single vSwitch and then use portgroups with explicit failover settings This is done for both vSS and dVS switches.

Configure networking to dedicate a pNIC to each vmKernel port

You can use separate vSwitches (which enforces separate pNICs) or use a single vSwitch and then use portgroups with explicit failover settings. Both methods work for standard and distributed switches.

Bind each vmKernel port to the software iSCSI initiator

This is the bit that's unfamiliar to many, and it can only be done from the command line (locally or via RCLI/vMA);

- `esxcli swiscsi nic add -n vmk1 -d vmhba33`
- `esxcli swiscsi nic add -n vmk2 -d vmhba22`
- `esxcli swiscsi nic list` (to confirm configuration)

NOTE: These configuration settings are persistent across reboots, and even if you disable and re-enable the SW iSCSI initiator.

Rescan storage

Do a rescan (GUI or `esxcfg-rescan`) to see the new paths;

Before;

Storage Adapters

Device	Type	WWN
iSCSI Software Adapter		
vmhba33	iSCSI	iqn.1998-01.com.vmware:zcglabsv9-33465858
Virtual Machine Chipset		
vmhba1	Block SCSI	
vmhba32	Block SCSI	
LSI Logic Parallel SCSI Controller		
vmhba0	scst	

Details

vmhba33
Model: iSCSI Software Adapter
iSCSI Name: iqn.1998-01.com.vmware:zcglabsv9-33465858
iSCSI Alias:
Connected Targets: 2 Devices: 7 Paths: 7

View: Devices | Paths

Runtime Name	Target	LUN	Status
vmhba33:C0:T1:L0	iqn.1986-03.com.hp:storage.msa2012i.0828d5a237.b:192.168.215.32:3260	0	Active (I/O)
vmhba33:C0:T0:L0	iqn.1986-03.com.hp:storage.msa2012i.0828d5a237.a:192.168.215.31:3260	0	Active (I/O)
vmhba33:C0:T0:L1	iqn.1986-03.com.hp:storage.msa2012i.0828d5a237.a:192.168.215.31:3260	1	Active (I/O)
vmhba33:C0:T1:L2	iqn.1986-03.com.hp:storage.msa2012i.0828d5a237.b:192.168.215.32:3260	2	Active (I/O)
vmhba33:C0:T1:L3	iqn.1986-03.com.hp:storage.msa2012i.0828d5a237.b:192.168.215.32:3260	3	Active (I/O)
vmhba33:C0:T0:L4	iqn.1986-03.com.hp:storage.msa2012i.0828d5a237.a:192.168.215.31:3260	4	Active (I/O)
vmhba33:C0:T0:L5	iqn.1986-03.com.hp:storage.msa2012i.0828d5a237.a:192.168.215.31:3260	5	Active (I/O)

After;

Storage Adapters

Device	Type	WWN
iSCSI Software Adapter		
vmhba33	iSCSI	iqn.1998-01.com.vmware:zcglabsv7-59973f1f
Virtual Machine Chipset		
vmhba1	Block SCSI	
vmhba32	Block SCSI	
LSI Logic Parallel SCSI Controller		
vmhba0	scst	

Details

vmhba33
Model: iSCSI Software Adapter
iSCSI Name: iqn.1998-01.com.vmware:zcglabsv7-59973f1f
iSCSI Alias:
Connected Targets: 4 Devices: 7 Paths: 14

View: Devices | Paths

Runtime Name	Target	LUN	Status
vmhba33:C1:T1:L0	iqn.1986-03.com.hp:storage.msa2012i.0828d5a237.a:192.168.215.31:3260	0	Active (I/O)
vmhba33:C1:T1:L1	iqn.1986-03.com.hp:storage.msa2012i.0828d5a237.a:192.168.215.31:3260	1	Active (I/O)
vmhba33:C1:T1:L4	iqn.1986-03.com.hp:storage.msa2012i.0828d5a237.a:192.168.215.31:3260	4	Active (I/O)
vmhba33:C1:T1:L5	iqn.1986-03.com.hp:storage.msa2012i.0828d5a237.a:192.168.215.31:3260	5	Active (I/O)
vmhba33:C0:T1:L0	iqn.1986-03.com.hp:storage.msa2012i.0828d5a237.a:192.168.215.31:3260	0	Active
vmhba33:C0:T1:L1	iqn.1986-03.com.hp:storage.msa2012i.0828d5a237.a:192.168.215.31:3260	1	Active
vmhba33:C0:T1:L4	iqn.1986-03.com.hp:storage.msa2012i.0828d5a237.a:192.168.215.31:3260	4	Active
vmhba33:C0:T1:L5	iqn.1986-03.com.hp:storage.msa2012i.0828d5a237.a:192.168.215.31:3260	5	Active
vmhba33:C1:T0:L0	iqn.1986-03.com.hp:storage.msa2012i.0828d5a237.b:192.168.215.32:3260	0	Active (I/O)
vmhba33:C1:T0:L2	iqn.1986-03.com.hp:storage.msa2012i.0828d5a237.b:192.168.215.32:3260	2	Active (I/O)
vmhba33:C1:T0:L3	iqn.1986-03.com.hp:storage.msa2012i.0828d5a237.b:192.168.215.32:3260	3	Active (I/O)
vmhba33:C0:T0:L0	iqn.1986-03.com.hp:storage.msa2012i.0828d5a237.b:192.168.215.32:3260	0	Active
vmhba33:C0:T0:L2	iqn.1986-03.com.hp:storage.msa2012i.0828d5a237.b:192.168.215.32:3260	2	Active
vmhba33:C0:T0:L3	iqn.1986-03.com.hp:storage.msa2012i.0828d5a237.b:192.168.215.32:3260	3	Active

2 Network

2.1 Implement and Manage Complex Virtual Networks

Knowledge

- Identify common virtual switch configurations

Skills and Abilities

- Determine use cases for and apply IPv6
- Configure NetQueue
- Configure SNMP
- Determine use cases for and apply VMware DirectPath I/O
- Migrate a vSS network to a Hybrid or Full vDS solution
- Configure vSS and vDS settings using command line tools
- Analyze command line output to identify vSS and vDS configuration details

Tools & learning resources

- Product Documentation
 - [vSphere Command-Line Interface Installation and Scripting Guide](#)
 - [vNetwork Distributed Switch: Migration and Configuration](#)
 - [ESX Configuration Guide](#)
 - [ESXi Configuration Guide](#)
- vSphere Client
- vSphere CLI
 - vicfg-*
- [TA2525 - vSphere Networking Deep Dive](#) (VMworld 2009 - free access)
- [TA6862 - vDS Deep Dive - Managing and Troubleshooting](#) (VMworld 2010)
- [TA8595 - Virtual Networking Concepts and Best Practices](#) (VMworld 2010)
- [Design considerations for the vDS \(Rich Brambley\)](#)
- [Catch-22 for vds and vCentre \(Jason Boche\)](#)
- TrainSignal – vSphere Pro Series 1 Nexus 1000v lessons
- [Kendrick Coleman's 'How to setup the Nexus 1000V in ten minutes' blogpost](#)

The VCAP-DCA lab is still v4.0 (rather than v4.1) which means features such as NIOC and load based teaming (LBT) aren't covered. Even though the Nexus 1000V isn't on the Network objectives blueprint (just the vDS) it's worth knowing what extra features it offers as some goals include knowing when to use the Nexus1000V or just the vDS.

2.1.1 Network basics (VCP revision)

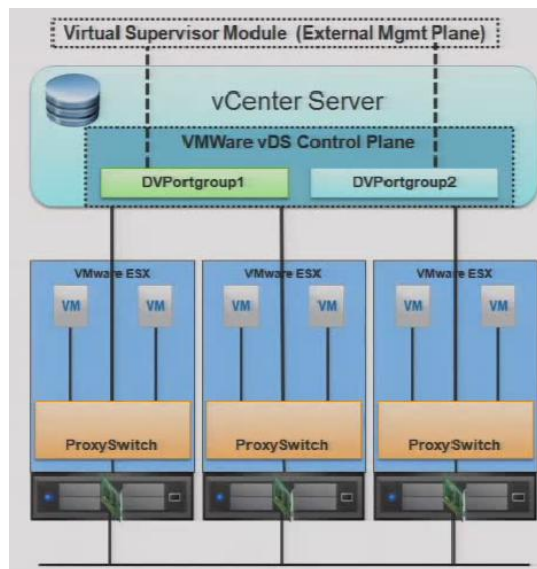
Standard switches support the following features;

- NIC teaming
 - Based on source VM ID (default)
 - Based on IP Hash (used with Etherchannel)
 - Based on source MAC hash
 - Explicit failover order
- VLANs (EST, VST, VGT)

vDS Revision

The vDistributed switch separates the control plane and the data plane to enable centralised administration as well as extra functionality compared to standard vSwitches. A good summary can be found at [GeekSilver's blog](#). Benefits;

- Offers both inbound and outbound traffic shaping (standard switches only offer outbound)
 - Traffic shaping can be applied at both dvPortGroup and dvUplink PortGroup level
 - For dvUplink PortGroups ingress is traffic from external network coming into vDS, egress is traffic from vDS to external network
 - For dvPortGroups ingress is traffic from VM coming into vDS, egress is traffic from vDS to VMs
 - Configured via three policies - average bandwidth, burst rate, and peak bandwidth
- Ability to build a third party vDS on top (Cisco Nexus 1000v)
- Traffic statistics are available (unlike standard vSwitches)



NOTES:

- CDP and MTU are set per vDS (as they are with standard vSwitches).
- PVLANS are defined at switch level and applied at dvPortGroup level.
- There is one DVUplink Portgroup per vDS
- NIC teaming is configured at the dvPortGroup level but can be overridden at the dvPort level (by default this is disabled but it can be allowed). This applies to both dvUplink Portgroups and standard dvPortGroups although on an uplink you CANNOT override the NIC teaming or Security policies.
- Policy inheritance (lower level takes precedence but override is disabled by default)
 - dvPortGroup -> dvPort
 - dvUplink PortGroup -> dvUplinkPort

NOTE: Don't create a vDS with special characters in the name (I used 'Lab & Management') as it breaks host profiles - see [VMwareKB1034327](#).

2.1.2 Determine use cases for and apply IPv6

The use case for IPv6 is largely due to IPv4 running out of address space - it isn't so much a VMware requirements as an Internet requirement. IPv6 is 'supported' on ESX/i, but there are a few features in vSphere which aren't compatible;

- ESX during installation - you have to install on an IPv4 network
- VMware HA
- VMware Fault Tolerance
- RCLI

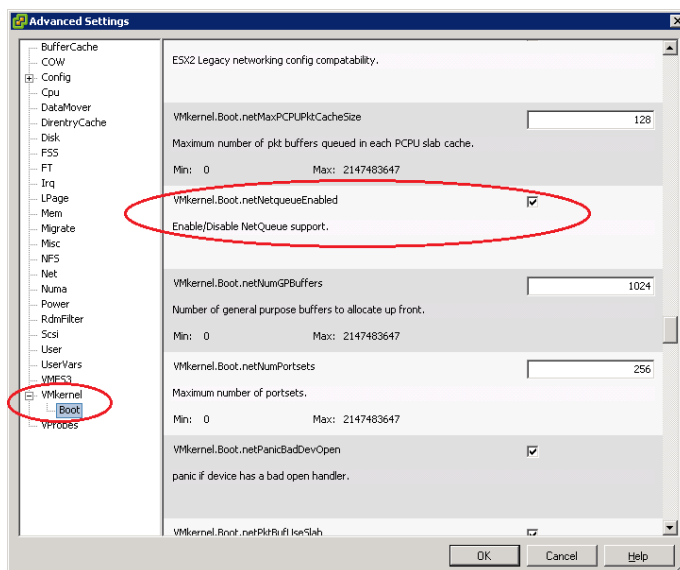
Enabling IPv6 is easily done via the VI client, Configuration -> Networking, click Properties and reboot host. You can enable IPv6 without actually configuring any interfaces (SC, vmKernel etc) with an IPv6 address.

See [VMwareKB1010812](#) for details of using command line to enable IPv6 and read this [blogpost by Eric Siebert on IPv6 support](#).

2.1.3 Netqueue

NetQueue (which was present in ESX 3.5 but is improved in vSphere) is a feature which improves network performance when sending or receiving large amounts of traffic to an ESX host, typically used with 10GB Ethernet. Without Netqueue it is normally impossible to achieve full 10GB throughput (read more in this [Dell whitepaper](#)). It does this by processing multiple queues in parallel, using multiple CPUs.

- Enabled by default
- Requires support from the pNIC
- More beneficial with NUMA architectures
- Enabled/disabled by;
 - Via host's Configuration -> Advanced Settings > vmKernel
 - `esxcfg-advcfg --set-kernel 1 netNetqueueEnabled` (to enable)
 - `esxcfg-advcfg --set-kernel 0 netNetqueueEnabled` (to disable)



As well as enabling the host functionality you may need to configure the NIC driver with vendor specific settings, typically using esxcfg-module. See [VMwareKB1004278](#) for details of enabling Netqueue for a specific 10GB NIC. In the real world it seems as if Netqueue performance is dependent on good driver support – as this [article at AnandTech](#) points out some drivers aren't much of an improvement over 1GB.

[TA2525 - vSphere Networking Deep Dive](#) covers this in detail around the 1hr 15 mark.

2.1.4 Configure SNMP

SNMP can be used to enhance management, typically by providing information either on request (polling) or when events are triggered (trap notification). Sending an SNMP trap is one of the standard alarm actions in vCenter.

Two network ports used;

- 161/udp - used to receive poll requests
- 162/udp - used to send trap notifications

Configuring SNMP for vCenter

- Go to Administration -> vCenter Settings -> SNMP
- vCenter can only send notification traps, it doesn't support poll requests.

Configuring SNMP for ESX/i hosts

- Use vicfg-snmp (RCLI) or directly edit configuration files on the ESX/i hosts
NOTE: There is NO esxcfg-snmp, and there is no GUI option for configuring hosts.

To check the current configuration;

```
[vi-admin@zcglabvma01 ~][zcglabsvr7.lab.co.uk]$ vicfg-snmp --show
Current SNMP agent settings:
Enabled : 0
UDP port : 161

Communities :

Notification targets :
```

```
vicfg-snmp --targets <SNMP Receiver>/community
vicfg-snmp --test
```

Misc;

- By default SNMP is disabled with no targets defined
- ESX has both the Net-SNMP agent and a VMware hostd agent. ESXi only has the VMware agent. VMware specific information is only available from the embedded (hostd) agent.
- Using vicfg-snmp configures the VMware SNMP agent (ie it modifies /etc/vmware/snmp.xml). It does NOT configure the Net-SNMP agent.
- When you configure SNMP using vicfg-snmp the relevant network ports are automatically opened on the firewall. If you edit the config files directly you'll need to do this yourself (esxcfg-firewall -o 162,udp,outgoing,snmpd for notification traps).

See [VMwareKB1022879](#) for details of editing configuration files or watch this video to see how to use vicfg-snmp - [Eric Sloof's How to configure SNMP](#). The [vSphere Command-Line Interface Installation and Scripting Guide](#) p.40 covers vicfg-snmp while the [Basic System Administration guide](#) covers it in more depth (page 50-65).

Configuring the SNMP management server

You should load the VMware MIBs on the server receiving the traps you've configured. These MIBs can be downloaded from VMware on the vSphere download page. Follow instructions for your particular product to load the MIBs.

2.1.5 Determine use cases for and apply VMware DirectPath I/O

This was covered in section 1.1. Obviously it can be used with 10GB NICs if high network throughput is required.

2.1.6 Migrate a vSS network to a hybrid of full vDS solution

Make sure you read the [VMware whitepaper](#) and experiment with migrating a vSS to a vDS. There is some discussion about whether to use hybrid solutions (both vSS and vDS) or not - see this [blogpost by Duncan Epping](#) for some background information.

Determining the best deployment method for a dvS - see [VMware white paper](#) but roughly;

- If new hosts or no running VMs - use host profiles
- If migrating existing hosts or running VMs - use dvS GUI and 'Migrate VM Networking'.

There are some catch-22 situations with a vDS which you may run into when migrating from a vSS;

- When vCenter is virtual - if you lose vCenter then you can't manage the vDS so you can't get a new vCenter on the network. See [Jason Boche's thoughts on the subject](#).
- If an ESXi host loses its management network connection you may not be able to reconfigure it using the command line (esxcfg-vswitch is limited with vDS operations). An alternative solution is to 'Restore Standard Switch' from the DCUI.
NOTE: The option above doesn't configure a VLAN tag so if you're using VLANs you'll need to reconfigure the management network after resetting to a standard switch.
- If you limited pNICs you may also run into problems like [Joep Piscaer's blogpost](#)

To check: how can you migrate templates? I think it errors if you try....

2.1.7 Command line configuration for vSS and vDS

Commands for configuring vSS

- esxcfg-nics -l
- esxcfg-vswitch
- esxcfg-vmknics
- esxcfg-vswif
- esxcfg-route

Typical command lines;

```
esxcfg-vswitch -a vSwitch2
```

```
esxcfg-vswitch -L vmnic2 vSwitch2
```

```
esxcfg-vswitch -A NewServiceConsole vSwitch2
```

```
esxcfg-vswif -a -i 192.168.0.10 -n 255.255.255.0 vswif2
```

NOTE: The above commands create a new standard vSwitch, portgroup and Service Console – typically used to recover from some vDS scenarios.

Commands for configuring a vDS

There are very few commands for configuring a vDS and what there is I've covered in section 2.4 which is dedicated to the vDS. See [VMwareKB1008127](#) (configuring vDS from the command line) and it's worth watching the following session from VMworld 2010 (although you'll need a current subscription) - [session TA6862 vDS Deep dive - Management and Troubleshooting](#).

2.1.8 Analyze command line output to identify vSS and vDS configuration details

Things to look out for above and beyond the basic vSwitch, uplink and portgroup information;

- MTU
- CDP
- VLAN configuration

See section 6.3, Troubleshooting Network Connectivity for more information.

2.2 Configure and maintain VLANs, PVLANS and VLAN settings

Knowledge

- Identify types of VLANs and PVLANS

Skills and Abilities

- Determine use cases for and configure VLAN Trunking
- Determine use cases for and configure PVLANS
- Use command line tools to troubleshoot and identify VLAN configurations

Tools & learning resources

- Product Documentation
 - [vSphere Command-Line Interface Installation and Scripting Guide](#)
 - [ESX Configuration Guide](#)
 - [ESXi Configuration Guide](#)
- vSphere Client
- vSphere CLI
 - vicfg-*
- [TA2525 vSphere Networking Deep Dive](#) (VMworld 2009)
- [Eric Sloof's video on configuring PVLANS and dvSwitches](#)
- [Carlos Vargas's video on VLAN configuration](#)

This is one of the smaller objectives plus only the PVLAN concepts and practices are new - VLAN support remains relatively unchanged from VI3 (although the vDS and its associated VLAN support is new).

2.2.1 Types of VLAN

VLANs are a network standard (802.1q) which are fully supported in vSphere. They can be used to minimise broadcast traffic and as a security measure to segregate traffic (although like any technology [there are weaknesses](#)). Typical uses for VLANs with vSphere are to isolate infrastructure (vMotion, iSCSI and NFS) traffic and VM traffic.

There are three main ways of using VLANs with vSphere ([covered in this VMware whitepaper](#));

- Virtual guest tagging (VGT) - requires VLAN driver support in the guest OS
- Virtual Switch tagging (VST) - common option, requires VLAN trunking on external switches
- External switch tagging (EST) - less flexible and requires more physical NICs

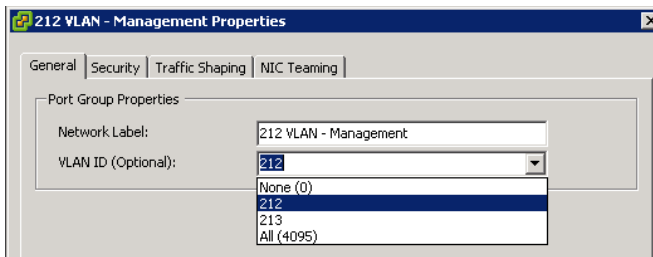
In the Cisco world you set a port to be an 'access port' or a 'trunk port' if it's going to carry multiple VLANs. VLAN IDs are 16 bit values giving a range of 0-4095. 4095 is used within vSphere to mean 'all VLANs' and is how you configure a portgroup when using VGT.

2.2.2 Configuring VLANs and VLAN trunking

For standard vSwitches you configure VLAN tags on portgroups. This configuration is done at the ESX host using the VI client (Configuration -> Networking);

- Use VLAN 0 when no VLAN tags are present (EST)
- Use VLAN 4095 to pass all VLANs (VGT)

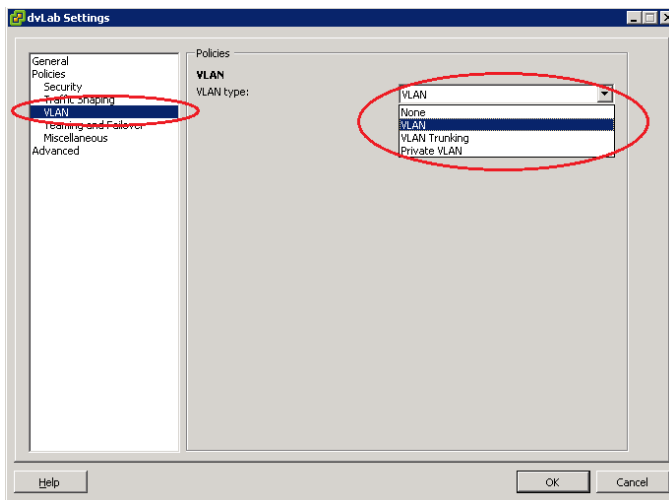
- Use a specific VLAN ID depending on the isolation required (VST)



NOTE: [Avoid using VLAN1](#) (native VLAN for most Cisco kit) as this can inadvertently expose traffic you may not mean to expose.

For distributed switches you configure VLANs on both dvPortGroups and dvUplinkPortGroups (with the option to override at the dvPort level when enabled). This is done in vCenter using the VI client;

- Use 'None' for EST
- Use 'VLAN' (and specify a VLAN ID) depending on the isolation requirement (VST)
- Use 'VLAN trunking' to pass either all VLANs (VGT) or a selection of VLANs (VST). This is an improvement over standard switches which either set a single VLAN ID or All. Restricting the VLANs this way is a form of [VLAN pruning](#).
- Use PVLANS when you need a subset of hosts within a single VLAN (see next section)



2.2.3 Types of PVLAN (Private VLANs)

From the ESXi Configuration Guide - "PVLANS are used to solve VLAN ID limitations". They allow more fine grained control over subsets of hosts without requiring a dedicated VLAN for each group, cutting down on network administration (here's [a good explanation and diagram](#)). [Eric Sloof's video on configuring PVLANS and dvSwitches](#) is also worth a watch (from 24mins for the PVLAN part).

Think of PVLANS as a VLAN within a VLAN (read [VMwareKB1010691](#) PVLAN concepts);

- Promiscuous VLAN - this is an extension of the original (parent) VLAN
- Secondary VLANs - there are two choices;
 - Isolated (one per primary VLAN)
 - Community (multiple per primary VLAN)

Requirements

- PVLANS are only available on vDistributed Switches
- The physical switches must support PVLANS and be configured with the VLAN IDs used for secondary VLANs.

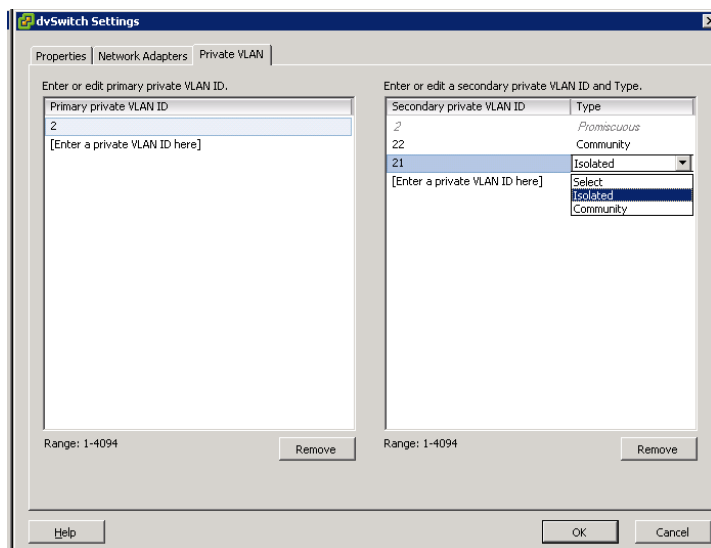
NOTE: To test PVLANS in a lab environment you could run multiple virtual ESX hosts on a single host (along with a virtual router such as [Vyatta Core](#)). This was the traffic never reaches the physical network so you don't need a PVLAN capable switch.

[VMwareKB1010691](#) offers a good overview of PVLAN concepts when used with vDS

2.2.4 Configuring PVLANS

PVLANS

- Configured via vCentre as a property of the vDS itself
- CANNOT be done from the command line
- MUST also be configured on physical switches



Read [VMwareKB1010703](#) (PVLAN implementation on a vDS) or the ESXi Configuration Guide page 32 onwards.

NOTE: Trying to remove PVLANS from the vDS when a dvPortGroup is still using the PVLANS will result in an error and no deletion occurring. Check the various dvPortGroups and remove the config before removing the PVLANS from the vDS.

2.2.5 Command line tools for VLAN configuration/troubleshooting

The usual commands support VLANs, typically using the -v parameter;

- vicfg-vswitch - how to use -v for vlan assignment. Use -v 0 to clear.
- vicfg-vswif
- vicfg-vmknic
- vicfg-route

NOTE: You can only administer VLANs at the command line - PVLANS are only configured in vCenter.

2.3 Deploy and maintain scalable virtual networking

Knowledge

- Identify VMware NIC Teaming policies
- Identify common network protocols

Skills and Abilities

- Understand the NIC Teaming failover types and related physical network settings
- Determine and apply Failover settings
- Configure explicit failover to conform with VMware best practices
- Configure port groups to properly isolate network traffic

Tools & learning resources

- Product Documentation
 - [ESX Configuration Guide](#)
 - [ESXi Configuration Guide](#)
 - [vSphere Command-Line Interface Installation and Scripting Guide](#)
- vSphere Client
- vSphere CLI
 - `vicfg-*`
- [Frank Denneman's blogpost on IP Hash vs LBT](#)

2.3.1 Identify, understand , and configure NIC teaming

The five available policies are;

- Route based on virtual port ID (default)
- Route based on IP Hash (MUST be used with static Etherchannel - no LACP). No beacon probing.
- Route based on source MAC address
- Route based on physical NIC load (vSphere 4.1 only)
- Explicit failover

NOTE: These only affect outbound traffic. Inbound load balancing is controlled by the physical switch.

Failover types and related physical network settings

Failover types

- Cable pull/failure
- Switch failure
- Upstream switch failure

Change NIC teaming for FT logging (use IP hash) - [VMwareKB1011966](#)

Use uplink failure detection (also known as [link state tracking](#)) to handle physical network failures outside direct visibility of the host.

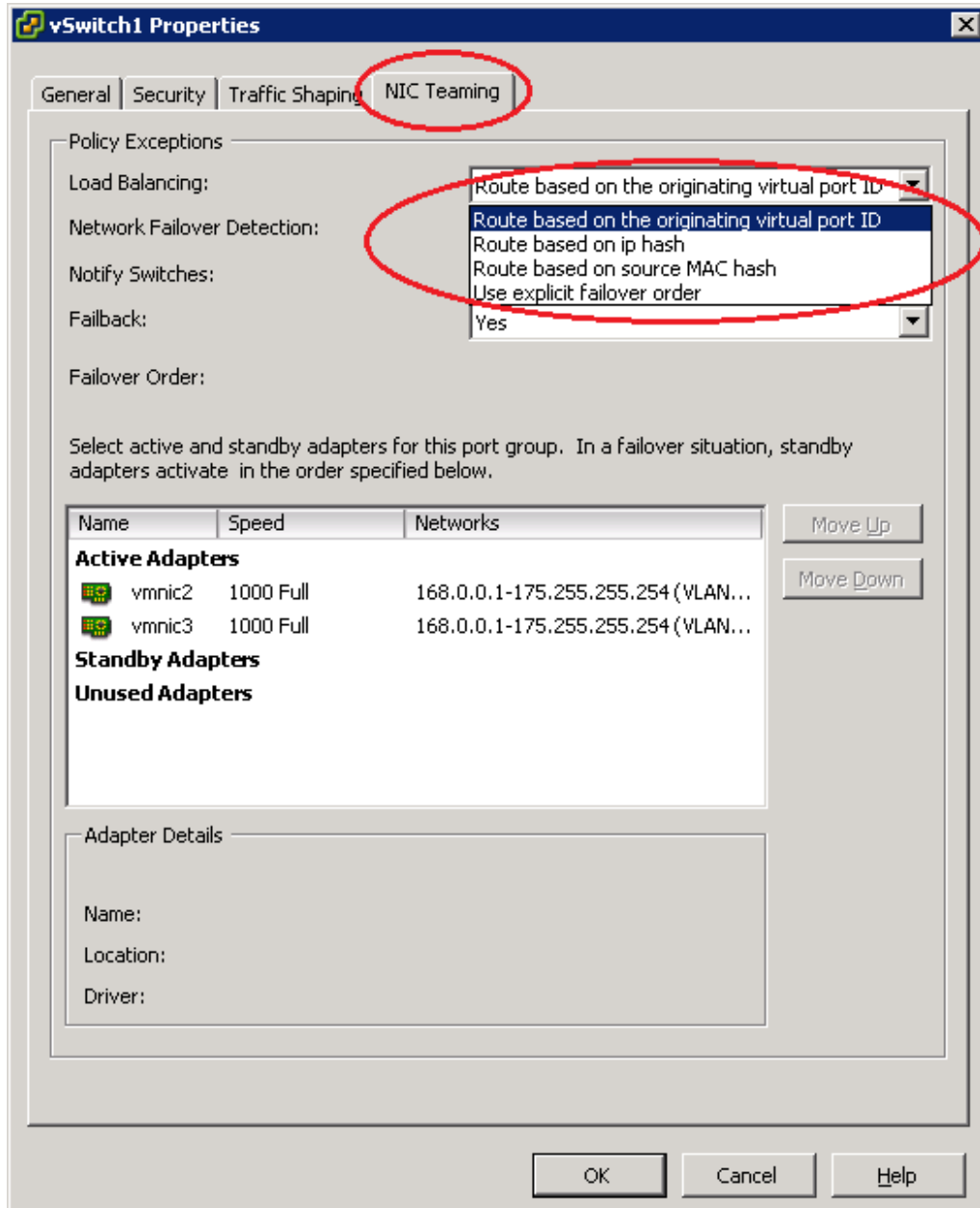
With blades you typically don't use NIC teaming as each blade has a 1 to 1 mapping from its multiple pNIC to the blade chassis switch. That switch in turn may use an Etherchannel to an upstream switch

but from the blade (and hence ESX perspective) it simply has multiple independent NICs (hence route on virtual port ID is the right choice).

Configuring failover settings

Failover settings can be configured at various levels on both standard and distributed switches;

- vSS
 - vSwitch, then port group
- vDS
 - dvPortGroup then dvPort
 - dvUplinkPortGroup (NOTE: you can't override at the dvUplinkPort level)



Explicit failover can be used to balance bandwidth while still providing resilience with minimal numbers of pNICs. If you only have two pNICs available;

- Configure a single vSwitch and add both pNICs
- Configure two portgroups with explicit failover orders;

- Configure the management traffic portgroup to use pNIC1 as active with pNIC2 as standby.
- Configure the VM network portgroup to use pNIC1 as standby and pNIC2 as active.

This achieves both separation of traffic over separate pNICs for optimal bandwidth as well as providing resilience to both portgroups. [VMwareKB1002722](#) describes this in more detail.

You can [configure NIC teaming using the CLI](#) (although this procedure isn't covered in the standard documentation so won't be available during the VCAP-DCA exam).

NOTE: With the vDS you get a diagram showing the actual path traffic takes through the switch. You can also confirm the actual NICs used (and therefore whether your teaming is working as expected using esxtop. More on this in section 6.3 Troubleshooting Network Connectivity.

2.3.2 Identify common network protocols

This has been covered elsewhere (in section 7.2 on the ESX firewall) and should be common knowledge. A few protocols which aren't so common but are supported in vSphere;

- CDP
- NTP (UDP port 123)

2.3.3 Isolation best practices

The following are generally accepted best practices (don't let [Tom Howarth](#) hear you say that);

- Separate VM traffic and infrastructure traffic (vMotion, NFS, iSCSI)
- Separate pNICs and vSwitches
- VLANs can be used to isolate traffic
- When using NIC teams use pNICs from separate buses (ie don't have a team comprising two pNICs on the same PCI card - use one onboard adapter and one from an expansion card)
- Keep FT logging on a separate pNIC and vSwitch
- Use dedicated networks for storage (iSCSI and NFS)

When you move to 10GB networks isolation is implemented differently (often using some sort of IO virtualisation like FlexConnect, Xsigo, or UCS) but the principals are the same. [VMworld 2010 session TA8440](#) covers the move to 10GB and FCoE.

2.4 Administer vNetwork Distributed Switch Settings

Knowledge

- Explain relationship between vDS and logical vSSes

Skills and Abilities

- Understand the use of command line tools to configure appropriate vDS settings on an ESX/ESXi host
- Determine use cases for and apply Port Binding settings
- Configure Live Port Moving
- Given a set of network requirements, identify the appropriate distributed switch technology to use
- Use command line tools to troubleshoot and identify configuration items from an existing vDS

Tools & learning resources

- Product Documentation
 - [ESX Configuration Guide](#)
 - [ESXi Configuration Guide](#)
 - [vSphere Command-Line Interface Installation and Scripting Guide](#)
- vSphere Client
- vSphere CLI
 - vicfg-*
- [TA2525 - vSphere Networking Deep Dive](#) (VMworld 2009 - free access)

2.4.1 Relationship between vSS and vDS

Both standard (vSS) and distributed (vDS) switches can exist at the same time - indeed there's good reason to use this 'hybrid' mode.

UPDATE: [A new post \(April 2011\) from Duncan Epping about 'full' dvS](#)

You can view the switch configuration on a host (both vSS and dvS) using `esxcfg-vswitch -l`. It won't show the 'hidden' switches used under the hood by the vDS although you can read more about those in this [useful article at RTFM](#) or at [Geeksilver's blog](#).

2.4.2 Command line configuration of a vDS

The command line is pretty limited when it comes to vDS. Useful commands;

- `esxcfg-vswitch`
 - `esxcfg-vswitch -P vmnic0 -V 101 <dvSwitch>` (link a physical NIC to a vDS)
 - `esxcfg-vswitch -Q vmnic0 -V 101 <dvSwitch>` (unlink a physical NIC from a vDS)
- `esxcfg-vswif -l | -d` (list or delete a service console)
- `esxcfg-nics`
- `net-dvs`

NOTE: `net-dvs` can be used for diagnostics although it's an unsupported command. It's located in `/usr/lib/vmware/bin`. Use of this command is covered in section 6.4 Troubleshooting Network connectivity.

NOTE: esxcfg-vswitch can ONLY be used to link and unlink physical adaptors from a vDS. Use this to fix faulty network configurations. If necessary create a vSS switch and move your physical uplinks across to get your host back on the network. See [VMwareKB1008127](#) or this [blogpost](#) for details.

Identify configuration items from an existing vDS

You can use esxcfg-vswitch -l to show the dvPort assigned to a given pNIC and dvPortGroup.

See the Troubleshooting Network connectivity section for more details.

2.4.3 Port Binding settings

With standard vSwitches all port bindings are 'ephemeral', meaning the port is created when the VM's powered on and deleted when the VM is powered off (or vMotioned to another host). With distributed switches there are now three types of port binding;

- Static
 - Default binding method for a dvPortGroup
 - Assigned to a VM when it's added to the dvPortGroup
 - Conceptually like a static IP address
 - Port assignment persists to the VM across reboots, vMotions etc
- Dynamic
 - Used when you approach port limits (either on the particular dvPortGroup or on the vDS itself which has a maximum of 6000 dvPorts). If you have 10,000 VMs you only allocate a dvPort to *powered on* VMs
 - Conceptually like DHCP for a pool of desktops
 - dvPort assignment can change when VM is powered off. vCenter will attempt to use the same dvPort but no guarantee.
 - *LIMITATION: Not all VMs can be powered on at the same time if you have more than 6000.*
 - *LIMITATION: vCenter must be available when powering on the VM, as it needs to assign a dvPort.*
- Ephemeral
 - Port binding does NOT persist.
 - Number of VMs can exceed the number of ports on a given dvPortGroup (but are still bound by the total number of dvPorts on a vDS)
 - Equivalent to standard vSwitch behaviour
 - You can power on a VM using either vCenter or the VI client connected directly to a host.
 - Typically used in emergency or recovery situations. You could create an ephemeral portgroup to be used with a virtual vCenter for instance.

[TA2525 - vSphere Networking Deep Dive](#) explains the port binding quite clearly (around the 30 minute mark) and [the Trainsignal Troubleshooting course](#) has a video dedicated to explaining port binding.

NOTE: vSphere 4.1 has increased the port [maximums](#) (there are now up to 20,000 dvPorts per vDS). There is some guidance in [VMwareKB1022312](#) on choosing the port binding type.

Configuring port bindings

- Port Binding are configured in vCenter
- Configured per dvPortGroup (can't be overridden on an individual dvPort)
- Must be configured before assigning VMs to the dvPortGroup.
- No command line configuration is available.

2.4.4 Configure Live Port Moving

Despite being on the blueprint there is very little information about what this actually is. The ESX Configuration Guide has a token mention (on page 35) where it refers to 'allowing live port migration' as a property on a vDS but I couldn't find the option (and I'm [not the only one](#)).

There is a [post on the VMware communities site](#) explaining a bit about it - let's just hope it's not tested!

2.4.5 Identify the appropriate distributed switch technology to use

This could mean knowing when to use the basic vDS or the Nexus 1000V - have a read of [Comparing vSS, dvS and Nexus 1000V white paper](#). Alternatively it could mean knowing then the extra features available with a vDS (compared to a vSS) are needed;

- to simplify network maintenance in larger environment (less configuration)
- delegation to a network team (relevant to Nexus 1000V)
- when Enterprise+ licencing is available!
- when you need PVLANs (isolation of hosts within a single VLAN for example)
- when you need network vMotion - VMSafe, vShield product suite etc

2.4.6 Use command line tools to troubleshoot an existing vDS

See the troubleshooting section 6.3 for details.

3 Deploy DRS Clusters and Manage Performance

3.1 Tune and Optimize vSphere Performance

Knowledge

- Identify appropriate BIOS and firmware setting requirements for optimal ESX/ESXi Host performance
- Identify appropriate ESX driver revisions required for optimal ESX/ESXi Host performance
- Recall where to locate information resources to verify compliance with VMware and third party vendor best practices

Skills and Abilities

- Tune ESX/ESXi Host and Virtual Machine memory configurations
- Tune ESX/ESXi Host and Virtual Machine networking configurations
- Tune ESX/ESXi Host and Virtual Machine CPU configurations
- Tune ESX/ESXi Host and Virtual Machine storage configurations
- Configure and apply advanced ESX/ESXi Host attributes
- Configure and apply advanced Virtual Machine attributes
- Tune and optimize NUMA controls

Tools & learning resources

- Product Documentation
 - [vSphere Resource Management Guide](#)
 - [vSphere Command-Line Interface Installation and Scripting Guide](#)
 - [Performance Troubleshooting for VMware vSphere 4](#)
- vSphere Client
 - Performance Graphs
- vSphere CLI
 - vicfg-*,resxtop/esxtop, vscsiStats
- [VMworld 2010 session TA7750 Understanding Virtualisation Memory Management](#) (subscription required)
- [VMworld 2010 session TA7171 - Performance Best Practices for vSphere](#) (subscription required)
- [VMworld 2010 session TA8129 - Beginners guide to performance management on vSphere](#) (subscription required)
- [Performance Troubleshooting in Virtual Infrastructure](#) (TA3324, VMworld '09)
- [Scott Sauer's blogpost on storage performance](#)
- [VMware's Performance Best Practices white paper](#)

It's hard to know what to cover in this objective as performance tuning often implies troubleshooting (note the recommended reading of Performance Troubleshooting!) hence there's a significant overlap with the troubleshooting section. Luckily there are plenty of excellent resources in the blogosphere and from VMware so it's just a case of reading and practicing.

3.1.1 Identify BIOS and firmware settings for optimal performance

This will vary for each vendor but typical things to check;

- Power saving for the CPU.
- Hyperthreading - should be enabled

- Hardware virtualisation (Intel VT, EPT etc) - required for EVC, Fault Tolerance etc
NOTE: You should also [enable the 'No Execute' memory protection bit](#).
- NUMA settings (node interleaving for DL385 for instance. Normally disabled - check [Frank Denneman's post](#).
- WOL for NIC cards (used with DPM)

3.1.2 Identify appropriate ESX driver revisions required for optimal host performance

I guess they mean the HCL. Let's hope you don't need an encyclopaedic knowledge of driver version histories!

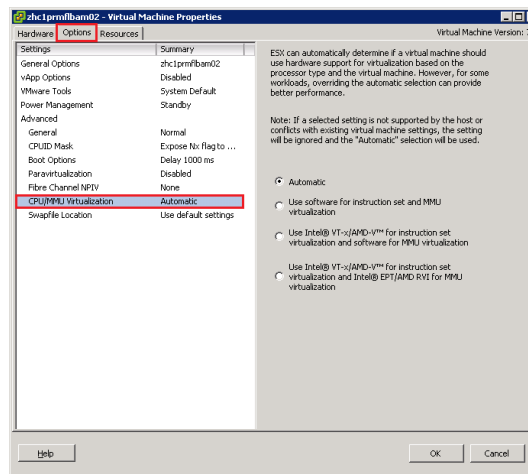
3.1.3 Tune ESX/i host and VM memory configurations

Read this great series of blog posts from Arnim Van Lieshout on memory management - part [one](#), [two](#) and [three](#). And as always the [Frank Denneman post](#).

Check your Service Console memory usage using esxtop.

Hardware assisted memory virtualisation

Check this is enabled (per VM). Edit Settings -> Options -> CPU/MMU Virtualisation;



NOTE: VMware strongly recommend you use large pages in conjunction with hardware assisted memory virtualisation. See section 3.2 for details on enabling large memory pages. However enabling large memory pages will negate the efficiency of TPS so you gain performance at the cost of higher memory usage. Pick your poison...(and read this [interesting thread on the VMware forums](#))

Preference for memory overcommit storage performance (most effective at the top);

1. Transparent page sharing (negligible performance impact)
2. Ballooning
3. Memory compression
4. VMkernel swap files (significant performance impact)

Transparent Page Sharing (TPS) - otherwise known as dedupe!

- Enabled by default
- Refreshed periodically
- Can be disabled;

- Disable per ESX - add Mem.ShareScanGHz = 0 in Advanced Settings ([VMwareKB1004901](#))
- Disable per VM - add sched.mem.pshare.enable = FALSE in .VMX (entry not present by default).
- Efficiency is impacted if you enable large memory pages (see this [discussion](#))

Balloon driver

- Uses a guest OS driver (vmmemctl) which is installed with VMware Tools (all supported OSs)
- Guest OS must have enough swapfile configured for balloon driver to work effectively
- Default max for balloon driver to reclaim is 65%. Can be tuned using sched.mem.maxmemctl in .VMX (entry not present by default). Read this [blogpost](#) before considering disabling!
- Ballooning is normal when overcommitting memory and *may* impact performance

Swapfiles

- VMware swapfiles
 - Stored (by default) in same datastore as VM (as a .vswp file). Size = configured memory - memory reservation.
 - Include in storage capacity sizing
 - Can be configured to use local datastore but that can impact vMotion performance. Configured at either cluster/host level or override per VM (Edit Settings -> Options - Swapfile location)
 - Will almost certainly impact performance
- Guest OS swapfiles
 - Should be configured for worst case (VM pages all memory to guest swapfile) when ballooning is used

NOTE: While both are classified as 'memory optimisations' they both impact storage capacity.

Memory compression

Memory compression is a new feature to vSphere 4.1 (which isn't covered in the lab yet) so I won't cover it here.

Monitoring memory optimisations

TPS

- esxtop;
 - PSHARE/MB - check 'shared', 'common' and 'savings' (memory overcommit)
 - Overcommit % shown on the top line of the memory view (press m). 0.19 = 19%.
 - NOTE: On Xeon 5500 (Nehalem) hosts TPS won't show much benefit until you overcommit memory ([VMwareKB1021095](#))
- vCenter performance charts (under 'Memory');
 - 'Memory shared'. For VMs and hosts, collection level 2.
 - 'Memory common'. For hosts only, collection level 2

Ballooning

- esxtop
 - MEMCTL/MB - check current, target.
 - MCTL? to see if driver is active (press 'f' then 'i' to add memctl columns)

- vCenter performance charts (under 'Memory');
 - 'memory balloon'. For hosts and VMs, collection level 1.
 - 'memory balloon target'. For VMs only, collection level 2.

Swapfiles

- esxtop
 - SWAP/MB - check current, r/s, w/s.
 - SWCUR to see current swap in MB (press 'f' then 'j' to add swap columns)
- vCenter performance charts (under 'Memory');
 - 'Memory Swap Used' (hosts) or 'Swapped' (VMs). Collection level 2.
 - 'Swap in rate', 'Swap out rate'. For hosts and VMs, collection level 1.

NOTE: Remember you can tailor statistics levels - vCenter Server Settings -> Statistics. Default is all level one metrics kept for one year.

Read Duncan Eppings blogpost for some interesting points on [using esxtop to monitor ballooning and swapping](#). See Troubleshooting section 6.46.2 for more information on CPU/memory performance.

3.1.4 Tune ESX/ESXi Host and Virtual Machine networking configurations

Things to consider;

- Check you're using the latest NIC driver both for the ESX host and the guest OS (VMTools installed and VMXNET3 driver where possible)
- Check NIC teaming is correctly configured
- Check physical NIC properties - speed and duplex are correct, enable TOE if possible
- Add physical NICs to increase bandwidth
- Enable Netqueue - see section 2.1.3
- Consider DirectPath I/O - see section 1.1.4
- Consider use of jumbo frames ([though some studies show little performance improvement](#))

Monitoring network optimisations

esxtop (press 'n' to get network statistics);

- %DRPTX - should be 0
- %DRPRX - should be 0
- You can also see which VM is using which pNIC in a team (assuming it's using virtual port ID load balancing), pNIC speed and duplex

vCenter (Performance -> Advanced -> 'Network');

- Network usage average (KB/s). VMs and hosts, collection level 2.
- Dropped rx - should be 0, collection level 2
- Dropped tx - should be 0, collection level 2

See Troubleshooting section 6.3 for more information on networking performance.

3.1.5 Tune ESX/ESXi Host and Virtual Machine CPU configurations

Hyperthreading

- Enable hyperthreading in the BIOS (it's enabled by default in ESX)

- Set hyperthreading sharing options on a per VM basis (Edit Settings -> Options). Default is to allow sharing with other VMs and shouldn't be changed unless specific conditions require it (cache thrashing).
- Can't enable with more than 32 cores (ESX has a 64 logical CPU limit)

CPU affinity

- Avoid where possible - impacts DRS, vMotion, NUMA, CPU scheduler efficiency
- Consider hyperthreading - don't set two VMs to use CPU 0 & 1 as that might be a single hyperthreaded core
- Use cases - licencing, copy protection

CPU power management (vSphere v4.1 only)

- Enabled in BIOS and ESX
- Four levels;
 - High performance (default) - no power management features evoked unless triggered by thermal or power capping events
 - Balanced
 - Low power
 - Custom

NOTE: VMware recommend disabling CPU power management in the BIOS if performance concerns outweigh power saving.

Monitoring CPU optimisations

esxtop (press 'c' to get CPU statistics);

- CPU load average (top line) - for example 0.19 = 19%.
- %PCPU - should not be 100%! If one PCPU is constantly higher than other check for VM CPU affinity
- %RDY - should be below 10%
- %MLMTD - should be zero. If not check for VM CPU limits.
- You can also use 'e' to expand a specific VM and see the load on each vCPU. Good to check if vSMP is working effectively.

vCenter (Performance -> Advanced -> 'CPU');

- '%CPU usage' - for both VMs and hosts, collection level 1
- 'CPU Ready' - for VMs only, collection level 1. Not a percentage like esxtop - see this [blog entry about converting vCenter metrics into something meaningful](#).

See Troubleshooting section 6.46.2 for more information on CPU/memory performance.

3.1.6 Tune ESX/ESXi Host and Virtual Machine storage configurations

In reality there's not that much tuning you can do at the VMware level to improve storage, most tuning needs to be done at the storage array (reiterated in the ESXTOP Statistics guide). So what can you tune? Watch [VMworld 2010 session TA8065](#) (subscription required).

- Increase VM memory. This may increase caching and reduce the impact on storage
- Disable AV scanning (often impacts storage quite heavily)

- Use svMotion to move VMs to a less busy datastore

Multipathing - select the right policy for your array (check with your vendor);

- MRU (active passive)
- Fixed (active/active)
- Fixed_AP (active/passive and ALUA)
- RR (active/active, typically with ALUA)

Check multipathing configuration using `esxcli` and `vicfg-mpath`. For iSCSI check the software port binding.

Storage alignment

You should always align storage at array, VMFS, and guest OS level.

Storage related queues

Use `esxcfg-module` to amend LUN (HBA) queue depth (default 32). Syntax varies per vendor.

Use `esxcfg-advcfg` to amend VMkernel queue depth (default ??). Should be the same as the LUN queue depth.

NOTE: If you adjust the LUN queue you have to adjust on every host in a cluster (it's a per host setting)

Using vscsiStats

See section 3.5 for details of using `vscsiStats`.

NOTE: Prior to vSphere 4.1 (which includes NFS latency in both vCenter charts and `esxtop`) `vscsiStats` was the only VMware tool to see NFS performance issues. Use array based tools!

Monitoring storage optimisations

`esxtop` (press 'd', 'u', or 'v' to get storage metrics for HBA, LUN and per VM respectively);

- KAVG/cmd should be less than 2 (delay while kernel empties storage queue)
- DAVG/cmd should be under 15-20ms (approx)
- ABRTS/s should be zero (this equates to guest OS SCSI timeouts)
- CONS/s should be zero (SCSI reservation conflicts. May indicate too many VMs in a LUN). V4.1 only.

vCenter (Performance -> Advanced -> Disk. Only available in vSphere 4.1)

- Kernel disk command latency - collection level 2
- Physical disk command latency - collection level 2
- Disk command aborts - if greater than 1 indicates overloaded storage. V4.1 only.

Generic tips for optimising storage performance

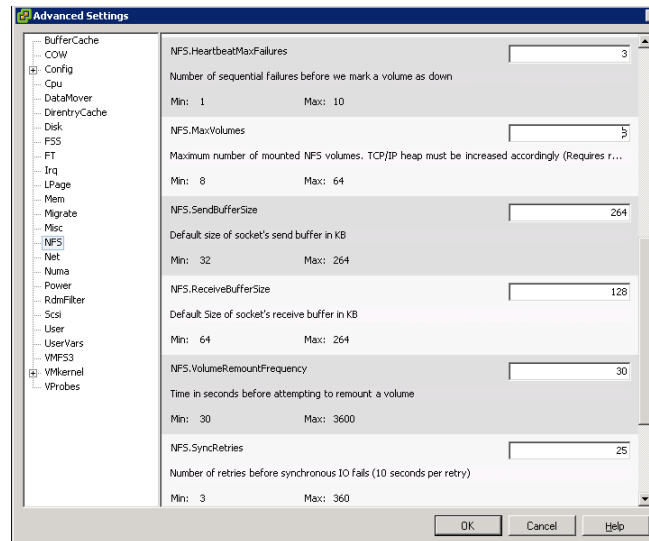
- Check I/Os
- Check latency
- Check bandwidth
- Remember for iSCSI and NAS you may also have to check network performance

See Troubleshooting section 6.46.46.2 for more information on storage performance.

3.1.7 Configure and apply advanced ESX/ESXi Host attributes

These can be configured via Configuration -> Advanced Settings. Things you'll have used this for;

- Checking if Netqueue is enabled/disabled (vmKernel -> Boot)
- Updating your NFS settings to apply Netapp recommendations (if you use Netapp storage)
- [Allowing snapshots on a virtual ESX host in your lab](#) (unsupported but very useful!)
- Disabling transparent page sharing
- Setting preferred AD controllers (when using AD integration in vSphere 4.1)

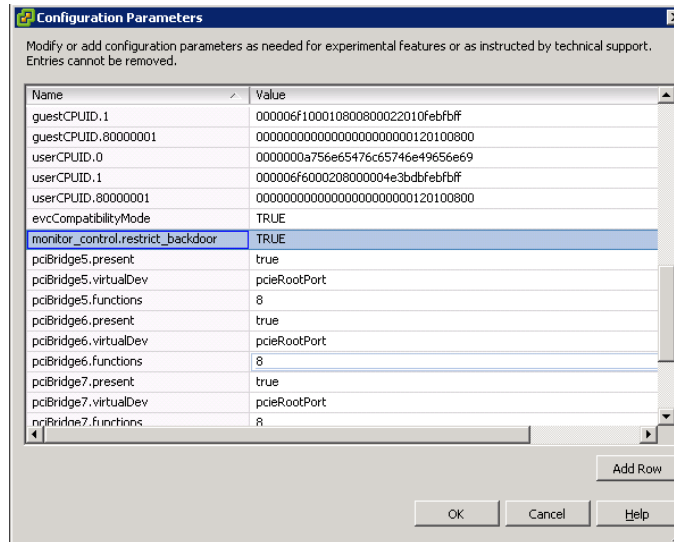


The [vSphere Resource Management Guide](#) lists some of the memory attributes but there are loads to investigate.

3.1.8 Configure and apply advanced Virtual Machine attributes

These are configured on a per VM basis via Edit Settings -> Options -> General -> Configuration Parameters. Things you'll use this at VMware support's recommendation;

- Disabling alerts about a missing SCSI driver (courtesy of this [blogpost](#))
- [Enabling Fault Tolerance on a virtual ESX host](#) for your lab (not that this worked for me)
- [Enabling nested VMs to run on a virtual ESX host](#)



3.1.9 Tune and optimize NUMA controls

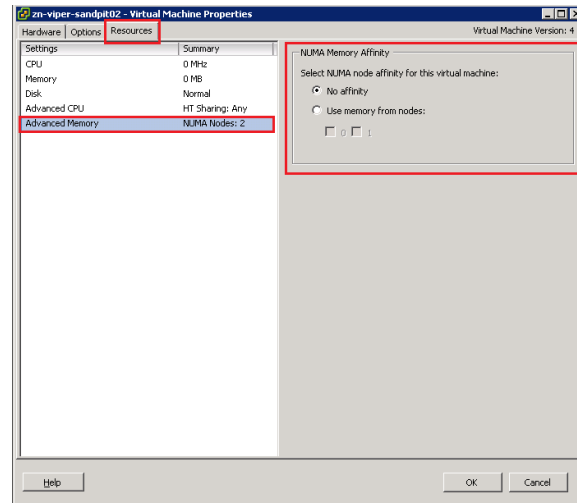
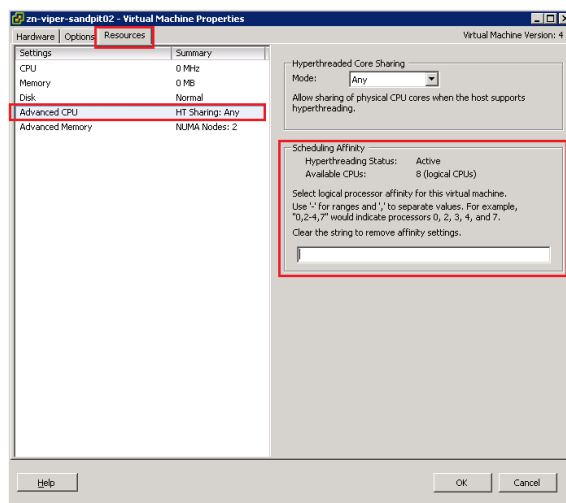
Non Uniform Memory access (NUMA) is a technology designed to optimise motherboard design. Rather than provide a single pool of physical memory to the various CPUs each CPU is given a set of 'local' memory which is can access very quickly. The disadvantage is that not all memory is instantly accessible to all CPUs. Read [VMware vSphere™ : The CPU Scheduler in VMware® ESX™ 4.1](#) for more info.

If you want to understand NUMA, you *need* to check out [Frank Denneman's site](#). As of March 2011 he's got 11 in-depth articles about NUMA.

Practical implications for VCAP-DCA exam?

- Configured per VM. Go to Edit Settings -> Resources tab -> Advanced CPU.
- Can have a performance impact if not balancing properly
- Can be monitored using esxtop ([Frank Denneman's post](#) shows how)
- Setting CPU affinity breaks NUMA optimisations

Configure CPU and memory NUMA affinity;



Monitoring performance impact of NUMA

Using esxtop go to the memory view (m). The first figure is the total memory per NUMA node (approx. 20GB in the screenshot below) and the figure in brackets is the memory free per node. To get more NUMA related statistics press 'f' (to select fields to add) and then 'g' for NUMA statistics.

```

5:18:35am up 237 days 21:47, 141 worlds; MEM overcommit avg: 0.00, 0.00, 0.00
PMEM /MB: 40950 total: 800 cos, 749 vmk, 27214 other, 12186 free
VMKMEM/MB: 39611 managed: 2376 minfree, 13385 rsvd, 25767 ursvd, high state
COSMEM/MB: 78 free: 1600 swap t, 1600 swap_f: 0.00 r/s, 0.00 w/s
NUMA /MB: 20326 ( 4), 19533 (12181)
PSHARE/MB: 878 shared, 138 common: 740 saving
SWAP /MB: 0 curr, 0 relmtgt: 0.00 r/s, 0.00 w/s
ZIP /MB: 0 zipped, 0 saved
MEMCTL/MB: 0 curr, 0 target, 17527 max

```

GID	NAME	NHN	NHIG	NRMEM	NLMEM	N*L	GST	NDO
47	zhclviperapp05	0	0	28.54	7888.63	99	7888.63	
44	zhclunodb01	0	0	3583.87	4336.13	54	4336.13	
46	zhclunoapp04	0	0	2.59	7911.41	99	7911.41	
45	zhclunoapp01	1	0	7.67	3822.33	99	7.67	
18	vobd.4278	-	-	-	-	-	-	-
20	vmware-vmkauthd	-	-	-	-	-	-	-
19	net-cdp.4286	-	-	-	-	-	-	-

As you can see this server in this example is very imbalanced which *could* point to performance issues (it's a 4.0u1 server so no 'wide NUMA'). Looking further you can see that zhclunodb01 only has 54% memory locality which is not good (Duncan Epping suggests under 80% is worth worrying about). Other [people have seen similar situations](#) and [VMwareKB1026063](#) is close but doesn't perfectly match my symptoms. vSphere 4.1 has improvements using 'wide' NUMA support - maybe that'll help...

3.2 Optimize Virtual Machine Resources

Knowledge

- Compare and contrast virtual and physical hardware resources
- Identify VMware memory management techniques
- Identify VMware CPU load balancing techniques
- Identify pre-requisites for Hot Add features

Skills and Abilities

- Calculate available resources
- Properly size a Virtual Machine based on application workload
- Configure large memory pages
- Understand appropriate use cases for CPU affinity

Tools & learning resources

- Product Documentation
 - [vSphere Resource Management Guide](#)
 - [vSphere Command-Line Interface Installation and Scripting Guide](#)
 - [Understanding Memory Resource Management in VMware® ESX™ Server 4.1](#)
 - [VMware vSphere™ : The CPU Scheduler in VMware® ESX™ 4.1](#)
- vSphere Client
 - Performance Charts
- vSphere CLI
 - resxtop/esxtop
- [VMworld 2010 session TA7750 Understanding Virtualisation Memory Management](#) (subscription required)
- [VMware's Performance Best Practices white paper](#)

This objective is focused on the VMs rather than the hosts but there's still a large overlap between this objective and the previous one.

3.2.1 Identify memory management techniques

The theory - read the following blogposts;

- [Frank Denneman's impact of memory reservations](#)
- [Duncan Epping's 'memory limits' post](#) (and work your way through the comments!)

The following memory mechanisms were covered in section 3.1 so I won't duplicate;

- transparent page sharing
- ballooning (via VMTools)
- memory compression (vSphere 4.1 onwards)
- virtual swap files
- NUMA

There are also various mechanisms for controlling memory allocations to VMs;

- reservations and limitations
- shares - disk, CPU and memory
- resource pools (in clusters)

Disable unnecessary devices in the VM settings (floppy drive, USB controllers, extra NICs etc) as they have a memory overhead.

3.2.2 CPU load balancing techniques

Read [VMware vSphere™ : The CPU Scheduler in VMware® ESX™ 4.1](#) and [Frank Denneman's blogpost](#) to understand the theory.

- NUMA architectures (see section 3.1.9 for full details). Don't allocate a VM more vCPUs than per NUMA node.
- Hyperthreading
- Relaxed co-scheduling

Use cases for CPU affinity

- licencing (although some companies such as Oracle [still don't recognise it](#)) where it's based on per physical CPU
- copy protection schemes which bind applications to a CPU (SafeEnd, FireDaemon etc)

Disadvantages

- Breaks NUMA optimisations (see Frank Denneman's post)
- A VM with CPU affinity set cannot be vMotioned, hence it can't be configured for a VM in a fully automated DRS cluster either.

3.2.3 Hot Add prerequisites

- Not enabled by default.
- Guest OS support (Check Jason Boche's [blogpost for details of guest OS and hotplug](#))
- Memory and CPUs can be hot added (but not hot removed) but not all devices can
- Enabled per VM and needs a reboot to take effect (ironically!).
 - Enable on templates
- Virtual h/w v7
- Not compatible with Fault Tolerance - use one of the other.

3.2.4 Calculate available resources

There are various places to check available resources;

- At the host level
 - The summary tab shows a high level view of free CPU and memory (with TPS savings accounted for)
 - The Resources tab shows more in-depth information broken down per VM
- On resource pools
 - Check the reservations and limits set for the resource pool. Is it unlimited (the default) and expandable?
- At cluster level
 - Use the DRS distribution chart to understand the resource allocation

One of the most common support issues is troubleshooting why a particular VM can't be powered on, typically with the error that it failed admission control. Check resource pool settings!

- In the VI client check the VM requirements - CPU, RAM, reservations, limits etc
- In a cluster check if admission control is enabled (Edit Settings -> HA)
- Check the parent resource pool to see if it has reservation or limits set
- Check shares for the VM - if it's in a cluster you can check it's 'worst case allocation' on the Resource tab.

NOTE: One common misconfiguration is to have VMs in the root resource pool. This can completely skew the resource allocation because it'll get a percentage of the root resource pool.

NOTE: Shares are relative to other VMs *on the same host* and only apply when there's contention.

[Food for thought from Duncan Epping](#).

NOTE: When admission control checks available memory (to guarantee to a VM as it's being powered on) it uses 'machine memory' rather than 'guest physical memory'. This means TPS savings are factored into admission control. [Explained by Frank Denneman](#).

Make sure you remember your VCP knowledge about resource pools;

- expandable reservations - the child resource pool can also use spare resources from its parent resource pool
- fixed reservations - the child resource pool has a fixed limit and attempt to allocate more resources will be denied.

Some good thoughts in this [VMware communities thread by Jase McCarty](#) and another [Frank Denneman post](#).

3.2.5 Properly size a VM based on application workload

Refer to section 3.1 regarding tuning memory, CPU, network and storage. Memory is often the main resource to get right. Use either vCenter or esxtop to monitor the memory statistics below;

- Active memory is key - shows an estimated amount of RAM needed by the VM
- Check ballooned or swapped memory
 - Ballooning is probably OK
 - Swap is BAD!

NOTE: Some applications will grab as much memory as they can (or are configured to) and manage it themselves rather than leaving it to the guest OS - Java and Oracle are typical examples.

Unfortunately this makes tuning the VM difficult as neither VMware or the guest OS know what memory is in use (and therefore can be paged/ballooned).

Check using esxtop that a vSMP server is equally balancing CPU across the vCPUs - otherwise it maybe misconfigured or simply not using multithreaded apps (in which case you could consider decreasing vCPU allocation). Consider HAL when doing this.

3.2.6 Configuring large memory pages

From the [VMware Performance Best Practices whitepaper](#);

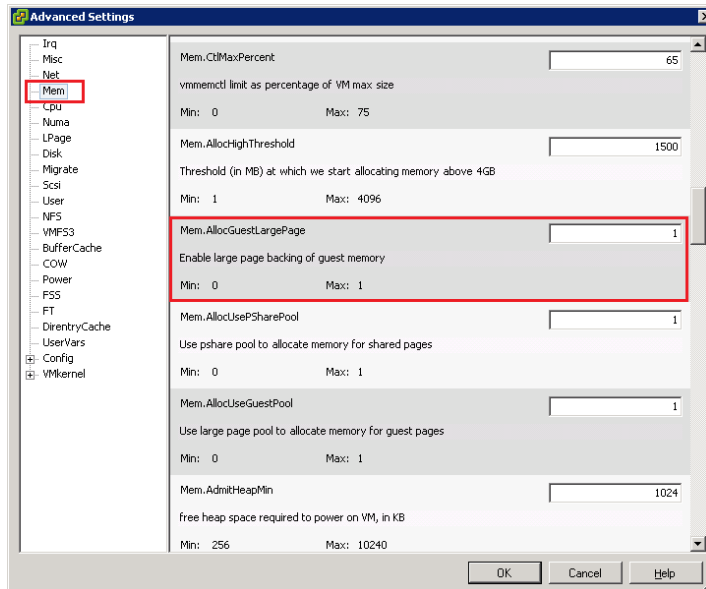
*"In addition to the usual 4KB memory pages, ESX also makes 2MB memory pages available (commonly referred to as "large pages"). By default ESX assigns these 2MB machine memory pages to **guest operating systems that request them**, giving the guest operating system the full advantage of using large pages."*

VMware recommend that when hardware memory virtualisation is enabled you also enable large page support in the guest OS (ESX is enabled by default). Some estimates show a 10-20% performance increase for large memory pages - more in the [VMware white paper on large page performance](#).

If you want to know more check this [good blogpost on large memory pages by Forbes Guthrie](#) and follow his links for some interesting discussion.

Enabling/disabling large pages in ESX

vSphere automatically enables large pages but you can disable it if you want. Go to Configuration -> Software -> Advanced Settings (this applies to all VMs on the host). After changing this setting you'll need to vMotion your VMs off and back onto the host for the memory to be reallocated into small pages.



NOTE: You can also override this setting per VM by adding the following value in the .VMX file; `monitor_control.disable_mmu_largepages = FALSE` (from this [thread](#))

To enable large pages in the guest OS

See vendor documentation (W2k3 and RHEL4 are also covered in the [VMware white paper on large page performance](#));

- Windows 2003 - enable permission to 'Lock memory' in Local Security Policy snapin. Assign rights to the appropriate user (maybe a SQL account if using large pages with SQL server for example)
- Linux - 'echo 1024 > /proc/sys/vm/nr_hugepages'

Impact on TPS

TPS doesn't work with large memory pages, so on Nehalem servers (or any server with hardware assisted memory virtualisation) the memory saving via TPS are minimised until there is memory contention. When memory is overcommitted TPS can break down a large memory page into small pages so TPS benefits can still be realised. ([VMwareKB1021095](#))

NOTE: The impact on TPS only comes into play if you've enabled large pages in BOTH ESX and the guest OS. If it's only enabled in ESX but the guest OS doesn't request large pages, small pages will still be used and TPS will kick in as usual.

Application support for large pages

Consider what applications are running and whether they benefit from large pages - otherwise it's all a waste of time!

[Some details about enabling large pages for SQL Server 2008](#)

3.3 Implement and Maintain Complex DRS Solutions

Knowledge

- Explain DRS affinity and anti-affinity rules
- Identify required hardware components to support DPM
- Identify EVC requirements, baselines and components
- Understand the DRS slot-size algorithm and its impact on migration recommendations

Skills and Abilities

- Properly configure BIOS and management settings to support DPM
- Test DPM to verify proper configuration
- Configure appropriate DPM Threshold to meet business requirements
- Configure EVC using appropriate baseline
- Change the EVC mode on an existing DRS cluster
- Create DRS and DPM alarms
- Configure applicable power management settings for ESX Hosts
- Properly size virtual machines and clusters for optimal DRS efficiency
- Properly apply virtual machine automation levels based upon application requirements

Tools & learning resources

- Product Documentation
 - [vSphere Resource Management Guide](#)
 - [vSphere Datacenter Administration Guide](#) (not listed in blueprint, but details EVC)
- vSphere Client
 - DRS Resource Distribution Chart
- [Frank Denneman's blogpost on using DPM with DRS](#)
- [Jason Boche's blogpost about DPM UI consistency](#)
- [Duncan Epping and Frank Denneman's HA and DRS book](#)
- [DRS limitations with vSMP](#)
- [Fine tuning the DRS algorithm](#)
- [The math behind the DRS algorithm](#)
- Community [blogpost](#) – good log files for troubleshooting EVC
- Frank Denneman's [excellent post on VM-Host affinity rules](#)

3.3.1 Advanced DRS

The basics

Use the (new to vSphere) DRS Faults and DRS History tabs to investigate issues with DRS

By default DRS recalculates every 5 minutes (including DPM recommendations), but it also does so when resource settings are changed (reservations, adding/removing hosts etc). For a full list of actions which trigger DRS calculations see Frank Denneman's HA/DRS book.

[Limitations of DRS with vSMP VMs](#)

[DRS deepdive at Yellow Bricks](#)

NOTE: It's perfectly possible to turn on DRS even though all prerequisite functionality isn't enabled - for example if vMotion isn't enabled you won't be prompted (at least until you try to migrate a VM)!

Affinity and anti-affinity rules

There are two types of affinity/anti-affinity rules;

- VM-VM (new in vSphere v4.0)
- VM-Host (new to vSphere 4.1)

The VM-VM affinity is pretty straightforward. Simply select a group of two or more VMs and decide if they should be kept together (affinity) or apart (anti-affinity). Typical use cases;

- Webservers acting in a web farm (set anti-affinity to keep them on separate hosts for redundancy)
- A webserver and associated application server (set affinity to optimise networking by keeping them on the same host)

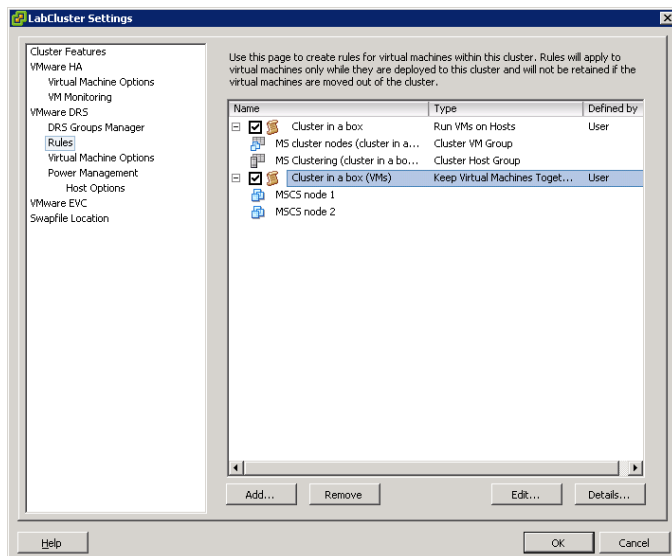
VM-Host affinity is a new feature (with vSphere 4.1) which lets you 'pin' one or more VMs to a particular host or group of hosts. Use cases I can think of;

- Pin the vCenter server to a couple of known hosts in a large cluster
- Pin VMs for licence compliance (think Oracle, although apparently they don't recognise this new feature as being valid – see the [comments in this post](#))
- Microsoft clustering (see section 4.3 for more details on how to configure this)
- Multi-tenancy (cloud infrastructures)
- Blade environments (ensure VMs run on different chassis in case of backplane failure)
- Stretched clusters (spread between sites. See this [Netapp post for Metrocluster details](#))

To implement them;

- Define 'pools' of hosts.
- Define 'pools' of VMs.
- Create a rule pairing one VM group with one host group.
 - Specify either affinity (keep together) or anti-affinity (keep apart).
 - Specify either 'should' or 'must' (preference or mandatory)

This is configured in the 'DRS Groups Manager' and 'Rules' pages of the DRS Properties;



NOTE: The 'Rules' tab was available in v4.0, but the DRS Groups Manager tab is new in v4.1. In v4.0 you could set VM-VM rules, it's only the VM-Host rules which are new.

NOTE: You only need vCenter v4.1 to get the new Host-VM affinity functionality – the hosts themselves can still be running v4.0.

Properly size VMs and clusters for optimal DRS efficiency

Use reservations sparingly – impacts slot size (see section 4.1 Complex HA for details of slot size algorithm).

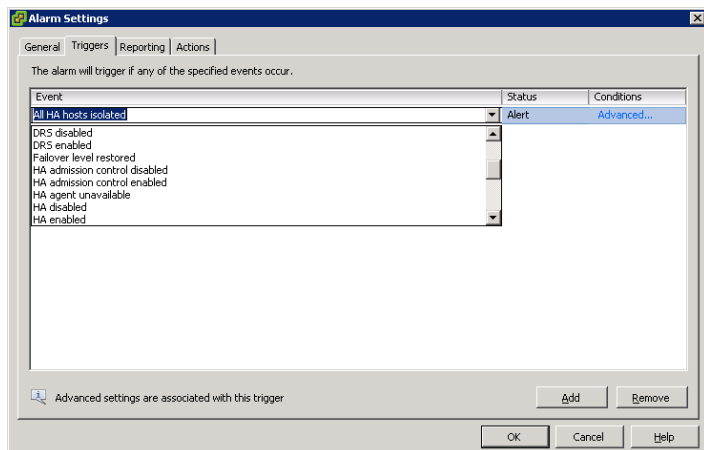
VM-Host affinity rules should be used with caution and sparingly, especially the mandatory rules. HA, DRS, and DPM are all aware of and bound by these rules which could impact their efficiency. VM-VM affinity rules are less of an issue because a) HA ignores them and b) they don't include hosts in their rules (and DRS load balancing is done amongst hosts). Operations which can be blocked by a VM-Host rule include;

- Putting a host into maintenance mode
- Putting a host into standby mode (DPM)

Size VMs appropriately – oversizing VMs can affect both slot size (if using reservations), free resources and the DRS load balancing algorithm.

DRS Alarms

There are no alarms predefined for DRS but there are plenty you can define;



NOTE: I'd have liked a way to get an email notification when a DRS or DPM recommendation was created (in Manual or Partial automation mode) but this doesn't seem possible.

DRS vs DPM automation level

As pointed out by Jason Boche, the [automation level sliders are inconsistent between DRS and DPM](#). This stems from a difference in the way recommendation are done for the two functions;

- For DRS the highest recommendation is rated 1, lowest is 5. Prior to vSphere 4.1 this was the opposite and used starts instead (ie a 5 star recommendation was the highest)
- For DPM the highest recommendation is rated 1, lowest is 5

Be careful when setting these that you're setting is as you planned!

3.3.2 Distributed Power Management (DPM)

BIOS requirements

- Ensure WOL is enabled for the NICs (it was disabled by default on my DL380G5)
- Configure the IPMI/iLO

Hardware requirements

Three protocols for using DPM depending on your hardware features/support;

- WOL is supported (or not) by the network cards (although it also requires motherboard support). It allows a server to be 'woken' up by sending a 'magic' packet to the NIC, even when the server is powered off. See [VMwareKB1003373](#) for details. *Largely unrelated tip – [how to send a 'magic packet' using Powershell!](#)*
- IPMI.
- iLO/DRAC cards

If a server supports multiple protocols they are used in this order: IPMI, iLO, WOL.

NOTE: Unfortunately you can't use DPM with virtual ESX hosts as the E1000 driver (which is used with vESX) doesn't support WOL functionality. This feature joins FT as a lab breaker! See this [post from vinf.net](#).

Configuring DPM

You'll need Advanced licencing or higher to get DPM.

For WOL

- you must use the vMotion (or VMkernel) port (so this must support WOL)
- check the port speed is set to 'auto' on the physical switch

- test that a host 'wakes' correctly from standby before enabling DPM.

192.168.8.15 VMware ESXi, 4.1.0, 260247

Summary Virtual Machines Performance Configuration Tasks & Events Alarms Permissions Maps SiteSurvey Storage Views Hardware Status

Hardware

- Processors
- Memory
- Storage
- Networking
- Storage Adapters
- Network Adapters
- Advanced Settings
- Power Management

Network Adapters

Device	Speed	Wake on LAN Supported	Configured	Switch	MAC Address
Broadcom Corporation NC373i Integrated Multifunction Gigabit Server Adapter					
vmnic1	1000 Full	No	1000 Full	vSwitch1	00:13:2
vmnic0	100 Full	No	100 Full	vSwitch0	00:13:2

Configuring IPMI and ILO/DRAC requires extra steps, done via Configuration -> Software -> Power Management. You need to provide;

- Credentials (vCenter uses MD5 if the BMC supports it else falls back to plaintext)
- the IP address of the IPMI/ILO card
- the MAC address of the IPMI/ILO card

192.168.8.15 VMware ESXi, 4.1.0, 260247

Summary Virtual Machines Performance Configuration Tasks & Events Alarms Permissions Maps SiteSurvey Storage Views Hardware Status

Hardware

- Processors
- Memory
- Storage
- Networking
- Storage Adapters
- Network Adapters
- Advanced Settings
- Power Management

Software

- Licensed Features
- Time Configuration
- DNS and Routing
- Authentication Services
- Power Management
- Virtual Machine Startup/Shutdown
- Virtual Machine Swapfile Location
- Security Profile
- System Resource Allocation
- Advanced Settings

IPMI/ILO Settings for Power Management [Properties...](#)

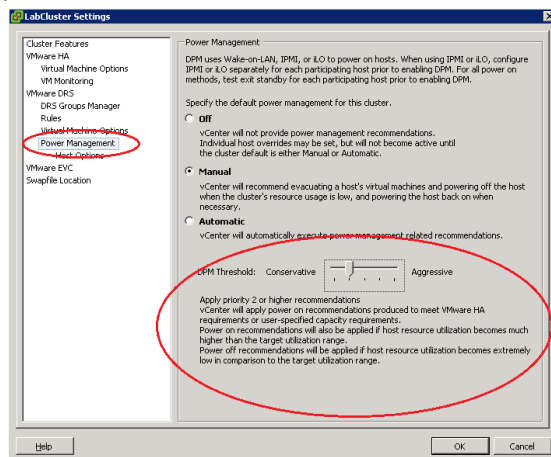
IPMI or ILO settings permit the host to participate in a Distributed Power Management (DPM) enabled cluster. In addition, manual Standby and Power On commands may be issued through vCenter.

User name: Administrator
 BMC IP Address: 192.168.8.25
 BMC MAC Address: 00:17:a4:46:70:e4

Enabling DPM

- Test each host individually to ensure it's able to 'wake' from Standby mode. Set DPM to disabled for any hosts that don't support the wakeup protocols.
- Enable DPM on the cluster
- Set the DPM threshold
- Priority-one recommendations are the biggest improvement and priority-five the least. This is the opposite of the e similar looking DRS thresholds – see [Jason Boche's blogpost about DPM UI consistency](#) for details.

- Disable power management for any hosts in the cluster which don't support the above protocols



Using alarms with DPM

- Define the default alarm for 'Exit standby error' so you can manually intervene if a host fails to return from standby.
- Using DPM means hosts will go offline frequently which makes monitoring host availability difficult as you can't distinguish a genuine outage from a DPM initialised power down.
- Optionally you can create alarms for the following events;
 - Entering standby mode
 - Exiting standby mode
 - Successfully entered standby mode
 - Successfully exited standby mode

3.3.3 Enhanced vMotion Compatibility (EVC)

EVC increases vMotion compatibility by masking off CPU features which aren't consistent across the cluster. It's enabled at cluster level and is disabled by default.

NOTE: EVC does NOT stop VMs from using faster CPU speeds or hardware virtualisation features that might be available on some hosts in the cluster.

NOTE: EVC is required for FT to work with DRS.

Requirements

- All hosts in the cluster must have CPU's from the same vendor (Intel or AMD)
- All hosts must have vMotion enabled (if not who cares about CPU compatibility?)
- Hardware virtualisation must be enabled in the BIOS (if present). This is because EVC runs a check to ensure the processor has the features it thinks should be present in that model of CPU.

NOTE: This includes having the 'No Execute' bit enabled.

Configuring a new cluster for EVC

1. Determine which baseline to use based on the CPUs in your hosts (check the Datacenter Administration Guide chapter 18 for a compatibility table if you don't know, or the more in-depth [VMwareKB1003212](https://kb.vmware.com/kb/1003212) although this won't be available during the exam)

2. Configure EVC on the cluster (prior to adding any hosts)
3. If the host has newer CPU features (compared to your EVC baseline) power off all VMs on the host.
4. Add hosts to the cluster.

Changing the EVC level on an existing cluster

- You can change EVC to a higher baseline with no impact. VMs will not benefit from new CPU features until each VM has been power cycled (a reboot isn't sufficient).
- When downgrading the EVC baseline you need to power off (or vMotion out of the cluster) all running VMs.

The image shows two screenshots related to EVC configuration in vSphere. On the left is the 'Change EVC Mode' dialog box. It has three radio buttons: 'Disable EVC', 'Enable EVC for AMD Hosts', and 'Enable EVC for Intel® Hosts'. The 'Enable EVC for Intel® Hosts' option is selected. Below the radio buttons is a dropdown menu for 'VMware EVC Mode' set to 'Intel® Xeon® Core™2'. The 'Description' section states: 'Applies the baseline feature set of Intel® Xeon® Core™2 ("Merom") processors to all hosts in the cluster.' It lists compatible processor types: Intel® Xeon® Core™2 ("Merom"), Intel® Xeon® 45nm Core™2 ("Penryn"), Intel® Xeon® Core™ i7 ("Nehalem"), and Intel® Xeon® 32nm Core™ i7 ("Westmere"). A 'Compatibility' section shows a warning icon and text: 'The cluster cannot be configured with the selected Enhanced vMotion Compatibility mode; CPU features disabled by that mode may currently be in use by powered-on or suspended virtual machines in the cluster.' with the IP address 192.168.8.12. On the right is the 'VM Summary' page for 'MSCS node 1'. The 'General' tab is active, showing details for a Microsoft Windows Server 2003, Enterprise VM. The 'EVC Mode' is listed as 'Intel® Xeon® Core™2' and is circled in red.

- An alternative approach is to create a new cluster, enable the correct EVC mode, and then move the hosts from the old cluster to the new cluster one at a time.

3.4 Perform Capacity Planning in a vSphere environment

Knowledge

- Understand the DRS slot-size algorithm and its impact on migration recommendations
- Identify tools needed for monitoring capacity planning
- Identify performance metrics related to resource contention and saturation

Skills and Abilities

- Predict when additional ESX/ESXi Host, network or storage resources will be required by observing an existing environment
- Determine when to expand or contract provisioned Virtual Machine resources based upon observed Virtual Machine utilization
- Interpret performance metrics from vCenter to properly size the environment

Tools

- Product Documentation
 - [vSphere Resource Management Guide](#)
 - [Overview Performance Charts Help](#)
- vSphere Client

Again there is a considerable overlap between this objective and the others in section three - the goal of understanding the DRS slot-size is an exact duplicate from section 3.3!

3.4.1 DRS slot size algorithm and its impact on migration recommendations

This was covered in section 3.3. You can always reread the [DRS deepdive at Yellow Bricks](#).

3.4.2 Identify tools needed for monitoring capacity planning

- vCenter Performance Charts
- vCenter Storage views
- esxtop (particularly in batch or reply mode)
- Perfmon
- Third party tools (not likely in VCAP-DCA exam though)

Consider SCSI reservations per LUN, number of VMs per LUN. Adaptive vs predictive LUN sizing.

3.4.3 Predict when additional ESX/ESXi Host, network or storage resources will be required by observing an existing environment

Refer to section 3.1 for the metrics to check. Ballpark;

- Memory - how much is in the host compared to active memory used? Factor in reservations etc
- Network – any dropped packets? Might imply greater bandwidth required...
- CPU – check for long term patterns using Performance Charts.
- I/O – high latency or lack of capacity are the main indicators to look for

3.4.4 Interpret performance metrics from vCenter to properly size environment

Be aware what the various metrics actually show you. For example what's the difference between Host Memory and Guest Memory in the screenshot below?? The answers can be found in VMworld session TA8129 Beginners guide to performance management.

Virtual Machines							
Tasks & Events							
Alarms							
Permissions							
Maps							
Update Manager							
Name, State, Host or Guest OS contains: <input type="text"/>							
Name	State	Host	Provisioned Space	Used Space	Host CPU - MHz	Host Mem - MB	Guest Mem - %
ZVSPRDB05	Powered On	zcgprvsh01.mfl.co.uk	220.00 GB	212.00 GB	2612	8000	25
LONW04138 - SAS PC	Powered On	zcgprvsh05.mfl.co.uk	206.02 GB	195.66 GB	773	6021	19
zcgprapp09	Powered On	zcgprvsh04.mfl.co.uk	45.00 GB	8.01 GB	26	4176	5
ZCGPRCTX14	Powered On	zcgprvsh04.mfl.co.uk	28.00 GB	9.83 GB	906	4169	5

vCenter and ESXTOP present statistics differently. While ESXTOP tends to display a more useful figure (%CPU ready for example) the value presented in vCenter needs to be calculated depending on the time interval.

Remember that vCenter summary statistics can sometimes mislead - memory per host looks fine in the screenshot above but you might find NUMA locality is low (for example).

3.5 Utilize Advanced vSphere Performance Monitoring Tools

Knowledge

- Identify hot keys and fields used with resxtop/esxtop
- Identify fields used with vscsiStats

Skills and Abilities

- Configure esxtop/resxtop custom profiles
- Determine use cases for and apply esxtop/resxtop Interactive, Batch and Replay modes
- Use vscsiStats to gather storage performance data
- Use esxtop/resxtop to collect performance data
- Given esxtop/resxtop output, identify relative performance data for capacity planning purposes

Tools & learning resources

- Product Documentation
 - [vSphere Resource Management Guide](#)
 - [vSphere Command-Line Interface Installation and Scripting Guide](#)
 - vSphere Client
- vSphere CLI
 - esxtop/resxtop
 - vscsiStats
- [VMworld 2008 session TA1440 - ESXTOP for Advanced users](#)
- [VMworld 2009 session TA3838 - ESXTOP for Advanced users](#)
- [VMworld 2010 session TA6720 - ESXTOP for Advanced users](#) (subscription required)

This is one objective where you definitely have to get hands on - there's no way you'll learn esxtop otherwise. Ideally you'll have a real infrastructure to play with as you want hosts with memory contention, ballooning, swapping, NUMA optimisations etc so you can play/understand the features.

3.5.1 Using resxtop

Two ways of invoking;

- `resxtop --server <esxi host>`
- `resxtop --server <vCenter server> --vihost <esxi host>`

3.5.2 Determine use cases for and apply esxtop Interactive, Batch and Replay modes

First things first - start by watching some VMworld presentations from [2008](#) and [2010](#) (subscription required). Then read some common counters to understand from the obligatory [Duncan Epping blogpost about esxtop](#) and the full list in [Interpreting esxtop statistics whitepaper](#).

Interactive (default. Used to check current performance)

- `esxtop`

Batch mode (used to gather performance statistics for further analysis)

- Run esxtop in interactive mode and configure the counter you want to monitor

- run `'esxtop -b > <filename.csv>'` (press CTRL-C to stop) or `'esxtop -b -n 100 -d 5'` to only monitor one hundred iterations at 5 second intervals
NOTE: you must redirect the output using '>', and you should use .CSV

Replay mode (used to analyse previously collected statistics)

- *Generate the performance statistics using `vm-support -S -d` (duration in seconds)*
- *Replay them using `esxtop -R <vm-support file>`*

You can see screenshots of this process on [Hany Michael's blogpost](#).

NOTE: resxtop isn't vifastpass aware, so you have to specify `--server`, `--username`, and `--password` when used from the RCLI or vMA. Also resxtop doesn't support the replay mode - is this still true?

3.5.3 Using esxtop custom profiles

A custom profile is where you choose to display different statistics from the default.

1. Run esxtop in interactive mode and use 'f' and 'o' to choose and order the statistics you're interested in.
2. Press 'W' and specify a filename to save to (you can override the defaults or use a new file)
3. Load you custom profile using `esxtop -c <filename>`

3.5.4 Hotkeys and fields in esxtop

Commonly used keys (note these are case sensitive);

h - show help	o - order fields	s - set refresh interval
f - choose which fields to display	W - save settings	e - expand an entity
V - view only VM stats		

3.5.5 Using vscsiStats

Prior to vSphere v4.0u2 there was very little information available about NFS storage performance, which is how I became familiar with vscsiStats (which provides storage profiles). It's useful for profiling the storage I/O, regardless of the storage protocol used (see section 1 for details on storage workloads). Read [the official vscsiStats manual](#). Note;

- Only available locally on an ESX/i host, not via RCLI or vMA.
- Not included on vSphere v4.0 hosts
- Included by default on vSphere v4.1 hosts

vscsiStats Theory

- Latency – anything above 15-20ms may indicate performance issues with underlying storage array.
- Seek distance – measures sequential vs random.

- Values to the extreme sides of the histogram imply random access, more towards the middle implies sequential
- Sequential = good, random = bad (generalisation!)
- IO size – the stripe size on the array may be optimised accordingly
- Read/write ratios

Gathering the above information lets you make informed decisions about which underlying RAID level may be most appropriate (depending also on the storage array).

Using vscsiStats - process

1. Find the world ID of the VM you're interested in profiling
2. Start monitoring the VM's storage (and optionally a specific disk)
3. Display statistics *You must do this before stopping monitoring.*
4. Stop monitoring

Using vscsiStats - the syntax

1. vscsiStats -l
2. vscsiStats -s -w <world ID>
3. vscsiStats -p all -w <world ID>
4. vscsiStats -x

NOTE: Follow [@virtualirfan](#) on twitter - he wrote vscsiStats! Read his [VMworld 2007 presentation](#).

Check [Duncan Epping's blogpost](#) for some screenshots of vscsiStats in action and [Gabe's blogpost](#) for a full walkthrough. Michael Poore also has an [interesting post on using Microsoft Chart Controls to visualise vscsiStats data](#) and Eric Zandboer even has [cool 3d Excel charts!](#)

4 Business Continuity

4.1 Implement and Maintain Complex VMware HA Solutions

Knowledge

- Identify the three admission control policies for HA
- Identify heartbeat options and dependencies

Skills and Abilities

- Calculate host failure requirements
- Configure customized isolation response settings
- Configure HA redundancy in a mixed ESX/ESXi environment
- Configure HA related alarms and monitor an HA cluster
- Create a custom slot size configuration
- Understand interactions between DRS and HA
- Create an HA solution that ensures primary node distribution across sites
- Analyze vSphere environment to determine appropriate HA admission control policy
- Analyze performance metrics to calculate host failure requirements
- Analyze Virtual Machine workload to determine optimum slot size
- Analyze HA cluster capacity to determine optimum cluster size

Tools & learning resources

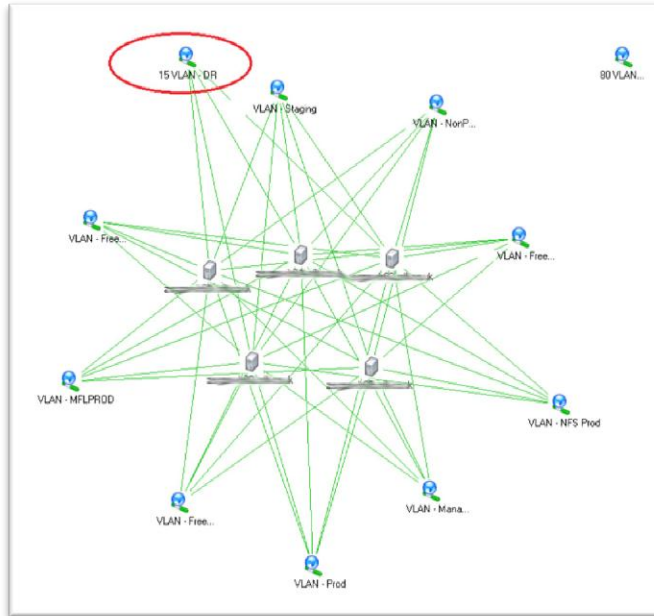
- Product Documentation
 - [vSphere Availability Guide](#)
 - [VMware HA: Deployment Best Practices](#)
- vSphere Client
- Blog posts
 - [Duncan Epping's HA Deepdive](#)
 - [Sample chapter from Duncan Epping and Frank Denneman's upcoming book](#)
- Session BC7803 (VMworld 2010)
- Knowledgebase articles
 - [VMware KB1006421 – Advanced configuration options for VMware HA](#)
 - [VMware KB1001596 – Troubleshooting HA](#)

4.1.1 HA basics

Requirements;

- shared storage
- Common networks
- Ideally similar (or identical) hardware for each host

A good way to check that all hosts have access to the same networks and datastores is to use the 'Maps' feature. Select your cluster then deselect every option except 'Host to Network' or 'Host to Datastore';



As you can see in this diagram the 15 VLAN is not presented to every host (it's slightly removed from the circle) and at least one VM in the cluster has a network assigned (in the top right) which isn't available in this cluster at all.

Clusters consist of up to 32 hosts. The first five hosts in a cluster will be primaries, the rest secondaries. You can't set a host to primary or secondary using the VI client, but you can using the AAM CLI (not supported, see how in this [Yellow bricks article](#)). One of the primaries will be the 'active primary' which collates resource information and places VMs after a failover event.

Heartbeat options and dependencies

- Heartbeats are used to determine whether a host is still operational
- Heartbeats use the service console networks by default, or the management network for ESXi hosts.
- They're sent every second by default. Can be amended using `das.failedetectioninterval`
- Primaries send heartbeats to both other primaries and secondaries, secondaries only send to primaries.
- After no heartbeats have been received for 13 seconds the host will ping its isolation address.

HA operates even when vCentre is down (the AAM agent talks directly from host to host), although vCentre is required when first enabling HA on a cluster.

Diagnosing issues with heartbeats – see [VMware KB1010991](#)

Ports required for HA: 8042-8045 (UDP). Presumably these ports are from ESX host to ESX host?

4.1.2 Cluster design

- Primary/secondary distribution
 - No more than four blades per chassis
 - At least one primary must be online to join new hosts to cluster
 - Can be configured with `aamCLI`, (but these settings are not persistent across reboots and not supported)
 - Use `Get-HAPrimaryVMHost PowerCLI cmdlet` (vSphere v4.1 onwards). [Example](#)

- Interactions between HA and DRS
 - Resource defragmentation (v4.1 onwards)
 - Restart VM on one host, then DRS kicks in and load balances. Priority is to restart the VM.
- Large vs small cluster size - a larger cluster reduces the overhead of N + 1 architecture, but consider other factors such as LUN paths per host (only 255 LUNs per host, 64 NFS datastores). See this [great post at Scott Drummonds Pivot Point blog](#) and [Duncan Epping's followup post](#).
- Enough capacity? Look at performance stats in vCentre for the running workloads.
- DRS host affinity rules may be useful depending on storage implementation. You can pin VMs to specific hosts if the storage is not 100% shared (see VMworld session BC7803 for details)
- Design networking to be resilient (dual pNICs for Service Console for example)
- Avoid using 'must' host-affinity rules (introduced in vSphere 4.1) where possible as it limits the ability of HA to recover VMs.

4.1.3 Admission Control

Admission Control is a mechanism to ensure VM's get the resources they require, even when a host (or hosts) in a cluster fails. Admission control is ON by default.

Three admission policies;

- No. of host failures to tolerate (default)
 - Generally conservative
 - Uses slots (can be customised). Reservations can cause sizing issues.
- % of resources
 - More flexible when VMs have varying resource requirements
 - Resource fragmentation can be an issue
- Dedicated failover host
 - Simple - what you see is what you get
 - Wastes capacity – specified host is not used during normal operations.
 - Often dictated by organisational policies

Both DRS and DPM respect the chosen admission control policy. This means hosts would not be put into suspend mode if the failover level would be violated for example. See [VMware KB1007006](#) for details.

Analyse a cluster to determine appropriate admission control policy

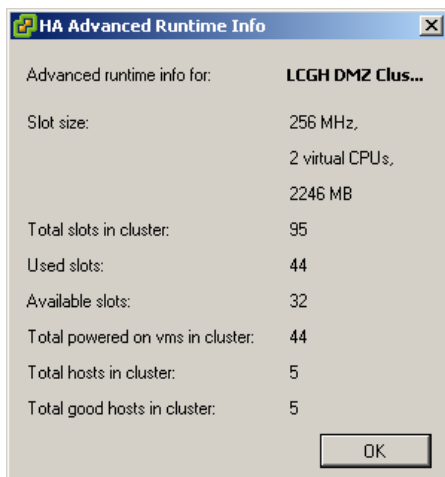
Factors to consider;

- Required failover capacity vs available failover capacity. Dedicated Failover host only allows one host maximum for example.
- Similarity of hosts (percentage of resources policy better for disparate h/w)
- Similarity of VMs (one oversized VM can affect slot sizing but also percentage of resources)

Analyse slot sizing (inc. custom sizes)

Slot sizing;

- Memory = smallest reservation + memory overhead for VM. Override using `das.slotMEMinMB`. Set a minimum using `das.vmMemoryMinMB`.
- CPU = smallest reservation or 256MHz (whichever is smaller). Override using `das.slotCPUinMHZ`. Set a minimum using `das.vmCPUMinMHZ`.
- Current slot size is shown in 'Advanced Runtime Info' for cluster
NOTE: This only shows the total slots in the cluster rather than slots per host. If a particular host has more memory or CPU compared to the other host it will have a higher number of slots.



NOTE:

- The 'available slots' figure shown in the 'Advanced Runtime Info' tab will be equal to (total slots – used slots) – *slots reserved for failover* (which isn't shown in the dialog). This is why 'used slots' and 'available slots' doesn't add up to 'total slots'.
- The total number of slots will take into account virtualisation overhead. For example in a cluster with 240GB RAM total only 210GB may be available to VMs (the rest being used for the vmKernel, service console (on ESX) and device drivers etc. If slot size is 2.2GB RAM there will be roughly 95 slots total. See this [VMware communities thread](#) for more info on virtualisation overhead.

To calculate failover capacity;

- Decide how many host failures you want to cope with
- Calculate the number of slots for each host in the cluster and therefore the total slots available. If all hosts are identical (CPU, mem) then simply divide the total number of slots by the number of hosts.(see Advanced Runtime Info)
- Subtract the largest x hosts from the number of slots (where x is the number of failures to tolerate)

This will give the number of slots that HA will keep reserved.

NOTE: Using 'No. of host failures' often leads to a conservative consolidation ratio.

Percentage of resources gotcha – if you set to 50% but you have more than 10 hosts in your cluster you can run into problems. In theory you can still reserve enough capacity but you can't guarantee that a primary role will still be working.

4.1.4 Isolation Response

Network isolation;

- Heartbeat pings every second (das.failedetectioninterval = 1000)
- 15 second timeout (das.failedetectiontimeout=15000)
 - Increase to 20 seconds (20000) for 2nd service console or second isolation address
 - Increase to 60 seconds (60000) if portfast is not set (to allow time for spanning tree)
- Advanced settings
 - das.isolationnetworkx – used to define multiple isolation networks
 - das.usevMotionNIC?? – used with ESXi (which has no service console)
- There is a small chance that HA could shutdown VMs and not restart them on another host. This only occurs when the isolated host returns to the network between the 14th and 15th second. In this case the isolation response is triggered by the restart isn't because by then the host is no longer considered failed ([VMware KB2956923](#))

Default settings for isolation response;

ESX 3.5	Poweroff
ESX 3.5(u3 through u5)	PowerOn
vSphere	Shutdown

Restart interval after a failover

- 2, 6, 14, 22, 30mins
- Hosts may be in standby mode (when using DPM) so could take several mins for the host to power-up and be ready to host VMs
- VM restart count – das.maxvmrestartcount

Split brain

- Occurs when both management network and storage fail (more likely with NFS, iSCSI or FCoE)
- VM is restarted on another host but continues to run in memory on the isolated host. When that host rejoins the network the VM is running simultaneously on two hosts. Bad!
- vSphere 4.0U2 solves this. For prior versions either avoid 'Leave powered on' as an isolation response or manually close processes on isolated ESX host before rejoining.

4.1.5 VM Monitoring & Application Monitoring

Not in the blueprint, but useful to know.

4.1.6 Operational considerations

You can monitor clusters using the following vCentre alarms;

- You can use the usual host alerts - host failed , thermal, memory usage over threshold etc
- Cluster high availability error – a specific error which you can set actions for

If you're doing network maintenance, put the cluster in maintenance mode (not the hosts) to avoid the isolation response being triggered.

You need to disable then re-enable HA if you make any of the advanced changes mentioned above before they become active.

4.2 Deploy and Test VMware FT

Knowledge

- Identify VMware FT hardware requirements
- Identify VMware FT compatibility requirements

Skills and Abilities

- Modify VM and ESX/ESXi Host settings to allow for FT compatibility
- Use VMware best practices to prepare a vSphere environment for FT
- Configure FT logging
- Prepare the infrastructure for FT compliance
- Test FT failover, secondary restart and application fault tolerance in a FT Virtual Machine

Tools & Learning materials

- Product Documentation
 - [vSphere Availability Guide](#)
- Knowledgebase articles
 - [VMware KB1008026](#) – Disabling and turning off FT
 - [VMware KB1008027](#) – Processors and guest OS support for FT
 - [VMware KB1017714](#) – Disable FT compliance checks
- [VMware white paper on FT Architecture and Performance](#)
- [VMware white paper on Creation and Design of FT](#) (academic and technical)
- [VMware forums for HA and FT](#) (a good place to learn typical issues)
- [VMworld s2010 session BC8274](#) – FT best practices and use cases (subscription only)
- [Brian Atkinson's blogpost](#)
- [Eric Siebert's Master of FT blogpost](#)
- [Barry Combs FT blogpost](#) (real world use cases)
- [Hany Michael's prize winning blogpost on FT](#) (with an uber diagram!)

The main document to work through for the VCAP-DCA is the Availability Guide but there are plenty of good white papers and blog posts which give useful background information. If you have access to the 2010 VMworld content it's worth watching session BC8274 which covers most of the material on the blueprint.

4.2.1 FT requirements (hardware, software and feature compatibility)

Compatibility

- Firstly you have to make sure your host hardware will support FT – it's more demanding than many other VMware features.
 - The main requirement is to have Intel Lockstep technology support in the CPUs and chipset. Rather than list the processor families which support FT you can read [VMwareKB1008027](#).
 - Hardware virtualisation must also be enabled in the BIOS (not always on by default).
- You need to ensure the guest OS and CPU combination is supported (as the Availability Guide states, Solaris on AMD is not for example).
- Must have HA enabled on the cluster
- Licencing– you need Advanced or higher to run FT

- Host certificates need to be enabled. If you did a clean install of vSphere 4.x this is enabled by default but if you upgraded from VI3.x you have to explicitly enable it (vCentre settings, SSL)
- Should avoid mixing ESX and ESXi hosts in a cluster with FT-enabled VMs ([VMwareKB1013637](#))
- There are also VM level requirements
 - No USB or sound devices
 - No NPIV
 - No paravirtualized guest OS
 - No physical mode RDMS
 - Hot plug (memory, CPU, hard disks etc) is automatically disabled for FT-enabled VMs
 - No Serial or parallel ports

Restrictions

FT places quite a few restrictions on the features you can use;

- No SMP
- No snapshots – this also means no VCB or any other backup technology which relies on an underlying snapshot. [VMwareKB1016619](#) describes how you can do backups using templates or storage array level snapshots. Both seem pretty awkward and far from ideal.
- No storage vMotion
- No thin provisioned disks

So what features are supported?

- HA, DRS and vMotion ([DRS with FT is new with vSphere v4.1](#), vMotion always worked)
NOTE: DRS is only available if EVC is enabled. If you disable EVC there is a warning that features such as FT with DRS will not work.

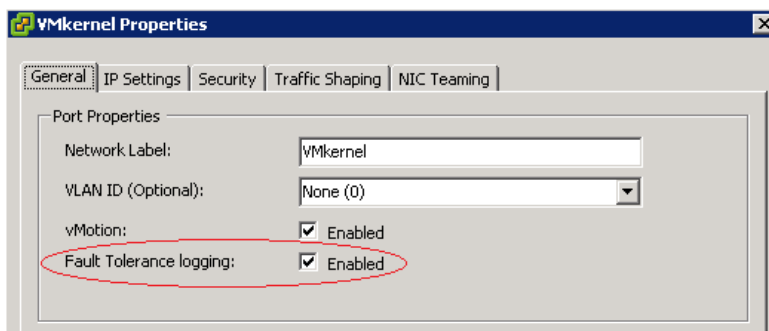
4.2.2 Preparing the infrastructure for FT compliance

The main requirement is for an additional GB capable NIC for FT logging.

- Configuring networking for FT
- Check compliance with FT
- Enable FT per VM

Configuring networking

Just like enabling vMotion, it's recommended (but not enforced) to have a dedicated NIC (preferably on a separate vSwitch and subnet) where you enable Fault Tolerance logging (on a VMkernel port);

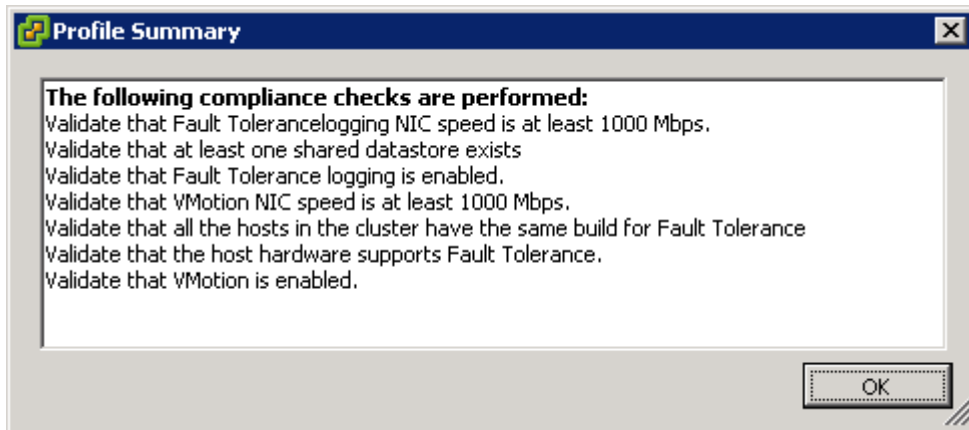


Sharing vMotion and FT logging (as in the screenshot above) might be OK in a lab environment or on 10GB NICs.

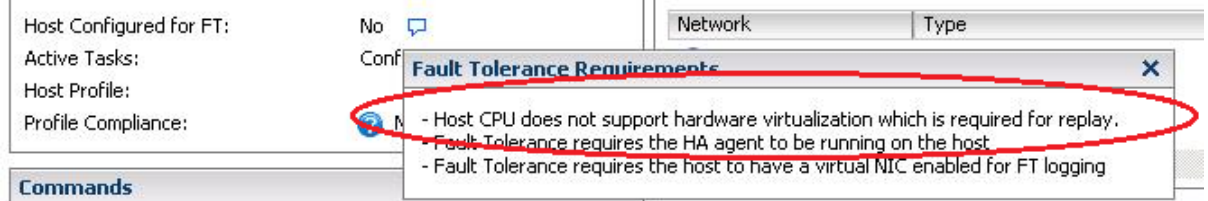
Checking compliance

So besides manually working through these extensive lists how can you check a host for FT compatibility? Various ways (only the first two are 'native' tools so these will probably be your only choice during the VCAP-DCA exam);

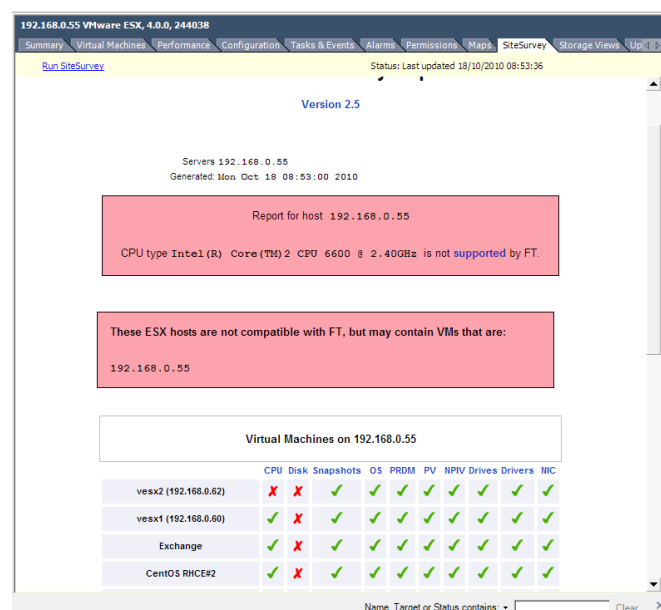
- Run a cluster compliance check (from the Profile Compliance tab of a cluster). The 'Description...' field will detail any issues;



- Slightly less comprehensive is the host's summary screen;



- Boot the server in question using the CPUInfo CD ([provided as an .ISO by VMware](#))
- Install and run the [SiteSurvey plugin](#) to vCenter. This simply integrates as a plugin to the VI client and illustrates any issues;



Enabling FT per VM

This is simply done from the context menu for the FT. If the prerequisites aren't satisfied the option will be greyed out. If the VMDK files are thin provisioned they will be converted to eager-thick-zeroed when you enable FT on the VM. For large VMs this could take a considerable time.

Determining if a VM can be FT-enabled using vim-cmd– [VMwareKB1026509](#)

You can see affinity rules which keep primary and secondary on separate hosts.

4.2.3 Miscellaneous issues

There are times when you need to disable FT, often due to its restrictions (for instance you can't have FT running while you patch the underlying host). Two options;

- Disable – the secondary remains but isn't up to date. Use when temporarily disabling FT
- Turn off – the secondary is deleted. Use when the change is permanent. [VMwareKB1008026](#) details these options.

You can enable a CPU reservation (which is then applied to both primary and secondary) if you're concerned about contention with other VMs (an FT-enabled VM will automatically have a full memory reservation).

Remember that FT-enabling a VM adds a second running VM, hence extra resources are being used.

Powering off the primary VM will also power off the secondary VM

You need to disable FT to perform a storage vMotion

There are several default alarms you can use with FT([VMwareKB1025755](#));

- VM Fault Tolerance state changed
- VM FT vLockstep interval status changed
- No compatible host for Secondary VM
- Secondary VM log latency exceeded

4.2.4 Best practices

Most of the best practices can be summarised as 'maintain consistency'. This applies at various levels;

- The hosts in the cluster should be of similar spec and performance (otherwise a secondary VM may not keep up with a primary VM or vice versa). This also enhances chances of compatibility. This is a best practice for VMware clusters in general but doubly so for FT.
- The FT-enabled VMs should be spread between the available hosts to avoid overloading either logging NICs or host CPU. By default you can't run more than four FT-enabled VMs on a given host anyway.
- Use identical power management and hyperthreading features on hosts used with FT. Typically disable power management.
- You can change the NIC teaming policy to better utilise multiple NICs for FT logging. See [VMwareKB1011966](#) for details.

4.2.5 Lab considerations

Testing and experimenting with FT is tricky if your home lab doesn't support FT, and it is one of the harder features to get compliant hardware for. Luckily several people have posted great articles about low cost hardware which is FT compliant – [here](#), [here](#), [here](#)

Some people have been successful [running virtual FT](#) ie running virtual ESX hosts and enabling FT on a nested VM, but I couldn't get this to work in my lab, even though the physical host's CPU was compatible.

4.3 Configure a vSphere Environment to support MSCS Clustering

Knowledge

- Identify MSCS clustering solution requirements
- Identify the three supported MSCS configurations

Skills and Abilities

- Configure Virtual Machine hardware to support cluster type and guest OS
- Configure a MSCS cluster on a single ESX/ESXi Host
- Configure a MSCS cluster across ESX/ESXi Hosts
- Configure standby host clustering

Tools & learning resources

- Product Documentation
 - [Setup for Failover Clustering and Microsoft Cluster Service](#)
- vSphere Client

The main guide for this section is the 'Setup for Failover clustering and Microsoft Cluster Service' whitepaper. Very little has changed in regards to running MSCS on VMware since the VI3 days so if you're familiar with that (and it was on the VCP syllabus) then don't read any further! If you want a refresher however (and a few tidbits which are new to vSphere 4.1), read on....

4.3.1 Supported MSCS configurations

Three options;

- Cluster in a box
- Cluster across boxes
- Standby (one physical node, one virtual node)

4.3.2 Solution requirements

Hardware

One of the main requirements is a FC SAN (this is one of the rare features which doesn't work with NFS).

Virtual hardware

- Use the correct SCSI adaptor.
 - LSI Parallel for all OSs except Win2k8 which needs the newer LSI SAS.
- Use the correct storage abstraction
 - Cluster in a box - use virtual disks (local or remote)
 - Cluster across boxes – use physical mode RDM (can be virtual mode for W2k3)
 - Standby clustering – use physical mode RDM
- Use thick provisioned disks (eagerzeroedthick)
- You must use h/w v7 with ESX/ESXi 4.1

Restrictions

Features you can't use in conjunction with MSCS;

- vMotion. Interestingly this is not *recommended* (a very vague support stance!) as opposed to not supported. See p30 of the [Setup for Failover Clustering and Microsoft Cluster Service](#) guide.
- HA and DRS clusters. Prior to vSphere 4.1 you couldn't put VMs running in an MSCS in a VMware cluster at all (it worked, but wasn't supported). With vSphere 4.1 this is now supported (you have to use VM-Host affinity rules – see chapter 5 of the [Setup for Failover Clustering and Microsoft Cluster Service](#))
- NPIV
- iSCSI and NFS
- FT (would be rather pointless)
- Round robin multipathing (when using the NMP) – see [VMwareKB1010041](#)

For all three solutions you follow a similar process;

1. Create the first VMs (including at least two vNICs, don't attach the storage yet)
2. Create the second server (either clone or from scratch, don't attach storage yet)
3. Attach the shared storage to the first server (using the applicable abstraction type)
4. Attach the shared storage to the second server (using the same abstraction type)

4.3.3 Cluster in a box

Follow this process;

- Create the first VM
 - Two NICs (one public, one heartbeat)
 - Set the disk timeout to 60 seconds (registry entry)
- Clone the first VM using vCenter, remembering to change the SID
- Add a quorum hard disk (and therefore SCSI adapter) to the first VM, assign to a new SCSI ID (1,0)
 - Select 'Support clustering features such as Fault Tolerance' to ensure the VMDK's are *eagerzeroedthick*. Follow [VMwareKB1011170](#) for details on how to check the VMDK format (it's not as simple as you think!).
 - Set the SCSI adapter type – LSI Parallel for W2k3, LSI SAS for W2k8
 - Set the SCSI bus sharing mode to virtual
- Add the same hard disk to the second VM
 - Use the same SCSI ID (1,0) etc
 - Use the same adapter type and adapter mode
- Setup MSCS (you can follow this useful [step by step guide to building a cluster](#))

4.3.4 Cluster across boxes

Follow an almost identical process to cluster in a box with a few exceptions;

- Create the two VMs as before
 - You should have an extra NIC (compared to cluster in a box) as both networks (public and private) need to span multiple hosts
- When adding the quorum disk;
 - Use an RDM (preferably in physical mode, but virtual is OK for W2k3) on an unformatted SAN LUN.

- Select 'Support clustering features such as FT', set the SCSI adapter type (LSI Parallel for W2k3, LSI SAS for W2k8), and use a new SCSI ID (1,0) (both as before)
- Set the SCSI bus sharing mode to physical (instead of virtual)
- Add the quorum disk to the second VM
 - Add an existing hard disk, specify the same RDM as the first VM, also physical mode
 - Use the same SCSI ID and SCSI adapter mode as for the first VM (ie physical)

4.3.5 Standby clustering

Follow a very similar process to cluster across boxes with a few exceptions;

- Create the first node using a physical server.
 - This server must have access to the same SAN LUN as the ESX host being used to host the standby VM
 - Unlike the other methods you attach the storage to the first node before creating the second (virtual) node
- Create a single VM for the second node
 - Ensure it has access to both the public and private networks available to the physical node
- Add the shared disk/s (quorum and optionally other disks) to the VM (second node);
 - Use an RDM (MUST be physical mode) pointing to the same LUN used by the physical node. As you're not creating a new disk you won't have to specify thick provisioning.
 - Set the SCSI adapter type (LSI Parallel for W2k3, LSI SAS for W2k8), and use a new SCSI ID (1,0) (both as before)
 - Set the SCSI bus sharing mode to physical (instead of virtual)
- Add the quorum disk to the second VM
 - Add an existing hard disk, specify the same RDM as the first VM, also physical mode
 - Use the same SCSI ID and SCSI adapter mode as for the first VM (ie physical)
- Install MSCS.
 - Note: If using Windows 2003 you must configure the Microsoft Cluster Service to use the 'minimum configuration' option.

There are a few extra constraints when using standby clustering;

- As stated previously you can't use Round Robin multipathing in the ESX host. Similarly you can't install multipathing software in the guest OS of either the VM or the physical node.
- Use the STORport Miniport driver for the FC HBA in the physical node (instead of the default)

4.3.6 Using MSCS in HA/DRS enabled clusters

This is a new feature introduced in vSphere 4.1. Previously an MSCS cluster wasn't supported (by either Microsoft or VMware) if it resided in a cluster with HA or DRS enabled, *even if you disabled those features for the VMs in question*. This meant using separate hardware which negated some of the benefits of virtualisation. By using VM-VM and VM-Host affinity rules it's now fully supported.

VM-VM affinity rules (for DRS)

VM-VM affinity rules are used with DRS to ensure VMs stay either together or apart

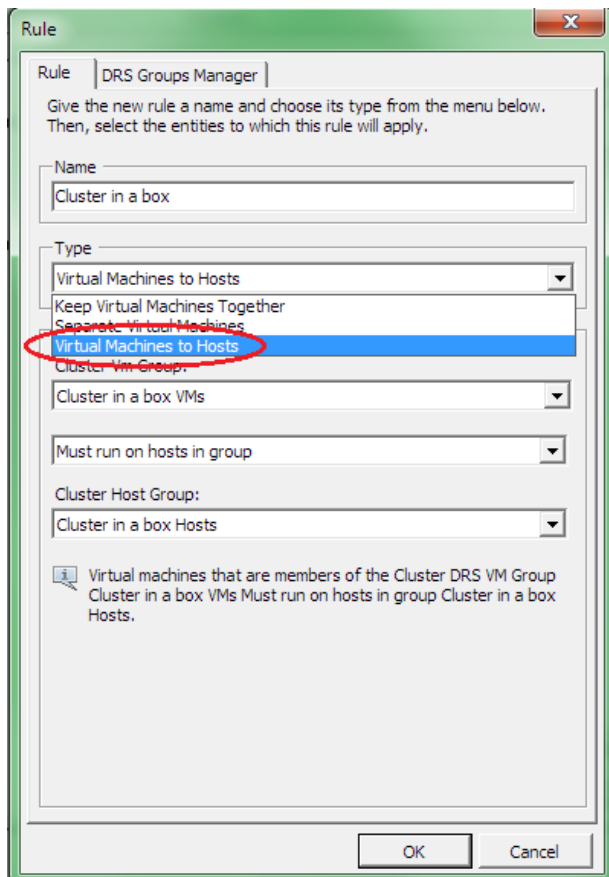
- Use VM-VM *affinity* rules to keep VMs together (cluster in a box)
- Use VM-VM *anti-affinity* rules to keep VMs apart (cluster across boxes)

VM-Host affinity rules (for HA)

VM-Host affinity rules are used to compensate for the fact that HA doesn't obey the above VM-VM rules (whereas DRS does).

For cluster in a box;

- Create a Host DRS Group containing a maximum of two hosts (one to run the VMs and one for failover)
- Create a VM DRS Group containing the MSCS VMs
- Setup an affinity rule between the above two groups using the new 'Virtual Machines to Hosts' type (new in v4.1). Use 'MUST run on hosts in group'



For cluster across boxes;

- Create two Host DRS Groups (each can contain more than two hosts but the same host must not occur in both groups)
- Create two VM DRS Groups with one cluster node in each
- Setup an affinity rule between each set of Host DRS groups and VM DRS groups. Use the same 'MUST run on hosts in group' setting.

4.3.7 Real world considerations

Personally we only run physical Microsoft clusters at both my current and previous employers, so I've no real world experience of implementing the solutions covered here.

Microsoft have **very limited support** for clustered solutions – it must be running a certified architecture (which includes hardware and software configuration);

- For Windows 2003 this only includes two EMC storage devices (full details [here](#))
- For Windows 2008 the policy is more flexible (and is documented [here](#) and [here](#))

Exchange CCR - see this [blogpost at VMGuru.nl](#) for a good discussion around this, also this [VMware community post](#). Exchange 2010 (when used with DAGs) has it's own issues – see this [post](#).

SQL2005/2008 – you need a custom installation of VMtools – see [VMwareKB1021946](#)

4.4 Deploy and Maintain vCenter Server Heartbeat

Knowledge

- Identify the five protection levels for vCenter Server Heartbeat
- Identify the three server protection options for vCenter Server Heartbeat
- Identify supported cloning options

Skills and Abilities

- Install and configure vCenter Server Heartbeat
- Determine use cases for and execute a manual switchover
- Recover from a failover
- Monitor vCenter Server Heartbeat and communication status
- Configure heartbeat settings
- Configure shutdown options
- Configure application protection
- Add/Edit Services
- Add/Edit Tasks
- Edit/Test Rules
- Install/Edit Plug-ins
- Add/Remove Inclusion/Exclusion Filters
- Perform Full System and Full Registry checks
- Configure/Test Alerts
- Troubleshoot common vCenter Server Heartbeat error conditions

Tools & learning resources

- Product Documentation
 - [vCenter Server Heartbeat QuickStart Guide](#)
 - [vCenter Server Heartbeat Reference Guide](#)
- vSphere Client
- [Protecting vCenter with vCentre Heartbeat](#) (TA15, VMworld '09. Covers V5.5, not v6.3)
- [VMware vCentre Heartbeat Best Practices](#) (VM2674, VMworld '09. Covers V5.5, not v6.3)
- [Mike Laverick's four part series at TechTarget](#)
- [Paul Richard's blog entry on testing vCSHB](#)
- Julian Wood on [why vCSHB isn't an ideal solution](#) (about half way down the post).
- [VMware's community forums for vCSHB](#)
- VMware KB articles
 - [vCSHB prerequisites \(VMware KB1017587\)](#)
 - [VMworld 2010 lab 10 – vCSHB setup and configuration](#) (requires valid login)

If you work your way through the vCSHB Reference Guide you'll have covered every objective in the VCAP-DCA blueprint, so that's where I'd recommend you start. If you have time view the VMworld sessions for a bit of background and reinforcement. I went into a bit more detail on this objective as it's something I wanted to evaluate for my company, so there's more 'real world' issues covered which I doubt you'll need for the exam.

4.4.1 Basics and architecture

vCenter Server Heartbeat (vCSHB) is a business continuity product which aims to increase availability for vCentre, increasingly a crucial piece of the infrastructure puzzle. You can [download a 60 day evaluation copy from VMware](#) to play with in your lab. Under the hood this is a customised version of Neverfail, an availability product that's been around for years. It works by having two copies of vCentre with one active and one 'passive' and then monitoring (at various levels) to ensure the primary is working as expected.

Previously some people have run vCentre in a Microsoft cluster but this is no longer (was it ever?) supported ([VMware KB1014414](#)). If vCentre is virtual you can benefit from VMware HA but that only covers ESX host failures.

The three protection options;

- vCentre and a local SQL database
- vCentre only (used when vCenter DB is on a separate server to vCentre)
- SQL only (used when vCentre DB is on a separate server to vCentre)

NOTE: vCSHB doesn't support Oracle databases. If the Oracle database resides on a separate server you can use vCSHB to protect vCentre and use Oracle resilience features for the DB.

NOTE: vCSHB doesn't support clustered SQL. If you're using application resilience you have no need of vCSHB (you can use app features such as log shipping for DR).

Aside from protecting vCentre it also works with the various products that integrate with vCentre (assuming they're on the same server as vCentre);

- SRM
- Orchestrator
- vCentre Linked Mode (all ADAM data is replicated)
- Update Manager

Protection levels;

- Server – protects from hardware and guest OS failure
- Application – monitors Windows services
- Network – pings network locations (default gateway, DNS server and Global Catalog (default every 10 seconds) to determine isolation.
- Performance – monitors various performance metrics for predefined thresholds
- Data – can protect a local or remote SQL instance.

NOTE: This doesn't protect against data corruption – that's still replicated (corrupted) to the mirror site but most replication technologies (except Oracle's Dataguard) accept the same limitations.

Two modes;

- High availability (LAN)
- Disaster Recovery (WAN)

- Integrates with DNS to handle the IP address change from active to passive. See [VMware KB1008571](#) (Microsoft DNS) or [VMwareKB1008605](#) (BIND).
- Uses compression and automatically optimises bandwidth (1MB minimum recommended for vCSHB)
- May require static routes. [VMware KB1023026](#)
- Requires a file filter exclusion when using Orchestrator. See vCSHB Reference Guide.
- Updates ESX hosts (/etc/opt/vmware/vpxa/vpxa.cfg) to update the vCenter IP

Supported cloning operations

The process used to create the identical second vCentre server varies depending on your use of virtual or physical servers. In each case the server specs should match (CPU, memory, OS and service pack, network connections, server name, SID, domain membership etc). There are three possibilities;

- V2V (both virtual). Use vCentre cloning to create the secondary server. Recommended to use separate ESX hosts and separate vSwitches for resilience (but neither are enforced).
- P2V (Primary physical, secondary virtual). Use P2V tools such as VMware Converter.
- P2P (both physical)– use the vCSHB installation routine to clone the physical server (uses NTBackup) or you can use third party tools such as Platespin etc. NOTE: Drive letters and ACPI compliance must match for this configuration.

Switchover = a managed transition from active to passive. Typically used when upgrading vCSHB (to newer hardware for example)

Failover = a mitigation action due to an unplanned outage.

Real world issues (not relevant to the VCAP-DCA exam but worth noting)

- Licencing vCSHB seems rather confusing. If vCentre is on a separate server to your vCentre DB then do you need a separate vCSHB licence to cover the database server? If the database server is only standby does it need to be fully licenced? Do you need a second vCentre licence plus vCSHB or does the vCSHB purchase cover it? These questions have been asked on the vCSHB forums but the answers seem to vary.
- As vCSHB doesn't keep *everything* on the two servers in sync, some operations (such as applying Windows patches, AV updates etc) have to be duplicated on both servers. This adds to the maintenance and complexity of running vCentre. See [VMware KB1010803](#) for details of how to apply Microsoft patches.
- What if your SQL server is on a separate server to vCentre, but hosts multiple databases? Does vCSHB protect them all? In WAN mode this could massively increase bandwidth requirements for example as well as causing licencing issues.
- The vCSHB admin guide states that all protected applications should be installed before installing vCSHB. However that's contradicted by guidance in [VMware KB1014266](#) which states that SRM should be installed after vCSHB.
- Services such as Update Manager and Orchestrator are only protected if they run on the same server as vCentre. For many enterprises (who I think are the target market for vCSHB given its cost) this is unlikely to be case.

- If you're using vCentre Linked Mode and you want to either join or leave while protected by vCSHB you have to disable vCSHB protection, join/leave, then re-enable vCSHB protection. Full details in [VMware KB1022869](https://kb.vmware.com/s/article/1022869).

4.4.2 Installation and configuration

Preinstall steps

Before running the installation make sure you've got the following information;

- What level of protection (vCentre only etc)? Dictated by existing vCentre architecture.
- Where your second vCentre server will reside (LAN/WAN mode)
- IP addresses for VMware Channel interfaces
- Check the vCSHB QuickStart guide to ensure you match all prerequisites (disk space, network connections, etc). Ideally do these checks before cloning or you risk doing mitigation work twice (I didn't have enough disk space and then had to increase a system disk on both the primary and clone. Doh!)

NOTE: Exclude vCSHB directories from file level AV scanning (see the vCSHB Reference guide, installation chapter). The exclusions should be made on both active and passive servers.

NOTE: The primary network card MUST be the first listed in the binding order (Network, Advanced Settings)

NOTE: The secondary server will have the same name, IP and DNS settings as the primary. This means if you bring it onto the network you'll get IP conflicts, name conflicts and possibly DNS issues. The vCSHB Reference Guide (though NOT the QuickStart guide) advises the following steps (which worked for me) on the secondary server;

- Disable (or disconnect) the primary network card
- Set the IP address on the VMware Channel to something different to the VMware Channel IP used on the primary
- Ensure 'Register this connection with DNS' is unchecked on the VMware Channel interface (otherwise you're DNS entry for vCentre will be wrong)

Installation

Installation is largely a 'next, next, Finish' type install (and is covered step by step in the vCSHB Reference guide and the VMworld lab referenced at the bottom of this post). Overview;

- Run install on primary vCentre server
- Run again on the secondary (cloned) vCentre server
- Run separately on the vCentre database server (if it's separate from vCentre) and again on the vCentre database secondary server.

NOTE: For Windows 2008 there is an additional post-installation step to run the vCSHB Setup Completion program. There's a shortcut on the desktop – double click and follow the prompts.

Default ports used:

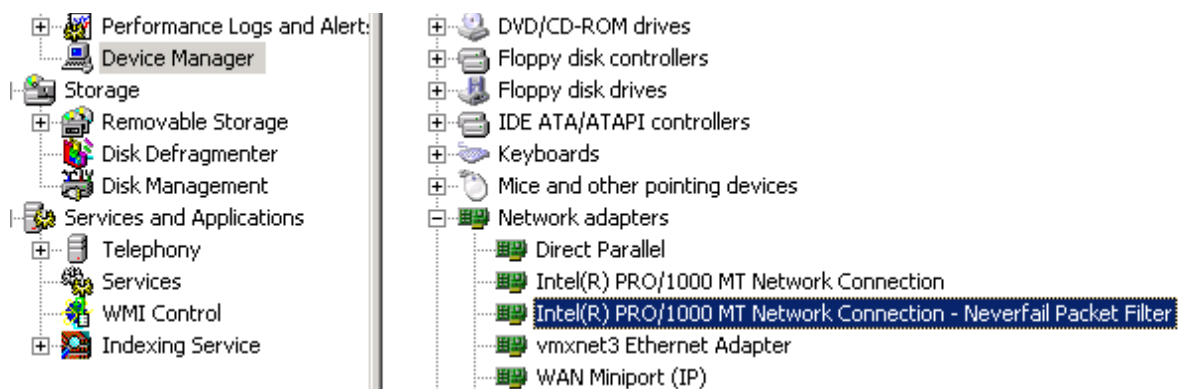
- 57348 – used for vCSHB heartbeat over VMware channel network
- 52267 – client tools connect to this port (Manage Server icon)

There are two Windows services installed. Check these when diagnosing issues;

- Neverfail SCOPE Data collector service (automatic)
- Neverfail Server R2 (automatic)

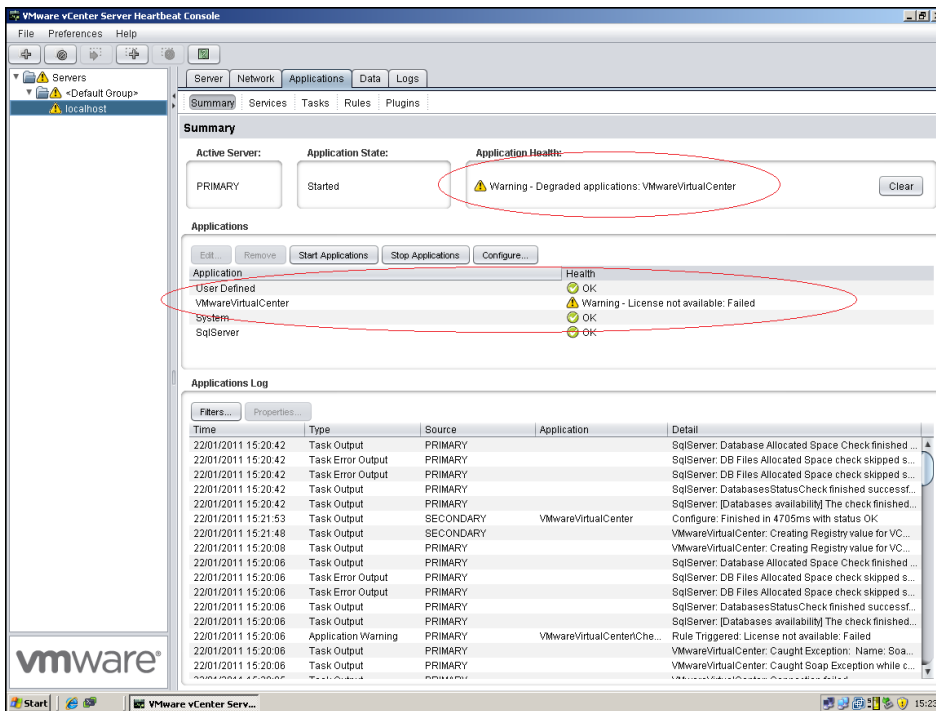
NOTE: The VMware VirtualCenter Server service will be changed from Automatic to Manual as it's now controlled by vCSHB.

A packet filter is installed and applied to all Primary network interfaces (but not the VMware Channel). You can check if the interface the filter is active on by looking at device manager (show hidden devices). It will be installed on both the primary and secondary;

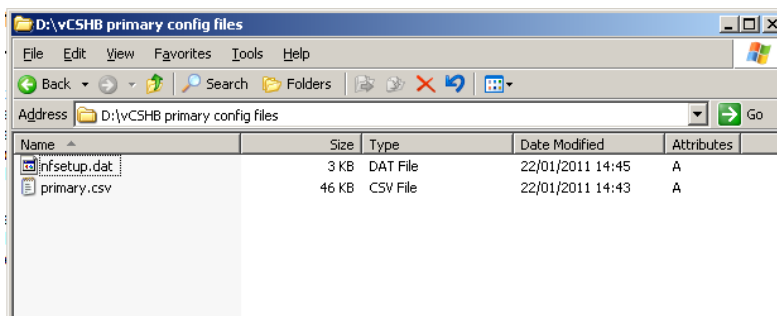


Post installation

After installation the vCSHB servers are replicating but haven't been given credentials to access vCentre. This causes an error about licencing (shown below). Go to the Applications tab, plug-ins, select and right click on the VirtualCentreNFPlugin.dll, choose Edit. Enter credentials with read access to vCentre.



NOTE: The install guide states that the primary config data (which is copied to a share on the secondary) can amount to GB's of data. This is only the case when choosing a P2P configuration where the application data is also backed up. For a V2V setup the files are tiny - mine were 49KB...



Upgrading/uninstalling (not part of VCAP-DCA syllabus)

It's possible to upgrade vCSHB without interruption, follow [VMware KB1014435](http://www.vmware.com/kb/1014435)

When uninstalling you simply run the uninstall routine on the primary and optionally on the secondary. You can delete the secondary if it's a VM. One issue to watch for is that both servers have the same NetBIOS name – either remove one server from the network or use the uninstall routine's option to rename one server. Full instructions can be found in [VMware KB1022877](http://www.vmware.com/kb/1022877).

4.4.3 Common Operations

Three utilities for managing;

- Manage Server. This is the main console used for day to day administration. This is the only tool available if installed on a separate client. Most admin operations can be performed on the primary OR secondary (changes are replicated to the other server).

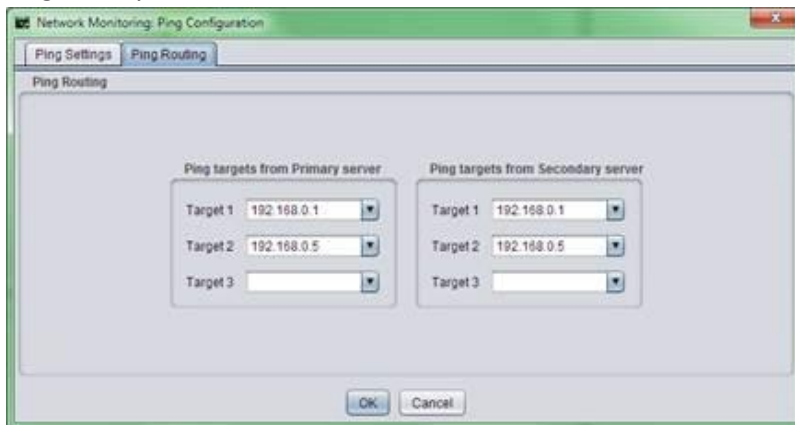
- Configuration Wizard. Used when changing server roles, IP addresses or network interfaces. Not needed for day to day administration. Only available on the servers.
- The tray utility. Provides 'at a glance' status along with right mouse button shortcuts. Only available on the servers.

Use the **Server tab** to complete the following tasks;

- Monitor server, replication, and heartbeat status
- Configure application startup and shutdown behaviour (Configure button)
- Perform switchover's from active to passive and vice versa
- Enable split brain avoidance (typically used with WAN setups). Monitoring tab, Configure.

Use the **Network tab** to complete the following tasks;

- Configure network ping settings. Particularly useful in WAN deployments where the ping targets may be different at the second site.



- Configure auto switchover if the client network fails (this default to 10 pings but is off by default). You might do this is you want HA to protect vCentre, in which case vCenter would take longer than ten pings to reboot after an ESX host failure.

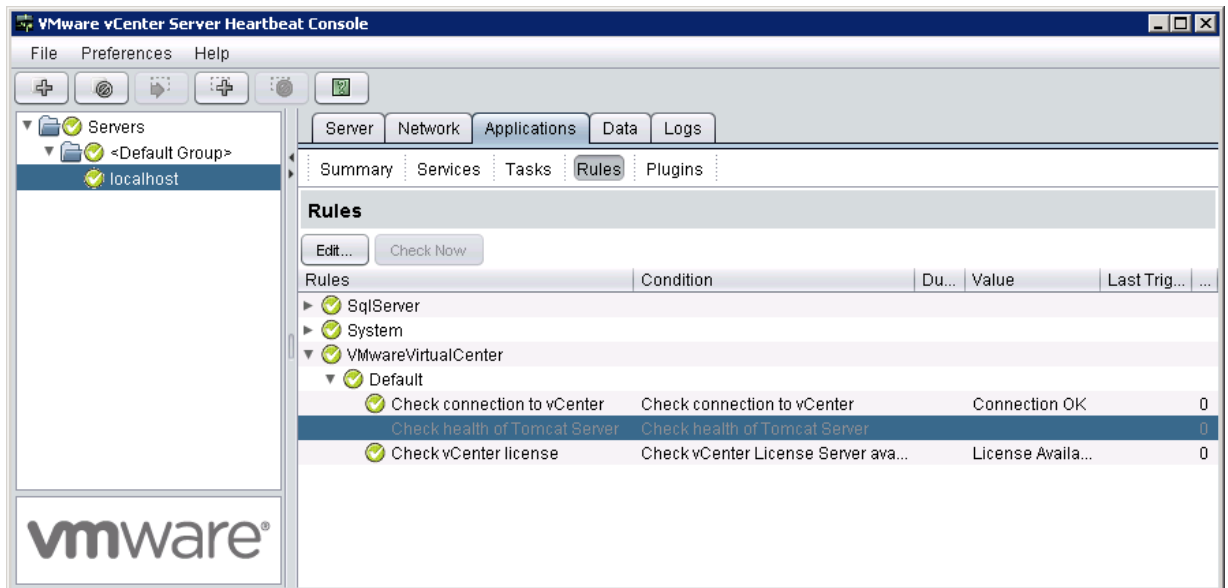
Use the **Applications tab** to complete the following tasks;

- Add/Edit/Remove protected applications. You choose to either monitor a service or manage it (which actively starts and stops the service) along with three failure actions to take;



- Add/Edit tasks. The available tasks are largely preconfigured and the user can only amend the timing. One example task is updating DNS when used in WAN mode. Another is Protected Service Discovery.

- Add/Edit Rules. Defined by the active plugins and used to protect performance using predefined metrics. A user can only enable/disable individual checks or configure timeouts and actions.



- Add/Edit Plugins. Not much to do here. There is a System plugin, vCentre plugin, and SQL plugin (if protecting SQL) by default and editing only offers authentication for vCentre and the option to protect index catalogue files for SQL.

Use the **Data tab** to complete the following tasks;

- Perform a full registry synchronisation – simply click the relevant button
- Perform a full system synchronisation – simply click the relevant button
- Add/Remove Filters. These let you include custom folders (a collection of PowerCLI scripts for instance) which are then replicated to the secondary server.
- Check replication queue lengths. This should be done prior to a switchover (best practice, in Reference Guide) or to understand WAN bandwidth requirements.

Use the **Logs tab** to complete the following tasks;

- Configure details of your SMTP server so email alerts are sent. Can also specify frequency and recipients based on alert level (red, yellow, green). Accepts mail servers which require authentication (unlike vCentre).
- Setup alerting. This can be email alerts or custom commands. To test simply click the 'Test Email Alerting' button.
- Review application logs

Recovering from a failover

1. Check vCSHB log files to determine status of both servers.
2. Identify cause of failover. Until all issues are resolved you should NOT try to restart vCSHB.
3. Ensure the secondary server is now active and working correctly (use systray icon or Manage Console utility)
4. Resolve issues with primary
5. (Optional) Switchover so the primary is active again.

Split Brain and vCSHB

As with many high availability solutions, split brain situations can happen. This can occur due to loss of the VMware Channel, power loss, or possibly misconfiguration of vCSHB. *Both* servers assume they are the active (or passive) server. **Data can be lost in this scenario.**

There is a split brain 'avoidance' option for vCSHB which lets it use the 'primary' network interfaces to test connectivity even if the VMware Channel fails. Enable the option 'Prevent failover if channel heartbeat is lost but Active server is still visible to other servers (recommended)' under the Server:Monitoring tab.. *Requires additional IP address to be configured on the primary network interfaces.* See the vCSHB Reference Guide p118 for full details.

To recover vCSHB from a split brain scenario the recommendation is to identify which server is most up to date (by looking at file timestamps) and reconfigure vCSHB to reset the roles. The full procedure is in [VMware KB1014405](#).

[Part three of Mike Laverick's vCSHB series](#) also specifically covers split brain. That article covers a known gotcha when using a remote vCentre database with vCSHB (covered by [VMware KB1027289](#)).

4.4.4 Troubleshooting

Check chapter 13 in the vCSHB Reference guide and Appendix A for a list of potential installation errors. Some useful VMware KB articles;

[VMware KB1008391](#) – log entries which may appear in the Application Event logs (in vCSHB console)

[VMware KB1008124](#) - Retrieving the VMware vCenter Server Heartbeat Logs and other useful information for support purposes.

[VMware KB1008572](#) – Troubleshooting vCSHB synchronisation errors. This links to other useful troubleshooting articles.

5 Perform Operations Maintenance

5.1 Implement and Maintain Host Profiles

Skills and Abilities

- Use Profile Editor to edit and/or disable policies
- Create sub-profiles
- Use Host Profiles to deploy vDS

Tools & learning resources

- Product Documentation
 - vSphere Datacenter Administration Guide
 - [VMware vSphere™ 4: Deployment Methods for the VMware® vNetwork Distributed Switch](#)
 - [VMware vNetwork Distributed Switch migration and configuration](#)
- vSphere Client
- [Technical Deep Dive – Host Profiles](#) (VM3433, VMworld '09)
- [VMware Management blog article on host profiles](#)

Host Profiles are a new feature to vSphere 4 but are only available to Enterprise+ licencees. As my company haven't yet found a need for Enterprise+ features I'd not really worked with them before so this section was new to me. Interestingly the main reference given in the blueprint is the Datacenter Administration Guide which has very little about host profiles. The ESX/ESXi configuration guides have a small section on host profiles but not much, so the best reference is probably the [VMware Host Profiles – Technical Overview whitepaper](#).

5.1.1 Host Profiles (VCP revision)

Basically host profiles are the equivalent of Microsoft's Group Policy, but for VMware hosts.

- Two primary uses
 - Ease deployment challenges (faster, more consistent)
 - Ongoing configuration control and audit reporting
- Policy options (determining how a configuration setting is applied)
 - Use a fixed configuration
 - Ask the user how to configure it
 - Use an intelligent policy (using one or multiple criterion)
 - Disregard a setting
- Works in a similar fashion to Update Manager;
 1. Create a baseline from a reference host.
 2. Attach the host profile to the hosts or clusters you want to configure
 3. Remediate (configure) the hosts or clusters
 4. Review compliance status
- Unlike VUM it can't remediate all the hosts in a cluster automatically (it won't put them into maintenance mode for you etc). You can attach a profile to the cluster but you have to apply to each host manually (this is largely because the host profile may require user input).
- Can only be used on vSphere hosts (not VI 3.x)
- Must be created using a reference host, or imported from a previously created host profile.

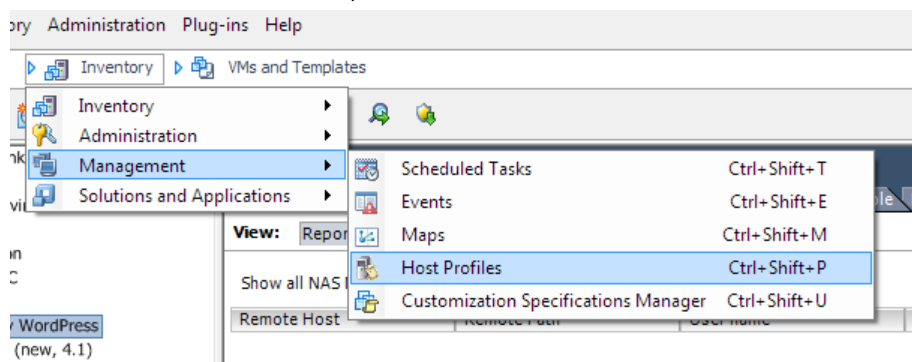
- Can be exported (in VMware Profile Format, *.vcp, which is XML content). Host Profiles are not shared using vCentre Linked Mode, you have to export/import to other vCentre instances.
NOTE: Administrator passwords aren't exported as a security measure.
- An ESX reference host can be applied to either ESX or ESXi. An ESXi reference host can ONLY be applied to another ESXi host.
- When updating a host using a host profile you have to manually put the host in maintenance mode first. This is a significant issue for some people (although if you're licenced for host profiles you've also got licences for vMotion and DRS so moving VMs off the host is potentially easier). Note that you need to enter maintenance mode even for trivial settings such as setting the time, timezone etc. Any setting which normally requires a reboot (changing service console memory for example) will still need a reboot.
- You must have both host profile privileges (create, delete, edit etc) AND privileges to configure the area in question (Networking, Storage etc) for the operation to be allowed.

REAL WORLD: When building a new ESX/ESXi host it will have a 60 day eval period with all features enabled so even if you don't have Enterprise+ licencing you can use host profiles for initial configuration.

REAL WORLD: Host profiles can't manage every configuration option you might want – for example they [can't configure jumbo frame support when using iSCSI](#), can't configure iSCSI HBAs (in fact host profiles won't even notice from a compliance point of view if iSCSI is enabled or not), can't set VM swapfile location, can't enable lockdown mode, and can't configure the load balancing algorithm for Pluggable Storage Path (PSP) among others. Some larger shops skip host profiles for this reason – if you have to use scripts at all you might as well script everything and have one place for all your settings being the logic. [Interesting post on SearchServerVirtualization.](#) They're also not that flexible – if you only want to configure storage for example (maybe because a scripted build configures networking) you can't turn off selected configurations. See this [blogpost](#) for details. Host Profiles are not hierarchical like MS Group Policy, so you can't apply multiple profiles to increase flexibility – it's one profile per host, and one profile only. Pretty limiting!

Once you've exported and reimported a host profile its connection to the original reference host is lost. This means you can't use the 'Update profile from reference host' option to refresh the profile if the source host changes.

You access host profiles under the Management node (using CTRL-SHIFT-P) or from a host's right mouse button context menu;



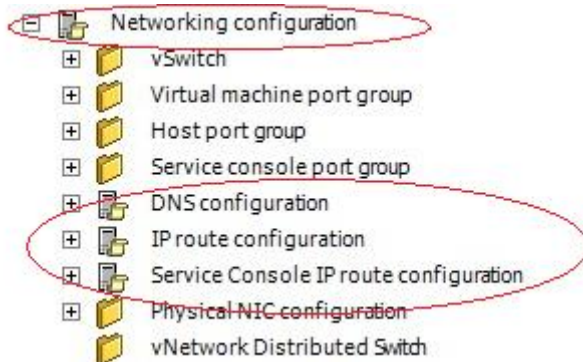
NOTE: Host Profiles will appear in the VI client regardless of your licencing level but if you try to create a host profile using a non-licenced host as the reference host you'll get an error message;

5.1.2 Sub Profiles

Each host profile consists of two parts;

1. Configuration settings. These determine the desired configuration of the host, for example 'connect vmnic0 and vmnic1 to vSwitch0'.
2. Compliance checks. These determine how the compliance check is validated, for example are vmnic0 & 1 connected to vSwitch0? If the host only has one pNIC the configuration won't apply correctly and this check will fail.

The configuration is split into various sub-profiles, each aligned to a functional area. Sub-profiles can also be nested (as you can see in the example for the networking sub-profile below;



The main (default) sub-profiles are;

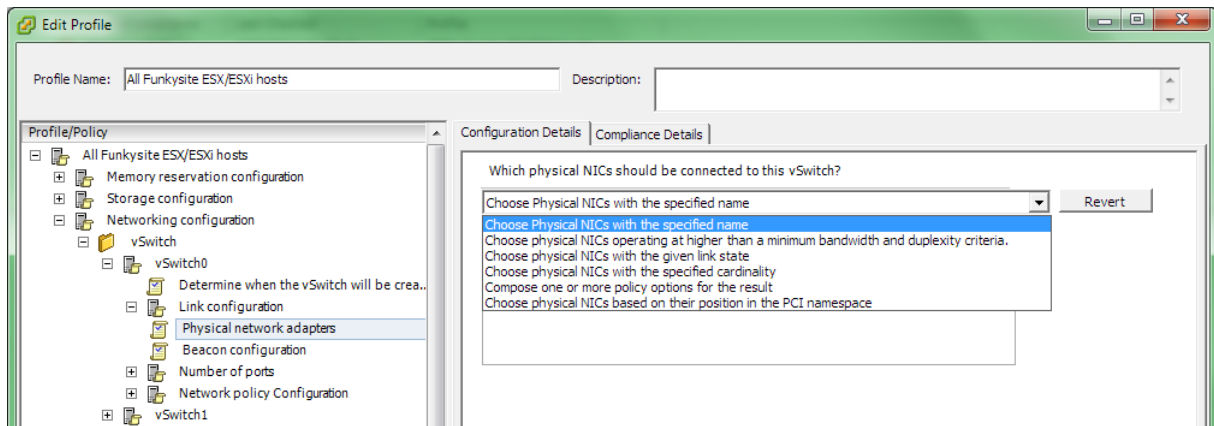
- Memory reservation (service console)
- Network configuration
- Storage configuration
- Date and time
- Firewall
- Security
- Users and user groups – these are NOT captured from the reference host and must be created manually. Likewise while a host profile can add users/group it will NOT remove extras or alert if extra users are added manually (Technical Overview p19).
- Authentication (new to vSphere 4.1). This can be used when AD integrated authentication is enabled on a host.

You'll probably need to create extra sub-profiles of your own, based on your infrastructure and what you need host profiles to do. If you wanted to mount four NFS datastores on every host for example you'd have to create an extra three sub-profiles under Storage -> NFS storage configuration. Each sub-profile would hold details for one NFS export (IP, path, read/write etc).

NOTE: Not all options can be 'ticked' for compliance checking (or not)

5.1.3 Host Profile policies

Policies allow host profiles to be more flexible (compared to fixed values) by specifying what to do instead of how to do it. A typical use is networking. Rather than specifying just the number of NICs for a vSwitch you may want to be more specific and choose the NICs based on name, PCI namespace, bandwidth etc (this is useful for blade environments in particular). Policies allow you to do this.



5.1.4 Deploying vDistributed switch using host profiles

Another typical use for host profiles is deploying the vDistributed Switch. There are two variations which both (as usual) require you to configure a reference host with a vDS first;

1. Configure a host with a single vSS (it needs network connectivity to join the vDS)
2. Create a vDS using the VI client connected to vCentre, with no hosts attached.
3. Add the host to the vDS, then delete the vSS on the host.
4. Create a host profile and apply to the remaining hosts
5. Reconfigure VMs to use the vDS instead of the vSS

The alternative method is very similar but minimises the 'per VM' configuration. At step 3 simply leave the vSS in place and continue to create and apply the host profile. On the vDS use the 'Migrate virtual machine networking' option to seamlessly migrate the VMs. Finally remove the vSS from the reference host, recreate the host profile and reapply. Voila!

Full details of this process and more details about the vDS can be found in [VMware's vNetwork Distributed Switch migration and configuration whitepaper](#).

5.1.5 Host profiles with PowerCLI

What about PowerCLI cmdlets for host profiles? [Good example at Damian Karlson's site](#), or [Hal Rottenberg's guide to using host profiles with PowerCLI](#).

Create a baseline;

- Get-VMHostProfile
- New-VMHostProfile
- Import-VMHostProfile
- Export-VMHostProfile



Attach a baseline to a host;

- Apply-VMHostProfile
- Remove-VMHostProfile
- Set-VMHostProfile

Remediate;

- Test-VMHostProfileCompliance





You can view from the host's summary tab whether it's in compliance with a profile;

Host Configured for FT: No 
Active Tasks:
Host Profile: All Funkysite ESX/ESXi hosts
Profile Compliance:  Noncom

Fault Tolerance

Fault Tolerance Version: 244038

Commands

-  New Virtual Machine
-  Enter Maintenance Mode
-  Reboot
-  Shutdown

For vSwitch vSwitch0 network policy property spec.policy.security doesn't match
For vSwitch vSwitch1 network policy property spec.policy.security doesn't match
Port group Production exists on vSwitch vSwitch0. Expected to be on vSwitch1
Port group VMkernel exists on vSwitch vSwitch0. Expected to be on vSwitch1
Device vmnic2 does not exist
Ruleset aam doesn't match the specification
Ruleset faultTolerance doesn't match the specification
Ruleset ntpClient doesn't match the specification
List of NTP servers doesn't match the specified list

5.2 Deploy and Manage Complex Update Manager Environments

Knowledge

- Identify firewall access rules for Update Manager

Skills and Abilities

- Determine use case for, install and configure Update Manager Download Service
- Configure a shared repository
- Configure smart rebooting
- Manually download updates to a repository
- Perform orchestrated vSphere upgrades
- Create and modify baseline groups
- Troubleshoot Update Manager problem areas and issues
- Generate database reports using MS Excel or MS SQL
- Upgrade vApps using Update Manager

Tools & learning resources

- Product Documentation
 - [VMware vCenter Update Manager Installation and Administration Guide](#)
- vSphere Client
- vmware-umds
- [VMware Update Manager forums](#)
- [Update Manager 4 Performance and Best Practice](#) (VM1882, VMworld '09)
- [Update Manager Performance and deployment best practices](#)
- [Sean Crookston's study notes](#)

Most people have used Update Manager to some degree so this objective is probably one of the easier ones to cover. The VUM Administration Guide covers pretty much everything on the VCAP-DCA blueprint and should be your first stop for study (apart from this blog obviously!).

Not listed in the blueprint (at least in this section) is the PowerCLI cmdlets for using Update Manager. Section 8 only lists 'Installing the Update Manager PowerCLI cmdlets' but if you have time it's probably worth giving them a spin.

5.2.1 Update Manager Basics (VCP revision)

The exam topics assume a certain amount of knowledge as Update Manager is on the VCP syllabus.

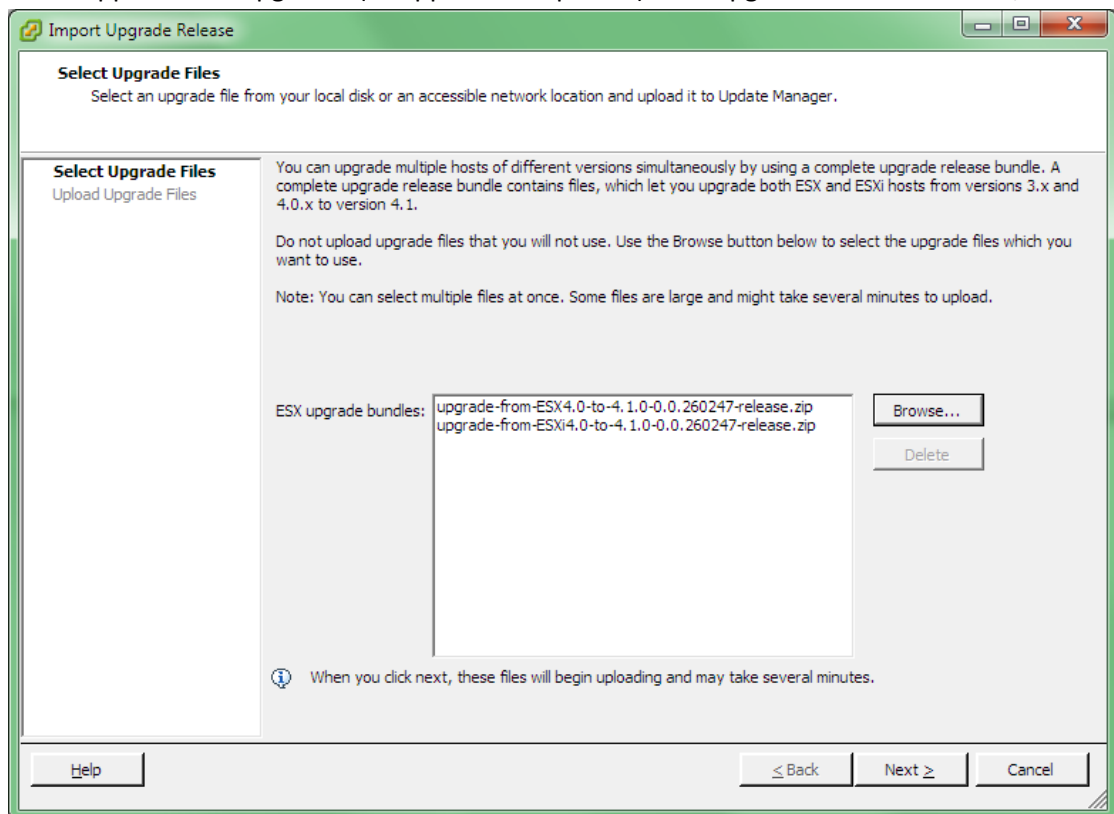
A quick recap;

- Installs as a plugin to vCentre
- Downloaded as part of the vCentre package
- Once the server component is installed you have to add the plugin to any VI client installations you use.
- Distinguishes between 'patches and security updates' vs 'product upgrades'.
- Typical lifecycle;
 - Download patches (from centralised repository or Internet)
 - Create baselines
 - Scan hosts (and/or guest OS)
 - View compliance reports

- Remediate (apply patches)
- Can patch;
 - ESX/ESXi hosts
 - Virtual hardware
 - VMware Tools
 - Virtual appliances
 - Guest OSs (Windows and Linux)
 - Third party components such as the Cisco Nexus 1000v distributed switch, Xsigo HCA drivers, and EMCs PowerPath.
- Patching guest OSs requires an agent to be installed to the guest. This is done automatically the first time a guest is scanned for patch compliance or can be done manually if required.
- Patches are downloaded accordingly to a defined schedule (default once a day)

New features in v4.1

- When installing for vSphere v4.1 you'll have to create a 32bit DSN (vCentre 4.1 requires a 64bit OS but VUM is a 32bit app). See [VMware KB1010401](http://www.vmware.com/kb/1010401) for details of how to do this.
- Patch recalls, actionable alerts, and new fixes are checked (and notified) via another defined schedule (default once an hour)
- Now supports host upgrades (as opposed to updates). Add upgrade binaries to VUM;



Recent Tasks

Name	Target	Status
Upload offline patches	VirtualCentre.thefunkysite.com	In Progress
Import host upgrade release	VirtualCentre.thefunkysite.com	In Progress
Upload offline patches	VirtualCentre.thefunkysite.com	In Progress
Import host upgrade release	VirtualCentre.thefunkysite.com	In Progress
Check new notifications	VirtualCentre.thefunkysite.com	Completed
Initiated guest OS shutdown	Xangati Appliance	Completed

Misc

- Be careful when patching a host containing your virtual vCenter server. Prior to vSphere 4.1 this could cause remediation to fail. Still requires DRS enabled for automatic remediation.
- VMs need to be powered on to report VMtools status correctly.
- Quite a few VMs may show as 'incompatible'. This can happen when no VMtools are installed or not managed via vSphere (appliances, virtual ESX hosts with no VMtools packages etc)

5.2.2 Firewall access rules for VUM

- Firewall ports: 8084 (SOAP), 9084 (patch repository)
- See the VUM Administration Guide (p.68) or [VMwareKB1004543](#) (which differs slightly from the admin guide)

5.2.3 Managing the patch repository

Four ways to populate VUM with patches;

1. Directly from VMware/Shavlik via the Internet (default)
2. Third party patch repositories (Cisco for example)
3. Via a manual patch download (offline bundle)
4. From an Update Manager Download Service (UMDS) repository

NOTE: Some drivers are provided as an offline bundle which can be deployed with VUM, but some need to be manually added to an ESX host using esxupdate or (PowerCLI). See this [post about Broadcom drivers at Julian Wood's site](#) for a good example.

5.2.3.1 Using third party patch repositories

Add a third party URL using the 'Add Patch Source' button on the VUM Configuration tab (see screenshot). See [Chad Sakacc's YouTube video of using VUM to install EMC PowerPath/VE](#).

5.2.3.2 Using offline bundles

An offline bundle (in .ZIP format) can be downloaded from VMware (typically linked to a VMware KB article) and imported manually. Use the 'Import Patches' button on the VUM Configuration tab (see screenshot)

Administration for VirtualCentre.thefunkysite.com

Configuration Events Notifications Patch Repository Host Upgrade Releases

Patch Download Settings Compliance View

Patch Download Sources Used for third party repositories

Direct connection to Internet - download new patches at intervals specified in **Patch Download Schedule** or click the **Download Now** button below Add Patch Source...

Enabled	Patch Type	Component	Patch Source	Description	Connectivity Status
<input checked="" type="checkbox"/>	VMware	ESX	https://hostupdate.vmware.com/software/...	Download ESX 4x patc..	Connected
<input type="checkbox"/>	VMware	ESX	https://www.vmware.com/PatchManagem..	Download ESX 3x patc..	Connected
<input checked="" type="checkbox"/>	Custom	ESX	https://hostupdate.vmware.com/software/...	Download ESX 4x patc..	Connected
<input type="checkbox"/>	Linux	VMs	https://xml.shavlik.com/unix/	Download Linux VM pa..	Connected
<input type="checkbox"/>	Windows	VMs	https://xml.shavlik.com/data	Download Windows V..	Connected

Use a shared repository [What's this?](#)

Note: you can also Import Patches manually from a local .zip file

Proxy Settings Used for offline bundles

Use proxy Proxy requires authentication

Proxy: Username:

Port: Password:

5.2.3.3 UMDS

Sometimes the VUM server doesn't have Internet access (due to security restrictions, infrastructure limitations etc). Two deployment models for Update Manager;

- Air gap model (Update Manager is isolated from other networks)
- Semi air gap model – Update Manager doesn't have internet connectivity but is connected to a server which does.

The Update Manager Download Service (UMDS) can address both these situations.

Installation

- Like VUM, install a database before running install. Requires a 32bit DSN if on 64 bit server. For lab installs you can use the bundles SQL Express 2005 (no prior db creation required).
- Installed on a separate Windows server to VUM. Can be x32 or x64 Windows OS.
- Install from vCenter install DVD, /umds directory. Run VMWARE-UMDS.EXE (not the .msi)
- Configured via the command line (no GUI). Default install path;
 - C:\Program Files\VMware\Infrastructure\VMware Update Manager\vmware-umds.exe

Post installation steps

- Configure patches to download;
 - vmware-umds --set-config --enable-host true --enable-win false --enable-lin false
 - vmware-umds --s -h 1 -w 0 -l 0
- Unlike VUM, UMDS actually downloads all the chosen patches immediately rather than waiting for a remediation action. Mine downloaded 10Gb of patches (host patches only).

- You can't specify via the command line the version of ESX that you want patches for (3.x, 4.x, ESX/ESXi etc) so the patch repository may be significantly larger than the VUM equivalent. Downloading all these patches can also take significant time depending on your internet connection. You can exclude V13 patches by following [VMware KB1015663](#).
- Download patches
 - `vmware-umds -download` (or `vmware-umds -D`)
NOTE: Quite a few of the patches didn't download first time for me. [VMware KB1026154](#) describes how to use an undocumented switch to increase timeouts.
 - Run this regularly (as a scheduled Windows job) to ensure patches are up to date
- Export patches. *NOTE: You need to export to a directory, not just a drive. Exporting to D:\ won't work, but exporting to D:/VUM will.*
 - `vmware-umds --export --export-store <directory/URL>`
 - Can be exported to a directory (C:\UMDSData) or a webserver (<http://server/folder>)
 - NOTE: [VMware KB1019288](#) gives details for configuring Apache or IIS to work as a remote source Of VUM patches (or use the Admin Guide p158-160).

Optional configuration

- Change patch download location;
 - `vmware-umds.exe --set-config --patch-store <directory path>`

See chapter ten of the VUM Admin guide for full details of installing and configuring UMDS.

5.2.4 Orchestrated upgrades and baseline groups

An orchestrated upgrade is simply a baseline group which contains multiple upgrade baselines. VUM is smart enough to apply them in the correct order and reboot accordingly;

- Upgrade VMTools to match host
- Upgrade virtual h/w to match host
NOTE: You can't use an Orchestrated upgrade on VMs where VMtools isn't installed or is managed by a third party (virtual appliances, virtual ESX hosts etc).

5.2.5 vApps (upgrading and smart reboot)

Smart rebooting is simply the ability for vCentre to restart vApps in a defined order, typically used when one server has a dependency on another (webserver and db server for example);

- Restarting individual VMs based on the vApp priority settings
- Configured via VUM Admin view, Configuration, vApp Settings
- ON by default

5.2.6 Reporting on VUM

There is no built in reporting so you'll have to use either SQL queries or Excel to produce reports. I could follow the example in the admin guide but it didn't produce an overly meaningful report – I suspect you need reasonable SQL skills to do this any justice in the real world;

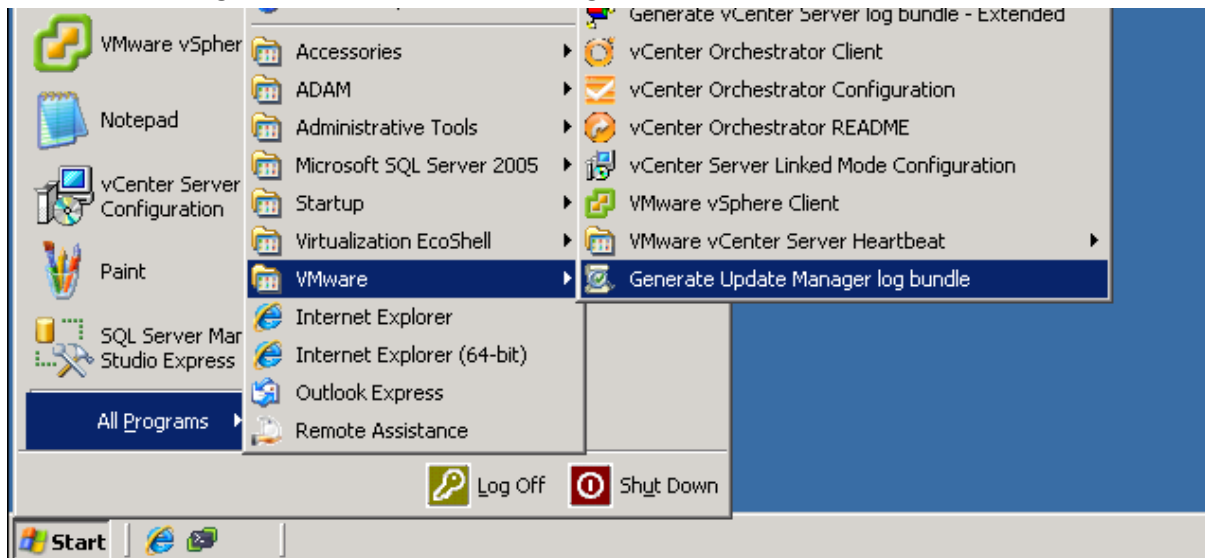
- Create an ODBC DSN (pointing to the VUM database) on the machine with Excel installed (any version)
- In Excel go to Import -> External Data source (pre 2010) or Data tab, select 'From Other Sources' and then either 'From SQL server' or 'from Microsoft Query' (Excel 2010).

5.2.7 Troubleshooting

Chapter 18 of the Update Manager Administration Guide lists a variety of troubleshooting steps you can take. General troubleshooting steps;

- Check the VMware Update Manager service is started (restart if necessary)
- Re-enable the plugin via the VI client. Sometimes you have to disable and re-enable.

Generate VUM log bundles from the host running VUM;



Determining patch level of a host (kernel build number vs vpxa build level etc ([VMware KB102154](#)));

- vmware -l = the base build (vSphere 4.0 Update 1 for example)
- vmware -v = the patch level (as reported in the VI client)

6 Perform Advanced Troubleshooting

6.1 Configure, Manage, and Analyse vSphere Log Files

Knowledge

- Identify vCenter Server log file names and locations
- Identify ESX/ESXi log files names and locations
- Identify tools used to view vSphere log files

Skills and Abilities

- Generate vCenter Server and ESX/ESXi log bundles
- Use vicfg-syslog to configure centralized logging on ESX/ESXi Hosts
- Test centralized logging configuration
- Configure the vMA appliance as a log host
- Use vilogger to enable/disable log collection on the vMA appliance
- Use vilogger to configure log rotation and retention
- Analyze log entries to obtain configuration information
- Analyze log entries to identify and resolve issues

Tools & learning resources

- Product Documentation
 - [vSphere Management Assistant Guide](#)
 - [vSphere Command-Line Interface Installation and Scripting Guide](#)
 - [vSphere Datacenter Administration Guide](#)
- vSphere Client
- vicfg-syslog, vilogger
- [Eric Sloof's Advanced Troubleshooting presentation](#) at the Dutch VMUG
- [VMware whitepaper on Troubleshooting Performance issues](#)

I'm covering the troubleshooting objectives last while preparing for the VCAP-DCA - it seems like the logical thing to do. Learn all the material then play with it, break it, fix it, recreate it etc. Practice makes perfect! I've been using the [Trainsignal's Troubleshooting for vSphere course](#) but the official VMware Troubleshooting course has been [getting good feedback](#).

6.1.1 vCenter log files

Located in;

- %ALLUSERSPROFILE%\Application Data\VMware\VMware VirtualCenter\Log (W2k3)
- C:\ProgramData\VMware\VMware VirtualCenter\Log (W2k8)

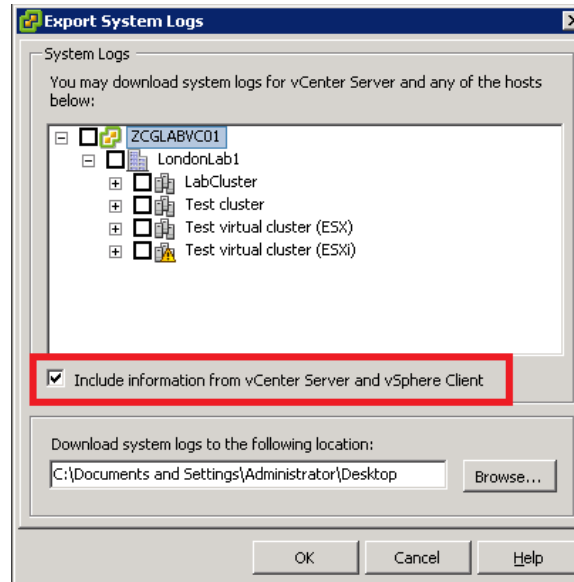
Available logs;

- sms.log Storage Management Service
- vpxd-xxxx.log vCenter logs
 - vpxd-xxxx.log.gz are archived logs. You have to unzip them to see contents.

You can change the logging level (which defaults to 'normal') by going to vCenter Server Settings -> Logging Options. This [VMwareKB](#) describes how to enable trivia logging in vCenter (even if vCenter isn't running) although this may have a performance impact and should only be used temporarily while diagnosing issues.

There are numerous ways to do this;

- On the vCenter server console (via RDP, ILO etc) go to Start -> Program Files -> VMware -> Generate vCenter Server log bundle (or log bundle enhanced, which contains ???)
- From the VI client go to File -> Export -> System logs. If you only want the vCenter logs (rather than logs for ESX/i hosts) leave all the tickboxes unchecked except the 'Include information from vCenter...'



- If you're feeling flash you can do it via PowerCLI (NOTE: This can take quite a while to run!)
 - `Connect-viserver <vCenter> | get-log -Bundle -DestinationPath d:\ -Server <vCenter>`

You might also want to check the VI client logfiles. They can be found at;

- C:\Documents and Settings\\Local Settings\Application Data\VMware\vpv (Windows XP)
- C:\Users\\AppData\Local\VMware\vpv (Vista and Windows 7)

6.1.2 ESX/ESXi logfiles

The logs on an ESX/i host are scattered around but most of the commonly used ones are;

- /var/log
- /var/log/vmware
- /var/log/vmware/vpxa
- /var/log/vmware/hostd

[VMwareKB1021800](#) details the logfiles for an ESX host.

[VMwareKB1021801](#) details the logfiles for an ESXi host.

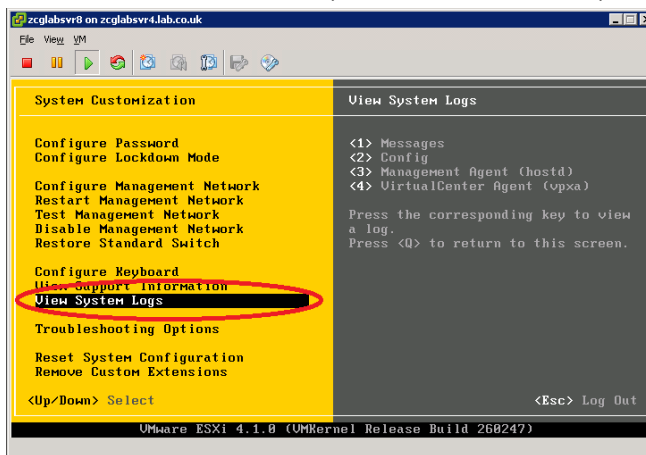
Like vCenter you can generate a log bundle for troubleshooting purposes;

- Connect the VI client to vCenter and follow the same instructions above (but select the host or hosts you want to generate a bundle for)
- Connect the VI client directly to a host and use File -> Export -> System Logs
- Using vm-support on the ESX/i console.
- Use PowerCLI (note the syntax is slightly different)
 - `Get-Log -VIHost <hostname> -Bundle -DestinationPath d:\`

6.1.3 Tools used to view logs

There are various ways to view logfiles both in the GUI and via command line;

- Connect the VI client to vCenter and go to Administration -> System Logs to see vCenter (not host) logs. You can search these logs using the search box in the top right.
NOTE: By default this only shows the logs since the last restart of the vCenter service.
- Connect the VI client directly to a host and go to Administration -> System Logs.
- Connect to a host using SSH and use standard Linux commands - cat, more, tail, and grep
 - A useful command is `'tail -f'` which watches a file - you see updates in real time
- Point a Web browser to a host
 - `http://<hostname>/host`
- Use the DCUI on an ESXi host (if not in lockdown mode). Then choose the log you want to see;



6.1.4 Centralised logging

VMware's migration to ESXi is inevitably why this objective exists, as by default it only logs into memory so all logs are lost when an ESXi host reboots. This means everyone is going to want to centralise their logging. [Trainsignal's Troubleshooting for vSphere course](#) has a great video for this.

There are two options;

1. Use vilogger (preferred solution)
2. Configure a syslog server (if you don't have a vMA server or have an existing syslog server)
 - a. Configure the central syslog server
 - b. Configure each host to send logs to the central server

Configuring vMA using vilogger

This is by far the quickest solution - for details of using vilogger refer to section 8.3.

NOTE: Occasionally when vilogger is configured the host authentication gets logged in vCenter, resulting in hundreds of logged events every minute. If you have even a small number of hosts (I tried it with just five) it can quickly fill your vCenter database with event logs. In a lab environment using SQL Server Express (with its 4GB limit) this can quickly cause your vCenter server to go offline. The solution is to disable vilogger and clear down the database tables (see [VMwareKB1025914](#)) and it's also been discussed in this [VMware community thread](#).

Description	Type	Date Time
User Administrator logged out	info	3/12/2011 2:55:16 PM
User Administrator@192.168.8.40 logged in	info	3/12/2011 2:55:16 PM
User Administrator logged out	info	3/12/2011 2:55:16 PM
User Administrator@192.168.8.40 logged in	info	3/12/2011 2:55:16 PM
User Administrator logged out	info	3/12/2011 2:55:15 PM
User Administrator@192.168.8.40 logged in	info	3/12/2011 2:55:15 PM
User Administrator logged out	info	3/12/2011 2:55:15 PM
User Administrator@192.168.8.40 logged in	info	3/12/2011 2:55:15 PM
User Administrator logged out	info	3/12/2011 2:55:15 PM
User Administrator@192.168.8.40 logged in	info	3/12/2011 2:55:15 PM
User Administrator logged out	info	3/12/2011 2:55:15 PM
User Administrator@192.168.8.40 logged in	info	3/12/2011 2:55:15 PM
User Administrator logged out	info	3/12/2011 2:55:15 PM
User Administrator@192.168.8.40 logged in	info	3/12/2011 2:55:15 PM
User Administrator logged out	info	3/12/2011 2:55:15 PM
User Administrator@192.168.8.40 logged in	info	3/12/2011 2:55:15 PM
User Administrator logged out	info	3/12/2011 2:55:15 PM
User Administrator@192.168.8.40 logged in	info	3/12/2011 2:55:15 PM
User Administrator logged out	info	3/12/2011 2:55:15 PM
User Administrator@192.168.8.40 logged in	info	3/12/2011 2:55:15 PM
User Administrator logged out	info	3/12/2011 2:55:15 PM
User Administrator@192.168.8.40 logged in	info	3/12/2011 2:55:15 PM

Configuring vMA as a central syslog server

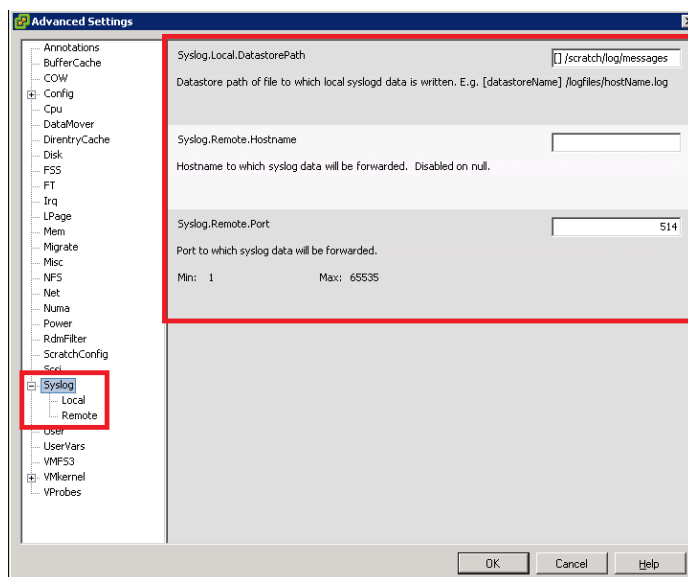
Setting up vMA as a syslog server involves the following steps;

1. Open the vMA firewall to allow incoming traffic (and make this persistent)
2. Amend the built-in syslog service (not vilogger) to receive logs from remote hosts
3. OPTIONAL: See [Simon Long's blogpost](#) which covers adding extra hard disks to cope with additional logging capacity.

Configuring the hosts to send logs to a syslog server

Configuring ESXi

- Using vicfg-syslog for an ESXi host (NOTE: vicfg-syslog only works for ESXi)
 - vicfg-syslog -s <hostname of syslog server>
 - Can be configured to log to any syslog server
- Setting logging in the Advanced Settings
 - Go to host Configuration -> Software -> Advanced Settings -> Syslog and specify the hostname of your syslog server



NOTE: Both vicfg-syslog and editing the Advanced Settings above modify /etc/syslog.conf under the hood although if you want to do it manually there are extra steps required - [VMwareKB1016621](#) goes into full detail.

Configuring ESX ([VMwareKB1005030](#))

- Editing /etc/vmware/syslog.conf for an ESX host
 - Add `'*. * @<IP_address_of_syslog-server>'` to the bottom of the file
- Open the ESX firewall to allow syslog traffic
 - `esxcfg-firewall -o 514, udp, out, syslog`
- Restart the syslog service
 - `service syslogd restart`

VMworld 2009 session VM3325 [vSphere and ESX/I logfiles 101 & 102](#) also covers configuring both ESX and ESXi to log to a syslog server (around the 30 minute mark).

6.1.5 Analyse log entries

This is a tough objective as the format and content of every logfile is different - the best study method is to simply look at every file you can find that might be useful. A good start is watching VMworld 2009 session VM3325 [vSphere and ESX/I logfiles 101 & 102](#) (requires free registration. Audio is crap!). Two main objectives;

- Resolve issues
 - Search for warnings, errors, etc
- Obtain configuration information
 - Need to know what you're looking for and which logfile to search
 - Use `'grep <string> *'` to search all files in the current directory for a given string (such as 'warning') if you're not sure which logfile to search (could take a long time in large directories).

NOTE: Many logfiles are time stamped using UTC - if you're host isn't configured to use UTC this may make correlating events and logs difficult.

6.2 Troubleshoot CPU and Memory Performance

Knowledge

- Identify resxtop/esxtop metrics related to memory and CPU
- Identify vCenter Server Performance Chart metrics related to memory and CPU

Skills and Abilities

- Troubleshoot ESX/ESXi Host and Virtual Machine CPU performance issues using appropriate metrics
- Troubleshoot ESX/ESXi Host and Virtual Machine memory performance issues using appropriate metrics
- Use Hot-Add functionality to resolve identified Virtual Machine CPU and memory performance issues

Tools & learning resources

- Product Documentation
 - vSphere Resource Management Guide
 - [vSphere Command-Line Interface Installation and Scripting Guide](#)
- vSphere Client
- vSphere CLI
- resxtop/esxtop

This is another objective that's hard to quantify – experience will be the main requirement! There are some great general purpose resources out there;

- [Performance Troubleshooting in Virtual Infrastructure](#) (TA3324, VMworld '09)
- [Trainsignal's Troubleshooting for vSphere course](#)
- [Eric Sloof's Advanced Troubleshooting presentation](#) at the Dutch VMUG
- [Understanding Host and Guest Memory Usage...](#) (TA2627, VMworld '09)
- Useful [command line cheat sheet](#)
- [VMware whitepaper on Troubleshooting Performance issues](#)

Note that resxtop (built in to the vMA) does not offer the 'replay' mode available in ESX classic.
Source: VMworld session MA6580, vMA Tips and Tricks.

6.2.1 Identify esxtop and vCenter metrics related to memory and CPU

[See section 3.1.3 in the Performance chapter for a list of metrics to check.](#)

6.2.2 Troubleshoot ESX/ESXi host and VM memory performance issues using appropriate metrics

Read the ESXTOP bible which covers metrics to look for and typical values for various problems.

Remedial actions

- Verify that VMtools is installed in every VM (otherwise the balloon driver won't be active and swapping will occur)
- Verify that the balloon driver is active (look for MCTL? in esxtop)
- Check for resource limits or insufficient reservations (both on the VM and any resource pools)
- If the host is in a cluster, enable DRS if not already enabled
- Check for VMs with a high reservation (compared to active memory). This may be a sign that the VM is oversized and memory is being wasted (be careful with Java and Oracle)

- Add more physical memory to the host

NOTE: On Xeon 5500 (Nehalem) hosts TPS won't show much benefit until you overcommit memory (assuming you use large memory pages in both the guest OS and ESX ([VMwareKB1021095](#)))

NOTE: vim-cmd is only available on ESXi and ESX hosts but NOT in the vMA.

6.2.3 Troubleshoot ESX/ESXi host and VM CPU performance issues using appropriate metrics

Read the ESXTOP bible which covers metrics to look for and typical values for various problems.

Remedial actions

For clusters

- enable DRS if not already enabled. This may alleviate hotspots.
- If DRS is already enabled
 - add hosts
 - check threshold setting. Setting a higher threshold may balance load more effectively.

For all hosts

- Enable CPU saving features such as TSO, h/w iSCSI initiators, TSO enabled pNICs, large memory pages, newer vNIC drivers (VMXNET3) etc
- Ensure VMtools is enabled in all VMs
- Right size VMs which are incorrectly allocated vSMP for single threaded apps

6.2.4 Use Hot-Add functionality to resolve identified VM CPU and memory performance issues

Covered in section 3.1

6.3 Troubleshoot Network Performance and Connectivity

Knowledge

- Identify virtual switch entries in a Virtual Machine's configuration file
- Identify virtual switch entries in the ESX/ESXi Host configuration file
- Identify CLI commands and tools used to troubleshoot vSphere networking configurations
- Identify logs used to troubleshoot network issues

Skills and Abilities

- Utilize net-dvs to troubleshoot vNetwork Distributed Switch configurations
- Utilize vicfg-* commands to troubleshoot ESX/ESXi network configurations
- Configure a network packet analyzer in a vSphere environment
- Troubleshoot Private VLANs
- Troubleshoot Service Console and vmkernel network configuration issues
- Troubleshooting related issues
- Use esxtop/resxtop to identify network performance problems
- Use CDP and/or network hints to identify connectivity issues
- Analyze troubleshooting data to determine if the root cause for a given network problem originates in the physical infrastructure or vSphere environment

Tools & learning resources

- Product Documentation
 - [ESX Configuration Guide](#)
 - [ESXi Configuration Guide](#)
 - [vSphere Command-Line Interface Installation and Scripting Guide](#)
- vSphere Client
- vSphere CLI
 - vicfg-*, net-dvs, resxtop/esxtop
- [Eric Sloof's Advanced Troubleshooting presentation](#) at the Dutch VMUG
- [VMware whitepaper on Troubleshooting Performance issues](#)
- [Trainsignal's Troubleshooting for vSphere course](#)
- [TA6862 - vDS Deep Dive - Managing and Troubleshooting](#) (VMworld 2010)

6.3.1 Identify virtual switch entries in a VMs configuration file

Contains both vSS and vDS entries;

```
35 sched.scsi0:0.throughputCap = "off"
36 sched.scsi1:0.shares = "normal"
37 sched.scsi1:0.throughputCap = "off"
38 ide1:0.present = "true"
39 ide1:0.fileName = ""
40 ide1:0.deviceType = "atapi-cdrom"
41 ethernet0.present = "true"
42 ethernet0.virtualDev = "e1000"
43 ethernet0.networkName = ""
44 ethernet0.addressType = "vpx"
45 ethernet0.generatedAddress = "00:50:56:ae:00:19"
46 guestOS = "winntenterprise"
47 annotation = "Used to test network connectivity on all VLANs"
48 uuid.bios = "42 2e 91 30 4f 95 25 69-4b 91 1c 00 21 56 db cb"
49 vc.uuid = "50 2e d2 fe d9 ce fe 1b-80 ad 8d e9 a2 35 05 76"
50 log.fileName = "vmware.log"
51 snapshot.action = "keep"
52 sched.cpu.min = "0"
53 sched.cpu.units = "mhz"
54 sched.cpu.shares = "normal"
55 sched.mem.minSize = "0"
```

In the example VM below it has three vNICs on two separate vDSs. When troubleshooting you may need to coordinate the values here with the net-dvs output on the host;

- NetworkName will show "" when on a vDS.
- The .VMX will show the dvPortID, dvPortGroupID and port.connectid used by the VM - all three values can be matched against the net-dvs output and used to check the port configuration details - load balancing, VLAN, packet statistics, security etc

NOTE: Entries are not grouped together in the .VMX file so check the whole file to ensure you see all relevant entries.

```
sched.swap.derivedName = "/vmfs/volumes/4c90802a-eb1b5cbf-d8d9-001b78373634/MSCS node 1/MSCS node 1-c9ae8d46.vswp"
migrate.hostlog = "./MSCS node 1-c9ae8d46.hlog"
sched.cpu.affinity = "all"

ide1:0.clientDevice = "TRUE"
ide1:0.startConnected = "FALSE"
ethernet0.dvs.switchId = "52 8f 2e 50 1c 69 4c be-a3 6d 89 df 4c d8 73 69"
ethernet0.dvs.portId = "1"
ethernet0.dvs.portgroupId = "dvportgroup-243"
ethernet0.dvs.connectionId = "1932334040"
config.readOnly = "FALSE"
ethernet1.virtualDev = "vmxnet3"
ethernet1.pciSlotNumber = "160"
ethernet1.startConnected = "TRUE"
ethernet1.allowGuestConnectionControl = "TRUE"
ethernet1.features = "1"
ethernet1.wakeOnPcktRcv = "TRUE"
ethernet1.addressType = "vpx"
ethernet1.generatedAddress = "00:50:56:ae:00:0f"
ethernet1.networkName = ""
ethernet1.dvs.switchId = "3f aa 2e 50 e0 e5 7c 46-b4 2e de 78 92 76 e3 a1"
ethernet1.present = "TRUE"
ethernet2.virtualDev = "vmxnet3"
ethernet2.pciSlotNumber = "192"
ethernet2.startConnected = "TRUE"
ethernet2.allowGuestConnectionControl = "TRUE"
ethernet2.features = "1"
ethernet2.wakeOnPcktRcv = "TRUE"
ethernet2.addressType = "vpx"
ethernet2.generatedAddress = "00:50:56:ae:00:1a"
ethernet2.networkName = ""
ethernet2.dvs.switchId = "52 8f 2e 50 1c 69 4c be-a3 6d 89 df 4c d8 73 69"
ethernet2.present = "TRUE"
ethernet1.dvs.portId = "261"
ethernet1.dvs.portgroupId = "dvportgroup-225"
ethernet1.dvs.connectionId = "389459040"
ethernet2.dvs.portId = "133"
ethernet2.dvs.portgroupId = "dvportgroup-244"
ethernet2.dvs.connectionId = "1933380915"
scsi1.present = "FALSE"
scsi1:0.present = "FALSE"
floppy0.present = "FALSE"
```

6.3.2 Identify virtual switch entries in the ESX/i host configuration file

The host configuration file (same file for both ESX and ESXi);

- /etc/vmware/esx.conf

Like the .VMX file it contains entries for both switch types although there are only minimal entries for the vDS. Most vDS configuration is held in a separate database and can be viewed using net-dvs (see section 6.3.7).

6.3.3 Command line tools for network troubleshooting

The usual suspects;

- vicfg-nics
- vicfg-vmknic
- vicfg-vswitch (-b) for CDP

- vicfg-vswif
- vicfg-route
- cat /etc/resolv.conf, /etc/hosts
- net-dvs
- ping and vmkping

6.3.4 Identify logs used to troubleshoot network issues

Check section 6.1 for the typical logs used for vCenter and ESX/i hosts. Also;

6.3.5 Utilize net-dvs to troubleshoot vDS switch configurations

Charming! VMware put an objective in the exam which references an *unsupported* command and then provides next to no documentation on how to use it. So generous!

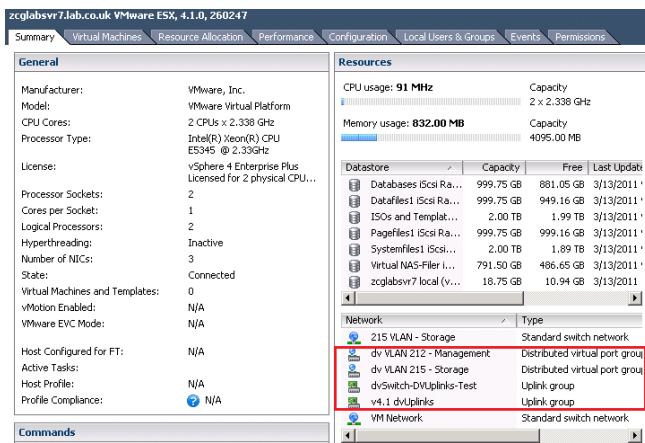
- Located in /usr/lib/vmware/bin (not in the PATH variable so just typing net-dvs won't work)
- Can be used to see the vDS settings saved locally on an ESX/i host;
 - dvSwitch ID
 - dvPort assignments to VMs
 - VLAN, CDP information etc
- Trainsignal's [Troubleshooting vSphere training course](#) covers the basics of net-dvs and how to match entries to a VM's .vmx file. Well worth the asking price.

Host is out of sync with vCenter vDS configuration

If the proxy switches (ie the local configuration on the ESX/i host) get out of sync with the vDS config as held in the vCenter database you'll get the following message for each host that's misconfigured;

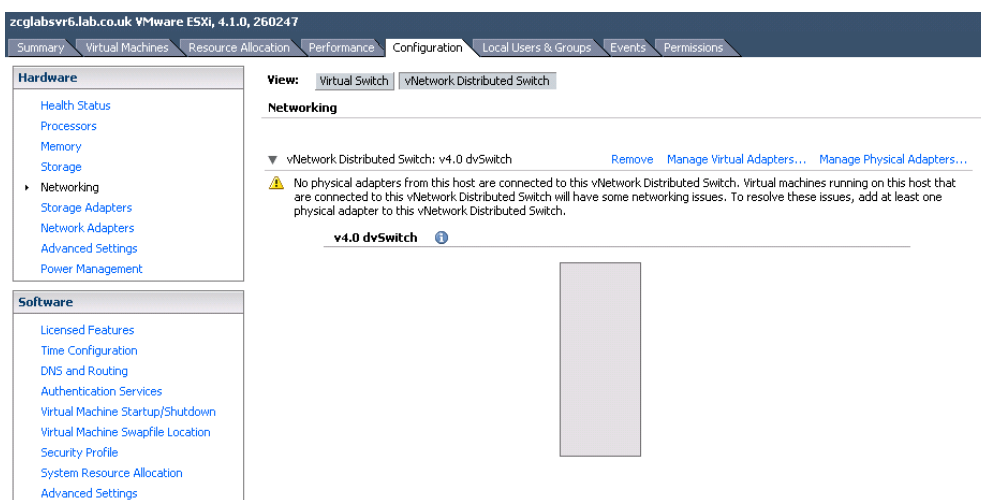
The screenshot shows the vCenter interface for a host named 'zcglabsvr7.lab.co.uk VMware ESX, 4.1.0, 260247'. The 'Configuration' tab is selected, and a yellow box highlights a 'Configuration Issues' message. The message states: 'HA agent on zcglabsvr7.lab.co.uk in cluster Test virtual cluster (ESX) in LondonLab1 has an error : Cannot complete the HA configuration. The vNetwork Distributed Switch corresponding to the proxy switches 0f 12 2e 50 8e 24 a3 02-0e 6e 02 7f 4b 9c 12 98 on the host does not exist in vCenter Server or does not contain this host. To resolve the issue, rejoin the host to v4.1 dvswitch. Go to Host > Configuration > Networking to manually remove the invalid proxy switches if the vCenter Server is not able to automatically remove them.'

When connecting the VI Client directly to the host in question you'll see it still has vDS network configuration even though vCenter shows nothing for this host. This shows that the host retains some of the vDS settings which may interfere with correct network functionality.



You can read some background and the solution to this problem on [Eric Sloof's blog](#) or in [VMwareKB1017558](#).

On another occasion the host failed while adding physical NICs to a dVS dvUplink PortGroup. This resulted in the host hanging and the vpxa agent failing. After rebooting the host I was presented with the following dVS configuration;



Steps to resolve;

1. Restart management agents
2. Reboot host
3. Connect VI client directly to host
4. Remove vDS config from host.

6.3.6 Configure a network packet analyser

There's a lot to cover with a packet analyser but remember you can do this in two places and each uses a different tools;

1. The guest OS (for virtual network traffic). This can use your choice of sniffer - Wireshark is a popular and free option ([check this post for a Wireshark tutorial](#)).
2. The service console or tech support mode (for management traffic, vMotion etc). You're more limited here as you can't install extra tools.

Packet sniffing within the virtual network

- Install your choice of packet sniffer in a guest OS (get [Wireshark here](#))
- Enable a promiscuous port for use with the VM running the packet sniffer.
NOTE: Either enable promiscuous mode on the vSwitch or create a dedicated portgroup and only enable promiscuous mode on that portgroup (slightly more secure). Check this [blogpost showing how to enable promiscuous mode](#).
- OPTIONAL: use a filter in Wireshark to track only certain types of traffic
 - tcp.port eq 902

Packet sniffing within the management network

The tool used varies between ESX and ESXi although usage is identical;

- tcpdump (ESX). Can capture both vmKernel and Service Console traffic.
- tcpdump-uw (ESXi). Only works on vmKernel interfaces.

[Tcpdump](#) isn't as full featured as Wireshark, it merely displays packets and lets you filter what to display using multiple criteria such as source IP, destination IP, ports etc. It won't do any analysis, highlighting or protocol recognition etc, so is less user friendly. You can output to a file using '-w' and open that in Wireshark for later analysis.

```
tcpdump -i vmk0 tcp -w /home/vi-admin/netcapture.pcap
```

NOTE: If you're using SSH to connect to a host (rather than a direct console connection via ILO etc) and then running tcpdump you'll see lots of traffic to port 22 as the screen updates are being sent to your screen! Exclude them using 'port not 22' on the end of the tcpdump syntax.

[VMwareKB1018217](#) shows a sample syntax used when diagnosing issues enabling HA.

[VMwareKB1031186](#) details how to capture packets on an ESXi host.

6.3.7 Troubleshoot vmKernel and Service Console network issues

vSphere 4.0 U2 introduced a new tool (console-setup) to make configuring the service console easier ([VMware KB1022078](#)). Prior to update 2 refer to [VMwareKB1000258](#). For general instructions on verifying service console connectivity refer to [VMwareKB1003796](#) (it's the same as RHEL5 if you're familiar with Linux).

```
5.1 vswif ID [0]
5.2 Name of service port group [Service Console]
5.3 vSwitch for service console [0]
5.4 IP Address [192.168.8.17]
5.5 Subnet mask [255.255.255.128]
5.6 Default gateway [192.168.8.1]
5.7 VLAN ID [0]
5.8 vmnic to use with service console [vmnic0]
5.9 Save Changes
5.10 Return to Menu
Enter your choice (1-10): _
```

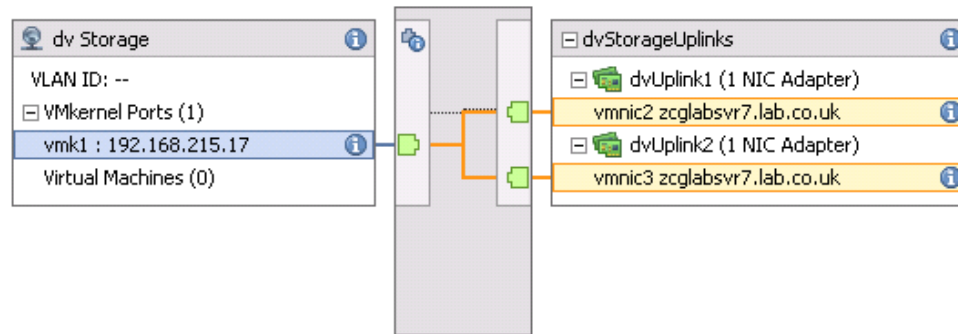
View: Virtual Switch vNetwork Distributed Switch

Networking

vNetwork Distributed Switch: dv Storage

[Manage Virtual Adapters...](#) [Manage Physical Adapters...](#)

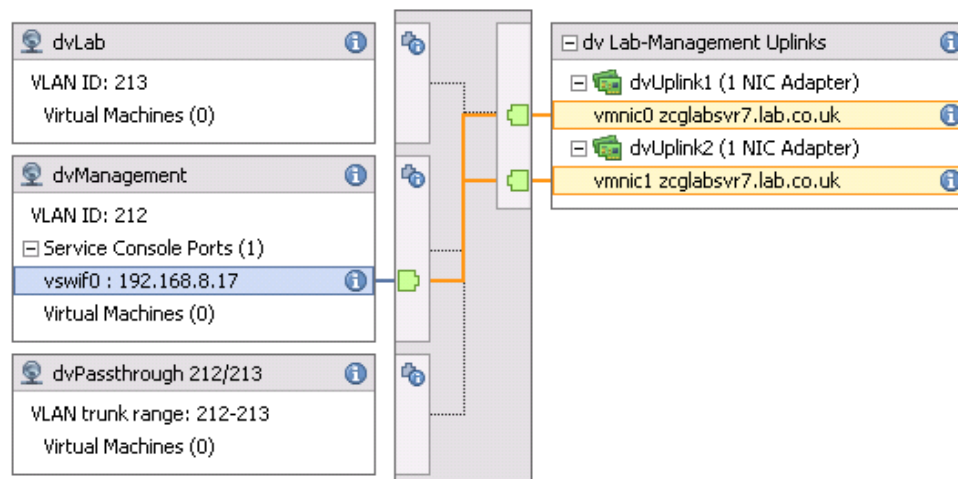
dv Storage ⓘ



vNetwork Distributed Switch: dv Lab_Management

[Manage Virtual Adapters...](#) [Manage Physical Adapters...](#)

dv Lab & Management ⓘ



Troubleshooting vMotion

- Check VM prerequisites
 - attached CD-ROM
 - persistent disks
 - snapshots
 - RDMs
 - Not on an internal-only vSwitch (unless you've overridden the default for vShield)
- Check host prerequisites
 - vmKernel ports
 - enough memory/CPU

Network troubleshooting steps

These steps are from [TA6862 - vDS Deep Dive - Managing and Troubleshooting](#);

- Ensure the vNIC is assigned to the correct portgroup
- Ensure the vNIC is 'connected'
- Verify the uplinks being used by the relevant portgroup
 - Is teaming correctly configured?
 - Is the physical switch configured correctly for all uplinks?
- Check VLAN tagging is correct at physical switch and portgroup
- Check MTU matches between vNIC and vSwitch (or portgroup)
 - Check vmknic MTU using `esxcfg-vmknic -m`
 - Run `ping -s <size> <destination IP>`
Here's a good [blogpost about MTU](#) - check the comments
- Check for dropped packets (either in esxstop or vCenter performance charts, and additionally on the Ports tab for a vDS)
- Use a packet sniffer to check traffic sent from the source is received at the destination. If not there may be issues in the physical network infrastructure.

Performance checks

- Check jumbo frames are enabled (especially if CPU usage is high)
- Verify that VMtools is installed in all VMs
- Locate VMs with high traffic requirements on the same host so traffic goes across the local bus rather than the network
- Enable TSO in the guest OS
- Where possible use VMXNET3 drivers which optimise performance

NOTE: [VMwareKB1003969](#) covers some good general troubleshooting steps.

NOTE: When configuring NIC teaming on an ESXi server and using 'route based on IP hash' with an Etherchannel link you can experience intermittent network access - see [VMwareKB1022751](#) for an explanation and solution.

6.3.8 Using resxtop/esxstop to diagnose network issues

You can use esxstop to check the NIC teaming is working as expected. Check the traffic per vmnic.

6.3.9 Using CDP and network hints to troubleshoot

- Enabled by default (in listen mode) on vDS
- Disabled by default on vSS
- CLI configuration (vSS only)
 - `vswitch -b <vSwitch>` Show CDP status for a given vSwitch
 - `vswitch -B both <vSwitch>` Enable CDP for a given vSwitch
- GUI configuration (vDS only)
 - Set on the vDS properties using the VI client

6.4 Troubleshoot Storage Performance and Connectivity

Knowledge

- Recall vicfg-* commands related to listing storage configuration
- Recall vSphere 4 storage maximums
- Identify logs used to troubleshoot storage issues
- Describe the VMFS file system

Skills and Abilities

- Use vicfg-* and esxcli to troubleshoot multipathing and PSA-related issues
- Use vicfg-module to troubleshoot VMkernel storage module configurations
- Use vicfg-* and esxcli to troubleshoot iSCSI related issues
- Troubleshoot NFS mounting and permission issues
- Use esxtop/resxtop and vscsiStats to identify storage performance issues
- Configure and troubleshoot VMFS datastores using vmkfstools
- Troubleshoot snapshot and resignaturing issues

Tools

- Product Documentation
 - ESX Configuration Guide
 - ESXi Configuration Guide
 - [vSphere Command-Line Interface Installation and Scripting Guide](#)
- vSphere Client
- vSphere CLI
 - vicfg-* , esxcli, resxtop/esxtop,vscsiStats, vmkfstools

There's obviously a large overlap between diagnosing performance issues and tuning storage performance, so check section 3.1 in tandem with this objective.

6.4.1 Recall vicfg-* commands related to listing storage configuration

- vicfg-scsidevs
- vmkiscsi-tool
- vicfg-mpath
- vicfg-iscsi
- esxcli corestorage | nmp | swiscsi
- vicfg-nas
- showmount -e
- esxtop/resxtop
 - look for CONS/s - this indicates SCSI reservation conflicts and might indicate too many VMs in a LUN. This field isn't displayed by default (press 'f' then 'f' again to add it)
- vscsiStats
- vmkfstools
- vicfg-module

6.4.2 Storage Maximums

Refer to section 1.2, Storage Capacity or read [VMware's vSphere Maximum's white paper](#).

6.4.3 Identify logs used to troubleshoot storage issues

- /var/log/vmKernel (ESX only)
- /var/log/messages
- /var/log/dmesg
- /var/log/vmkscsid.log
- /var/log/vmware/vpxa/vpxa.log
- /var/log/hostd/hostd.log
- vCenter logs
- Use tail -f to watch /var/log/vmkernel to monitor a svMotion

6.4.4 Describe the VMFS file system

I suspect this objective is carried over from the VI 3.5 Enterprise exam, and hence a good place to check is these [Enterprise 3.5 study notes](#). There's also a useful recent [blogpost by Deinos Cloud](#).

Be aware of VMFS alignment - [Vaughn Stewart's latest blogpost](#)

6.4.5 Use vicfg-* and esxcli to troubleshoot multipathing and PSA-related issues

Refer to section 1.3.

6.4.6 Use vicfg-module to troubleshoot VMkernel storage module configurations

Refer to section 9.1.3, advanced ESX builds.

6.4.7 Use vicfg-* and esxcli to troubleshoot iSCSI related issues

There are various ways to configure iSCSI at the command line;

- vicfg-iscsi - the most powerful CLI tool as it can configure pretty much anything – targets, host adapters, authentication, etc
- esxcli swiscsi – typically used to configure multipathing for iSCSI. See section ?? for details.
- esxcfg-swiscsi – can enable, disable, and query status of iSCSI but not much more.
- esxcfg-hwiscsi – can enable, disable, and query status of iSCSI but not much more.

Things to check;

- Are any LUNs masked? (see section 1.1.9)
- Which HBAs are available?
 - vicfg-iscsi -H -l
- Are the expected number of paths available?
 - vicfg-mpath -l
- Are the targets correctly configured?
 - vicfg-iscsi -T -l vmhba33
- Are the LUNs correctly configured?
 - vicfg-iscsi -L -l vmhba33
- Is authentication correctly configured?
 - vicfg-iscsi -A -l vmhba33

6.4.8 Troubleshoot NFS mounting and permission issues

There's a good video on this in the TrainSignal Troubleshooting Course, but in reality there's not much to check inside VMware – permissions are handled on the storage array or NAS server.

- Use vmkping to ensure the vmKernel has connectivity to the NAS server
- 'esxcfg-nas -l' to show configured NFS datastores
- For ESX hosts you may be able to use the SC to diagnose permissions issues;
 - Ensure the nfsClient service is enabled on the firewall (*esxcfg-firewall -e nfsClient*)
 - Ensure the service console can see the storage target using ping (the target may be on an isolated network as per best practices)
 - Use 'showmount -e <IP of storage array>' (*showmount -e 192.168.215.33*)

```
[root@zcglabsvr7 ~]# vmkping 192.168.215.33
PING 192.168.215.33 (192.168.215.33): 56 data bytes
64 bytes from 192.168.215.33: icmp_seq=0 ttl=64 time=1.103 ms
64 bytes from 192.168.215.33: icmp_seq=1 ttl=64 time=0.735 ms

--- 192.168.215.33 ping statistics ---
2 packets transmitted, 2 packets received, 0% packet loss
round-trip min/avg/max = 0.735/0.919/1.103 ms
[root@zcglabsvr7 ~]# esxcfg-firewall -q
Chain INPUT (policy ACCEPT 11M packets, 6719M bytes)
  pkts bytes target      prot opt in      out     source      destination

Chain FORWARD (policy DROP 0 packets, 0 bytes)
  pkts bytes target      prot opt in      out     source      destination

Chain OUTPUT (policy ACCEPT 11M packets, 6867M bytes)
  pkts bytes target      prot opt in      out     source      destination

Neither incoming nor outgoing blocked by default
Enabled services: nfsClient CIMSLP ntpClient aam VCB CIMHttpsServer vpxHeartbeats
  CIMHttpServer swISCSIClient faultTolerance sshServer

Opened ports:
  hostdSnmp          : port 0 udp.out
  hostdSnmp          : port 162 udp.out

Added Iprules:

[root@zcglabsvr7 ~]# showmount -e 192.168.215.33
Export list for 192.168.215.33:
/mnt/nfs1-openfiler/nfs1-openfiler/share1 192.168.215.17/255.255.255.255,192.168.
215.16/255.255.255.255,192.168.215.15/255.255.255.255,192.168.215.12/255.255.255.
255,192.168.215.14/255.255.255.255,192.168.215.13/255.255.255.255
[root@zcglabsvr7 ~]#
```

6.4.9 Use esxtop/resxtop and vcsisiStats to identify storage performance issues

Refer to section 3.5.

6.4.10 Configure and troubleshoot VMFS datastores using vmkfstools

Refer to section 1.2.1.

6.4.11 Troubleshooting snapshots and resignaturing

Use vmware-cmd;

vmware-cmd /vmfs/volumes/LocalRAID5/testVM/testVM.vmx hassnapshot

- Ruben Garcia's [excellent post on troubleshooting snapshots](#)

- Monitoring snapshot deletion ([VMwareKB1007566](#))
- Consolidating snapshots - a long but very good article - ([VMwareKB1007849](#))
- Best practices for snapshots in a VMware environment ([VMwareKB1025279](#))

For resignaturing refer to section 1.1.8.

6.4.12 Further Reading

- [TA1394 – Advanced Storage Log Analysis](#) (VMworld 2009 session)
- [Eric Sloof's Advanced Troubleshooting presentation](#) at the Dutch VMUG
- [VMware whitepaper on Troubleshooting Performance issues](#)
- [Trainsignal's Troubleshooting for vSphere course](#)
- [A combined multivendor post about iSCSI](#)
- Troubleshooting SCSI reservations - [VMwareKB1002293](#), [VMwareKB1005009](#)
- [Troubleshooting SCSI reservations](#)
- Chad Sakacc's [blogpost about storage performance](#), and [another about VMFS/NFS limits](#)
- [VMware article on troubleshooting storage performance](#)
- A good [presentation on troubleshooting storage by a VMware engineer](#)

6.5 Troubleshoot vCenter Server and ESX/ESXi Host Management

Knowledge

- Identify CLI commands and tools used to troubleshoot management issues

Skills and Abilities

- Troubleshoot vCenter Server service and database connection issues
- Troubleshoot the ESX Service Console firewall
- Troubleshoot ESX/ESXi server management and connectivity issues
- Determine the root cause of vSphere management or connectivity issue

Tools

- Product Documentation
 - [ESX Configuration Guide](#)
 - [ESXi Configuration Guide](#)
 - [vSphere Command-Line Interface Installation and Scripting Guide](#)
- vSphere Client
- vSphere CLI
 - vicfg-*
- [Virtual Center – Troubleshooting Unleashed](#) (VM2409, VMworld '09)
- [VMware whitepaper on Troubleshooting Performance issues](#)
- [Trainsignal's Troubleshooting for vSphere course](#)

6.5.1 Identify CLI tools used to troubleshoot management issues

- vicfg-vswitch
- vicfg-vmknic
- vicfg-vswif
- vpxd.exe -s

There are a few more covered later in this objective for restarting management agents on ESX/i hosts. This [VMware article on resolution paths](#) is a great place to start learning about troubleshooting.

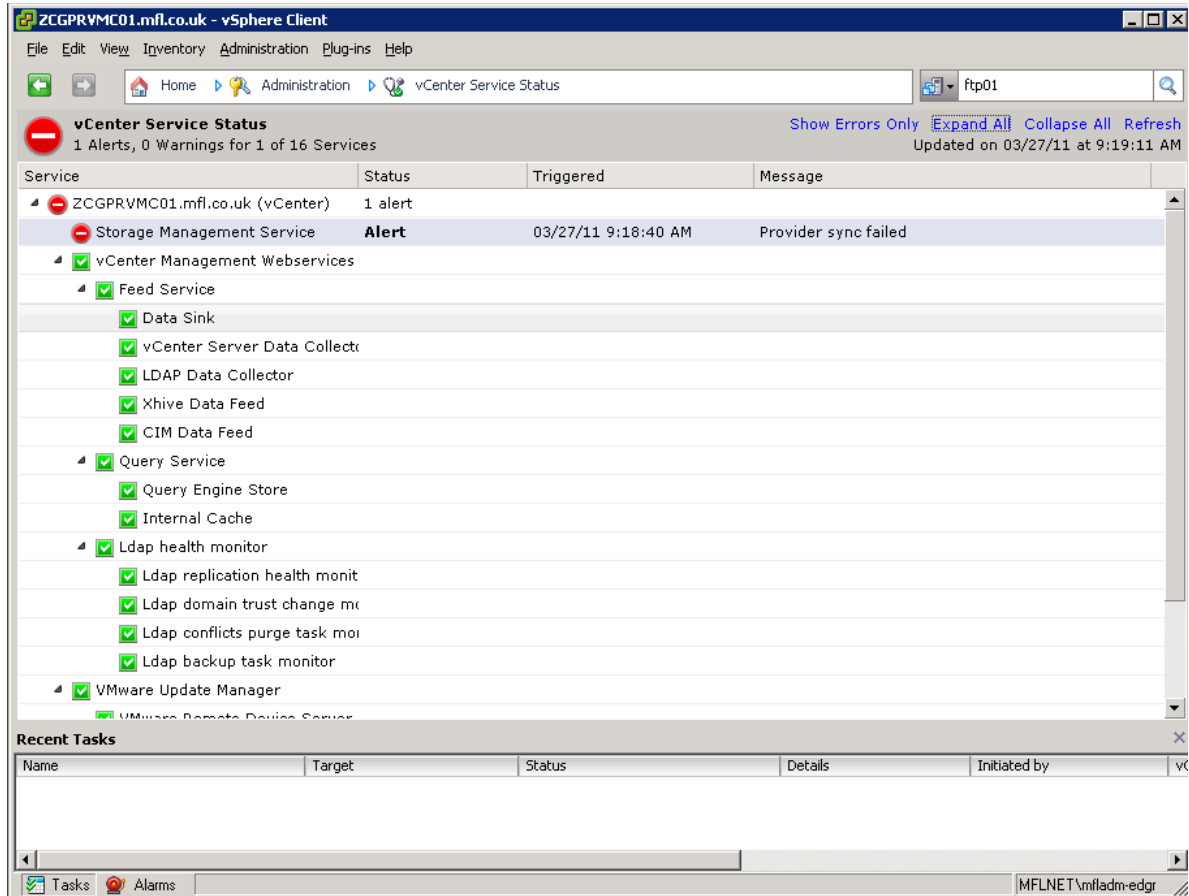
6.5.2 Troubleshoot vCenter Server service and database connection issues

- Check the VMware vCenter service is started and the account it's configured to run as. Check that account isn't locked out.
- Start vCentre using vpxd.exe;
 - 'vpxd.exe -s' to start it as an application rather than a service. This will show error messages in plain text rather than the cryptic service codes.
 - 'vpxd.exe -p' refreshes the password hash used to connect to the database. Used after replacing the default SSL certificates ([VMwareKB1003070](#))
- How to set SQL as a service dependency – [blog post](#)
- With a lab setup and SQL Express the database often grows to the 4GB limit, at which point the vCenter service will fail. Follow [VMwareKB1025914](#) for details of how to clear down data in the vCenter database.

- Check the ODBC connectivity using the 'Test' button. Check the SQL security logs to see failed authentication attempts.

[VMwareKB1003979](#) gives a good overview of the previous processes.

Check the status page using the VI client (Home -> Administration -> vCenter service status);



The vCenter Management Webservices

Under the hood vCenter uses a Tomcat webserver to provide functionality for Performance Charts, Storage View plugin, Service status and hardware status. If these components aren't available check the Windows service is started. Useful logs can be found here;

- C:\Program Files\VMware\Infrastructure\tomcat\logs
- %ALLUSERSPROFILE%\Application Data\VMware\VMware VirtualCenter\Log\sms.log

More for curiosity than likelihood of any exam related questions, check out this blogpost by Jason Boche about [modifying vCenter configuration via the vpxd.conf file](#).

6.5.3 Troubleshoot the ESX firewall

There's not that much to the Service Console firewall. One good step is to allow all outbound and incoming traffic and see if that resolves the issue - if so you can reenale rules and narrow down the specific ports required.

- Check the enabled firewall services in the GUI

- Telnet to a port to check connectivity. vCenter needs these ports. For a full list check out the [great firewall diagram at vReference.com](#);
 - 80
 - 443
 - 902
- Check esxcfg-firewall -q to see if individual ports are opened
- Use esxcfg-firewall -l to reload the firewall rules and ensure they're up to date.

NOTE: When enabling or disabling services using esxcfg-firewall they're case sensitive (like most things at the command line).

6.5.4 Troubleshoot the ESX/ESXi server management and connectivity issues

ESX

service mgmt.-vmware restart

service vmware-vpxa restart

Check logfiles - see section 6.1

ESXi

services.sh restart

DCUI - Troubleshooting tools - restart management agents

DCUI - Restart management network

Check logfiles - see section 6.1

7 Secure a vSphere Environment

7.1 Secure ESX/ESXi hosts

Knowledge

- Identify configuration files related to network security
- Identify virtual switch security characteristics

Skills and Abilities

- Add/Edit Remove users/groups on an ESX Host
- Customize SSH settings for increased security
- Enable/Disable certificate checking
- Generate ESX Host certificates
- Enable ESXi lockdown mode
- Replace default certificate with CA-signed certificate
- Configure SSL timeouts
- Secure ESX Web Proxy
- Enable strong passwords and configure password policies
- Identify methods for hardening virtual machines
- Analyze logs for security-related messages

Tools & learning resources

- [ESX Configuration Guide](#)
- [ESXi Configuration Guide](#)
- [vSphere Command-Line Interface Installation and Scripting Guide](#)
- [vSphere Command-Line Interface online reference](#)
- vSphere CLI
 - vicfg-user
 - vifs
- [William Lam's vsphere Health Check and Security Report](#)
- [VMworld 2010 session SE8206 - Security hardening guidelines for vSphere](#)
- [vSphere 4.0 Security Hardening Guide](#)
- [vSphere 4.1 Hardening Guide \(still in draft at time of writing, Feb 2011\)](#)
- [Virtualisation security podcast with Edward Haletky](#)
- [VMware security community](#)
- [VMware white paper – Replacing vCenter Server 4.0 certificates](#)

Security is a large topic and one you could spend a lifetime mastering. The blueprint isn't too helpful in clarifying what level of detail you're expected to know for this as the ESX/ESXi configuration guides cover issues not in the 'skills and abilities' section. More in depth still is the vSphere Hardening Guide. I guess the main thing is to focus on practical issues as the VCAP-DCA is a practical exam – knowing that the VMkernel uses memory hardening is no use in an exam if it can't be configured or tweaked! Some of this section seems to have been added for the sake of it – how often will an admin need to modify the SSL timeouts? I could only find one KB article about it!

7.1.1 Virtual network security

vSwitch security (layer2) settings (can be overridden at portgroup level);

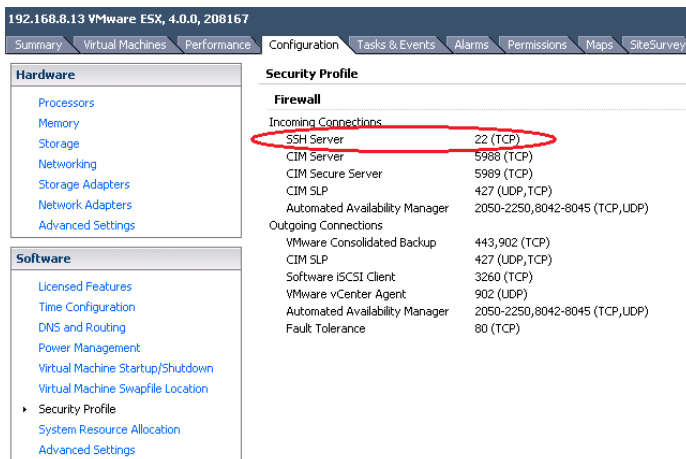
- Promiscuous mode – needed for packet sniffing (and virtual ESX hosts). Disabled by default.
- MAC address changes –affects inbound traffic. May need to be enabled if you’re using MS load balancing in Unicast mode, or the iSCSI software initiator with certain storage arrays. Disabled by default.
- Forged transmits – affects outbound traffic. Disabled by default.

Other network security measures (IPSec, VLANs, PVLANS etc) are dealt with in section 2, Networking.

7.1.2 Host security

Customise SSH settings (ESX only)

- Edit `/etc/ssh/sshd.conf` and set `'PermitRootLogin'` to YES (default is NO). See [VMwareKB](#) for a list of other settings you can adjust (including the available ciphers).
- By default only SSH server is enabled. Configuration -> Security Profile to enable SSHClient, or use `'esxcfg-firewall -e SSHClient'`.



- You can also use PKI to authenticate using SSH without being prompted for a password. This is a standard Linux procedure – for step by step instructions see [VMwareKB1002866](#).

Adding/deleting/modifying user/group security

- Connect VI client directly to the ESX/ESXi host (assuming lockdown mode isn’t enabled)
- Use `vicfg-user` (this is only available via the RCLI (hence vMA) but not directly on the ESX/ESXi host).

```
[vi-admin@zcglabvma01 ~]$ vifptarget -s zcglabsvr7
[vi-admin@zcglabvma01 ~][zcglabsvr7]$ vicfg-user -e user -o add -l
testuser2
Enter password for the user:
Enter password for the user again:
Created user testuser2 successfully.
[vi-admin@zcglabvma01 ~][zcglabsvr7]$ vicfg-user -e user -o delete -l
testuser2
Removed the user testuser2 successfully.
[vi-admin@zcglabvma01 ~][zcglabsvr7]$
```

Password policy

For both ESX and ESXi include a mix of characters from four character classes with passwords:

- lowercase letters
- uppercase letters
- numbers
- special characters such as an underscore or dash

The longer the password the less complex it needs to be (see the [ESXi Configuration Guide](#) for full details). For calculating complexity the first character is ignored if it's a capital, as is the last character if it's a number. This is to prevent Password01 (for example) being a legitimate password!!

esxcfg-auth is used to amend password policy on a host-wide basis (see Chapter 14 of the [ESX Configuration Guide](#));

- esxcfg-auth --probe Used to display the current settings
- esxcfg-auth --passmaxdays=DAYS Sets the maximum password age (default 99999)
- esxcfg-auth --passmindays=DAYS Sets the minimum password age (default 0)
- esxcfg-auth --passwarnage=DAYS Sets the password expiry warning period (default 7)

You can also amend the same settings on a per user basis. Log into the root console and run;

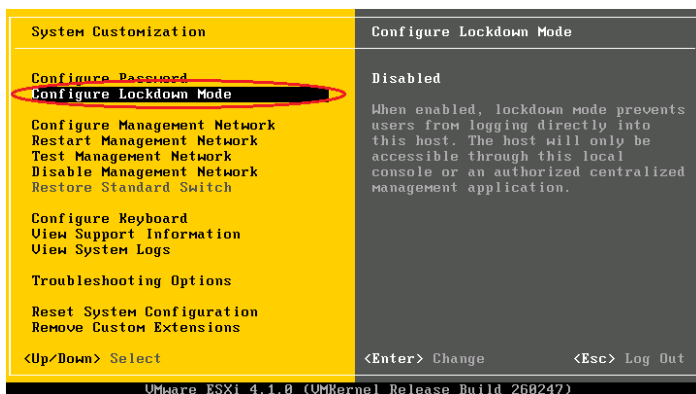
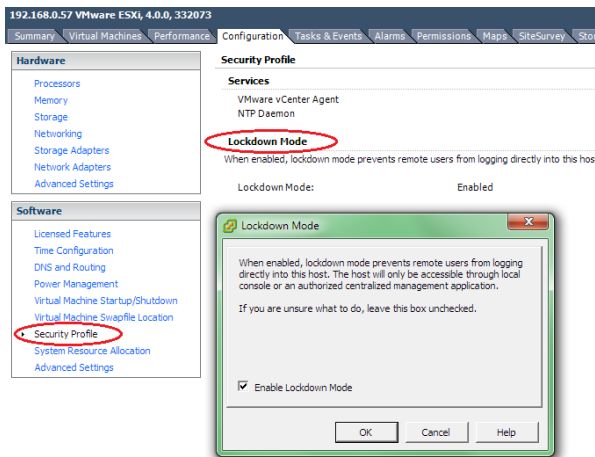
- chage -M 10 testuser Set the maximum p/w age to 10 days for testuser
- chage -m 1 testuser Set the minimum p/w age to 1 day for testuser
- chage -W 7 testuser Set the warning period to 7 days for testuser

ESX uses Pluggable Authentication Mechanism (PAM) via the service console. In a lab environment you can simplify the password policy – see [VMwareKB1012033](#) or this [article at vm-help](#).

7.1.3 Lockdown mode for ESXi

Further details can be found in this [VMware community post](#) (which covers v4.1) and David Davis has done a very useful [video showing how to activate it](#).

- ESXi lockdown mode can be enabled via the DCUI, VI client, or via RCLI (using vimcmd). It is NOT possible to use host profiles. 'vim-cmd' is only available on the ESX/ESXi hosts (not the vMA).
- Once enabled you can't use RCLI to disable it – you have to use the DCUI or VI client.
- You can disable the DCUI for total lockdown. Go to Configuration -> Security Profile (vSphere v4.1 only)
- It's disabled by default.
- Significantly different between v4.0 and v4.1. [VMwareKB1017628](#) explains the differences;
 - v4.0 only removes permissions for the root user but leaves other users (and Tech Support Mode) unaffected.
 - v4.1 removes all permissions and also disables Tech Support Mode (and more). *As a consequence ESXi lockdown mode disables the vi-admin account used by fastpass authentication from the vMA – [details on vGhetto](#)*
- When enabled all configuration must be done via vCenter (which authenticates against the host using the special 'vpxuser' account). See [VMwareKB1008077](#).

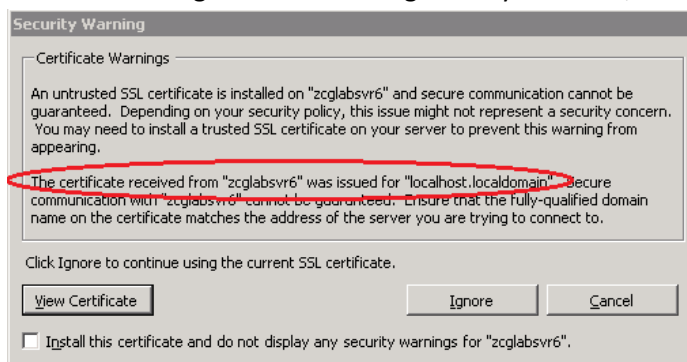


7.1.4 SSL settings and certificates

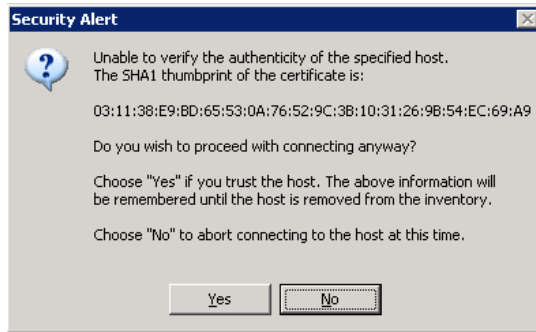
Certificates are used by vCenter to secure communications with the ESX hosts via SSL. Certificates are installed on both vCenter and each ESX/ESXi host.

NOTE: Certificate checking is a prerequisite for FT.

- Each host has a self-signed certificate generated during install. This is for localhost.localdomain however, so as soon as you rename the host it becomes invalid.
 - VI client warning when connecting directly to a host;



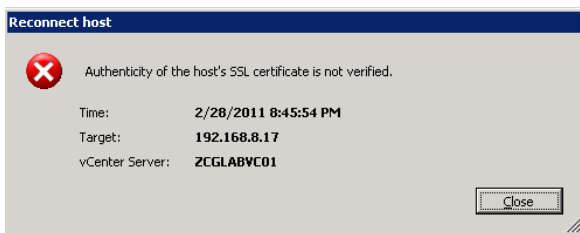
- Warning shown by vCenter when adding a host to inventory



- Certificate checking is enabled by default in vSphere (was disabled in VI3)
- Certificate checking can be disabled in vCenter via Administration -> vCenter Settings -> SSL
- On an ESX host (but not ESXi) you can regenerate the self-signed certificates by simply deleting the existing one. On restart the host will regenerate certificates;
 - Delete (or preferably backup) the two files (ru1.crt, ru1.key) in /etc/vmware/ssl

```
[root@zcglabsvr7 tmp]# service mgmt-vmware restart
Stopping VMware ESX Management services:
  VMware ESX Host Agent Watchdog          [ OK ]
  VMware ESX Host Agent                   [ OK ]
Starting VMware ESX Management services:
Generating VMware ESX SSL certificate...
Generating a 2048-bit RSA private key
.....+++
.....+++
writing new private key to '/etc/vmware/ssl/ca.key'
-----
Generating a 2048 bit RSA private key
.....+++
.....+++
writing new private key to '/etc/vmware/ssl/ru1.key'
-----
unable to write 'random state'
Signature ok
subject=/C=US/ST=California/L=Palo Alto/O=VMware, Inc/OU=VMware ESX Server Default Cert
2e
Getting CA Private Key
unable to write 'random state'
  VMware ESX Host Agent (background)      [ OK ]
  Availability report startup (background) [ OK ]
[root@zcglabsvr7 tmp]#
```

You'll need to remove and re-add the host to vCenter after changing the certificate otherwise you'll receive the following;



Using new certificates from a CA

1. Generate certificates using a certificate authority (CA). This can be a commercial CA or an internal CA. Either way you'll need the OpenSSL tools. See VMwareKB1023688 for full details.
2. Put the certificates on the host using *vifs* (for certificate and key respectively);


```
vifs --server <hostname> --username <username> --put ru1.crt /host/ssl_cert
vifs --server <hostname> --username <username> --put ru1.key /host/ssl_key
```

NOTE: You need to restart the management agents for this to take effect as you'll need to re-add the host to vCenter.

3. Add the CA certificate to the vCenter server's local certificate store to establish trust between the vCenter server and the ESX hosts.
4. Add the CA certificate to a VI client's local certificate store to establish trust between client sessions and vCenter.

Here's a good [walkthrough of applying a new root CA](#).

Generating new certificates has the potential to be disruptive;

- You may have to restart every host! (p5 of the [VMware certificate white paper](#)). In my tests you only had to restart the management agents however.
- After generating a new vCenter certificate you have to rejoin all the hosts to vCenter (to refresh the SSL handshakes). Unless you have vMotion and a cluster this could be very disruptive.
- vCenter may fail to start until you refresh the database connection (when you change the vCenter certificate) ([VMwareKB1003070](#))
- Deploying VMs with customisations may fail ([VMwareKB1019893](#))
- ESXi hosts are disconnected from vCenter due to lockdown mode ([VMwareKB1033572](#))
- Some vCentre plugins may fail ([VMwareKB1017577](#))

SSL timeouts

SSL timeouts may need adjusting when a host is very busy ([VMwareKB1022449](#)). There are two timeouts both of which can be modified by adding an entry to `/etc/vmware/hostd/config.xml`;

- SSL handshake timeout
- SSL read timeout

```
<ssl>
  <doVersionCheck> false </doVersionCheck>
  <useCompression>true</useCompression>

  <!-- to set handshake timeout add the line below -->
  <handshakeTimeoutMs>20000</handshakeTimeoutMs>

  <!-- to set read timeout add the line below -->
  <readTimeoutMs>20000</readTimeoutMs>
</ssl>
```

For ESXi either use tech support mode (SSH) or `vifs` to retrieve and upload the changed file;

```
vifs --get /host/hostAgentConfig.xml ~config.xml
```

NOTE: Using `vifs` this way uses an HTTP GET which means the path to the config.xml file is based on the webserver's root directory rather than the absolute path. It's also case sensitive so remember – [camelCase](#) FTW!

7.1.5 Securing the ESX web proxy

From the [ESXi Configuration Guide](#);

“To protect against misuse of ESXi services, most internal ESXi services are accessible only through port 443, the port used for HTTPS transmission. Port 443 acts as a reverse proxy for ESXi”

Where security is an issue you may want to disable some of the services offered by the built-in ESX/ESXi webserver (such as the default webpage). This can be done by editing `/etc/vmware/hostd/proxy.xml`. As with certificates, either use tech support mode or `vifs` to retrieve/put the file to ESXi hosts and remember to restart the management agents after any changes.

There are some good blogposts around this topic – [Duncan Epping](#), [Maish](#), and [William Lam](#) (the usual suspects!)

7.1.6 Other

Hardening virtual machines;

- Copy/paste from console (can be configured in the .VMX file or via VMtools)
- VM logging can be disabled or controlled via a log rotation size;
 - Edit Options -> Advanced -> General -> Enable logging (clear the tickbox)
 - add `log.rotateSize=100000` (100KB) to the .VMX file

Analyse logs – see troubleshooting section 6!

7.2 Configure and Maintain the ESX Firewall

Knowledge

- Identify vicfg-firewall commands
- Explain the three firewall security levels
- Identify ESX firewall architecture with/without vCenter Server

Skills and Abilities

- Enable/Disable pre-configured services
- Configure service behavior automation
- Open/Close ports in the firewall
- Create a custom service
- Set firewall security level

Tools & learning resources

- Product Documentation
 - [ESX Configuration Guide](#)
 - [ESXi Configuration Guide](#)
 - [vSphere Command-Line Interface Installation and Scripting Guide](#)
- vSphere Client
- vSphere CLI
 - vicfg-firewall
- [vSphere 4.0 Security Hardening Guide](#)
- [vSphere 4.1 Hardening Guide \(still in draft at time of writing, Feb 2011\)](#)
- [Virtualisation security podcast with Edward Haletky](#)

A blessedly quick objective this one! Quite why the ESXi Configuration Guide is listed in the blueprint is anyone's idea as ESXi doesn't contain a firewall! The blueprint also lists vicfg-firewall which is a typo – they mean esxcfg-firewall, as [vicfg-firewall doesn't exist!](#)

7.2.1 Firewall architecture

The ESX Configuration Guide talks very generally about where to put firewalls to protect traffic. In reality I can't see much difference in architecture whether you have a vCenter server or not.

Figure 12-1. Sample vSphere Network Configuration and Traffic Flow

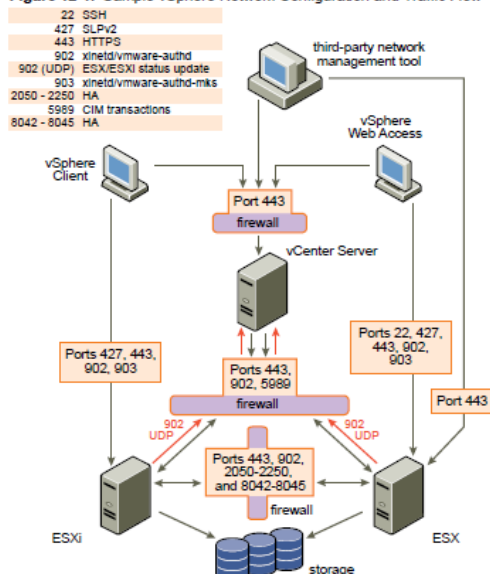
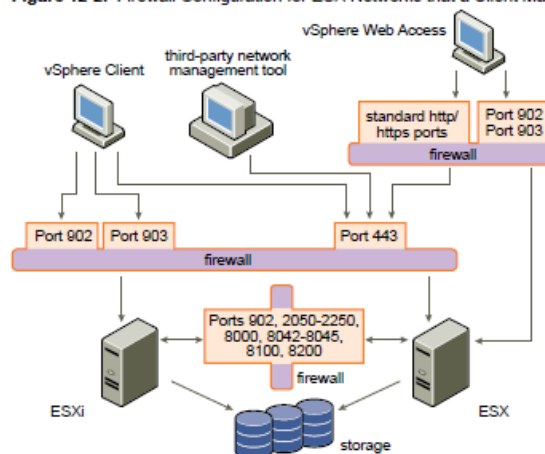


Figure 12-2. Firewall Configuration for ESX Networks that a Client Manages Directly



The firewall is ESX only (there's no ESXi firewall as no service console).

Three firewall security levels (high is default);

- High (outbound blocked, limited inbound allowed (902, 443,22,123 and a few other including ICMP).
- Medium (outbound allowed, inbound blocked apart from allowed services)
- Off

7.2.2 Services

A service is a pre-defined set of ports used to enable common functionality. This includes;

SSH server	22	TCP, inbound	
SSH client	22	TCP, outbound	
NTP	123		
FTP	20/21		
HTTP/HTTPS	80/443		
VI client	902 & 903		
NFS client			
NFS server	2049		Traffic from VMkernel to storage device
SNMP	161		
Syslog			
iSCSI	3260		On both service console and VMkernel
HA	8042-8045, 2050-2250		Between ESX hosts
FT	8100, 8200		Between ESX hosts on the FT logging port?
VUM	8084, 9084 (SOAP)		
vMotion	8000		Between ESX hosts on the VMkernel port

The default services for an ESX host (in an HA cluster) are shown below;

Hardware
Processors
Memory
Storage
Networking
Storage Adapters
Network Adapters
Advanced Settings

Software
Licensed Features
Time Configuration
DNS and Routing
Power Management
Virtual Machine Startup/Shutdown
Virtual Machine Swapfile Location
▶ Security Profile
System Resource Allocation
Advanced Settings

Security Profile

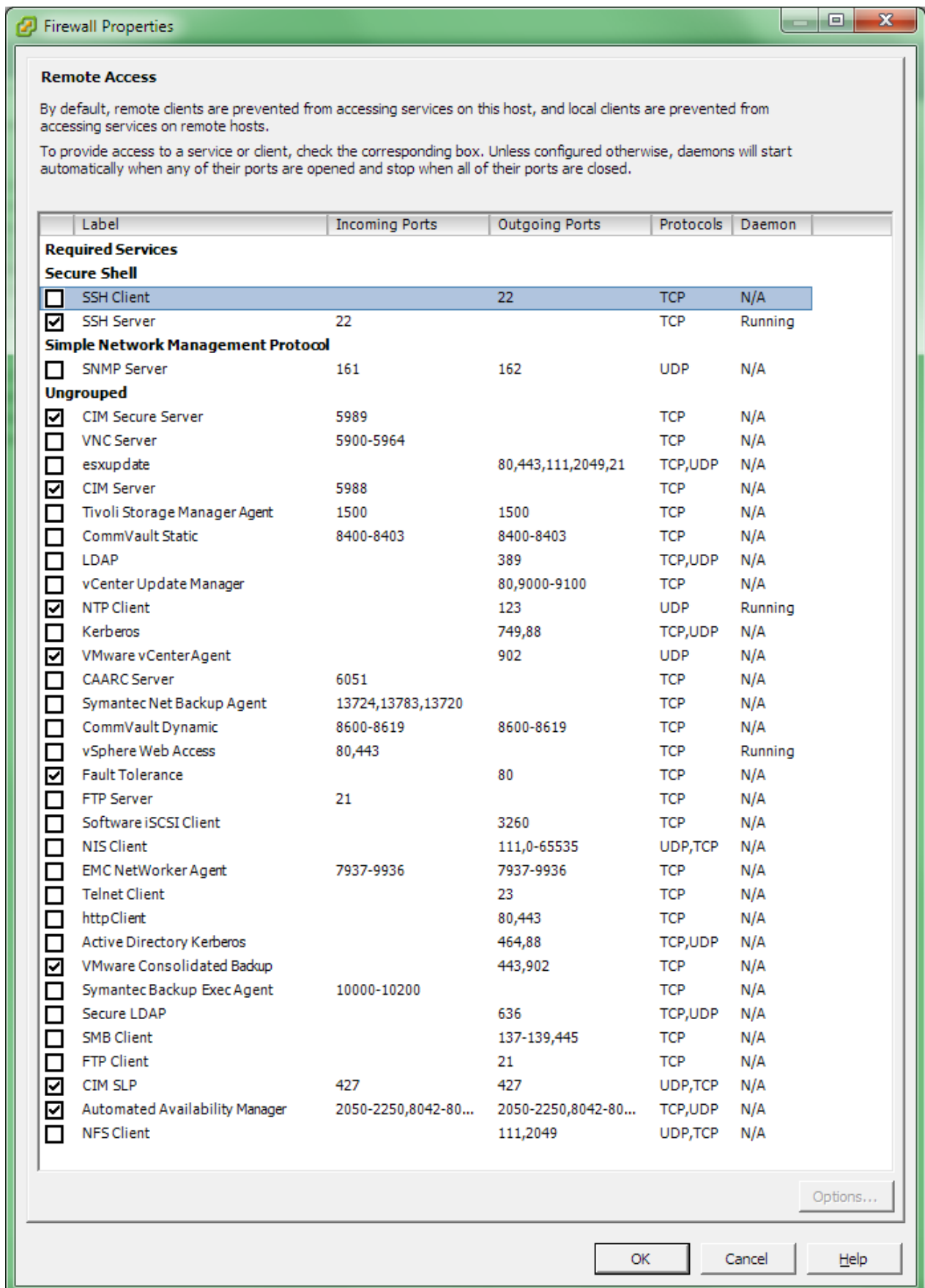
Firewall

Incoming Connections

SSH Server	22 (TCP)
CIM Secure Server	5989 (TCP)
CIM SLP	427 (UDP,TCP)
Automated Availability Manager	2050-2250,8042-8045 (TCP,UDP)
CIM Server	5988 (TCP)

Outgoing Connections

VMware vCenter Agent	902 (UDP)
CIM SLP	427 (UDP,TCP)
VMware Consolidated Backup	443,902 (TCP)
NTP Client	123 (UDP)
Fault Tolerance	80 (TCP)
Automated Availability Manager	2050-2250,8042-8045 (TCP,UDP)



Service behaviour automation

From the Configuration Guide, “ESX can automate whether services start based on the status of firewall ports”. This can be edited from the Options button for a particular service (under Configuration -> Security Profile). There are three options;

1. Start/stop when ports are open/closed. This is the default and recommended option.
2. Start/stop with the host.
3. Start/stop manually.

One example where you may need to do this is the DCUI service, part of ESXi’s lockdown feature in vSphere 4.1. See section 7.1 for details.

Create a custom service

If you have a management service, device or service which isn’t already in the predefined list it’s possible to create your own. You need to;

1. Create an .XML definition file and store it in the /etc/vmware/firewall folder
2. Enable this new definition (`esxcfg-firewall -e <your service>`)
3. Restart the management services (`service mgmt-vmware restart`)

For instructions see [Duncan Epping’s blogpost from 2007](#) which is still relevant and works just fine for vSphere. This also allows you to delegate the administration to a GUI user rather than requiring them to manually open ports by command line (which can be tiresome if it’s a large range for example).

7.2.3 Managing the firewall at the command line (esxcfg-firewall)

While the GUI is pretty simple to use there are occasions when it doesn’t offer enough flexibility. If you need to open ports not in the predefined list (but don’t want to go to the trouble of defining a service as it’s a temporary change) or you need to troubleshoot, the command line’s what you need;

- Set the firewall security level to high;
 - `esxcfg-firewall --blockIncoming --blockOutgoing`
- Set the firewall security level to medium;
 - `esxcfg-firewall --allowIncoming --blockOutgoing`
- Set allow all traffic through the firewall (see [VMwareKB1003634](#));
 - `esxcfg-firewall --allowIncoming --allowOutgoing`
- Open/Close ports. You’ll need to do this if you want to allow access to a service, device, or management agent which is not in the pre-defined services previously listed;
 - `esxcfg-firewall -o <ports>`
 - `esxcfg-firewall -c <ports>`
- To restart the firewall
 - `esxcfg-firewall -r`
- Show all predefined services
 - `esxcfg-firewall -s` (or `esxcfg-firewall --services`)
- Enable/disable pre-configured services (SSH etc)
 - `esxcfg-firewall -e <service>`
 - `esxcfg-firewall -d <service>`

The ESX firewall uses a modified [IP Tables](#), the default firewall shipped with Red Hat Enterprise Linux v5. If you want to know more about how it works under the hood here's [a good tutorial](#). All esxcfg-firwall does is call iptables under the hood – if you scan the firewall logfile you'll see multiple calls to iptables.

NOTE: Even though you can modify the firewall by using 'iptables' (the commands are still present in ESX) you shouldn't as it might break ESX functionality (and is certainly not supported).

7.2.4 Logging

The firewall logfile can be found at /var/log/vmware/esxcfg-firewall.log. This is largely an audit file containing changes to the firewall rules, NOT a log of the packets being either allowed or disallowed. Unlike many firewalls there isn't a live console where you can see this info although it is possible to enable logging of denied packets following advice in this [VMware communities post](#) (refers to VI3 but still works fine for vSphere). The '-v' parameter is undocumented although not sure why - seeing denied packets seems like useful functionality to me!

```
[root@esxi ~]# esxcfg-firewall -g
Chain INPUT (policy DROP 7526 packets, 1082K bytes)
pkts bytes target prot opt in out source destination
806K 597M ACCEPT all -- lo * * 0.0.0.0/0 0.0.0.0/0
33717 6119K valid-tcp-flags tcp -- * * 0.0.0.0/0 0.0.0.0/0
38158 6493K valid-source-address !udp -- * * 0.0.0.0/0 0.0.0.0/0
168K 49M valid-source-address-udp udp -- * * 0.0.0.0/0 0.0.0.0/0
49 2532 valid-source-address tcp -- * * 0.0.0.0/0 0.0.0.0/0 tcp flags:0x17/0x02
4441 373K icmp-in icmp -- * * 0.0.0.0/0 0.0.0.0/0
60521 21M ACCEPT all -- * * 0.0.0.0/0 0.0.0.0/0 state RELATED,ESTABLISHED
1 52 ACCEPT tcp -- * * 0.0.0.0/0 0.0.0.0/0 tcp dpt:902 state NEW
0 0 ACCEPT tcp -- * * 0.0.0.0/0 0.0.0.0/0 tcp dpt:80 state NEW
47 2428 ACCEPT tcp -- * * 0.0.0.0/0 0.0.0.0/0 tcp dpt:443 state NEW
143 47524 ACCEPT udp -- * * 0.0.0.0/0 0.0.0.0/0 udp spts:67:68 dpts:67:68
1 77 ACCEPT udp -- * * 0.0.0.0/0 0.0.0.0/0 udp dpt:427
0 0 ACCEPT tcp -- * * 0.0.0.0/0 0.0.0.0/0 tcp dpt:427 state NEW
0 0 ACCEPT tcp -- * * 0.0.0.0/0 0.0.0.0/0 tcp dpts:2050:2250 state NEW
3 1812 ACCEPT udp -- * * 0.0.0.0/0 0.0.0.0/0 udp dpts:2050:2250 state NEW
0 0 ACCEPT tcp -- * * 0.0.0.0/0 0.0.0.0/0 tcp dpts:8042:8045 state NEW
133K 33M ACCEPT udp -- * * 0.0.0.0/0 0.0.0.0/0 udp dpts:8042:8045 state NEW
0 0 ACCEPT tcp -- * * 0.0.0.0/0 0.0.0.0/0 tcp dpt:5989 state NEW
0 0 ACCEPT tcp -- * * 0.0.0.0/0 0.0.0.0/0 tcp dpt:5988 state NEW
1 52 ACCEPT tcp -- * * 0.0.0.0/0 0.0.0.0/0 tcp dpt:22 state NEW
0 0 ACCEPT tcp -- * * 0.0.0.0/0 0.0.0.0/0 tcp dpts:2500:3000

Chain FORWARD (policy DROP 0 packets, 0 bytes)
pkts bytes target prot opt in out source destination

Chain OUTPUT (policy DROP 0 packets, 0 bytes)
pkts bytes target prot opt in out source destination
806K 597M ACCEPT all -- * lo 0.0.0.0/0 0.0.0.0/0
29568 10M valid-tcp-flags tcp -- * * 0.0.0.0/0 0.0.0.0/0
4481 386K icmp-out icmp -- * * 0.0.0.0/0 0.0.0.0/0
49 3334 ACCEPT udp -- * * 0.0.0.0/0 0.0.0.0/0 udp spts:1024:65535 dpt:53
0 0 ACCEPT tcp -- * * 0.0.0.0/0 0.0.0.0/0 tcp spts:1024:65535 dpt:53
48369 15M ACCEPT all -- * * 0.0.0.0/0 0.0.0.0/0 state RELATED,ESTABLISHED
0 0 ACCEPT tcp -- * * 0.0.0.0/0 0.0.0.0/0 tcp dpt:902 state NEW
0 0 ACCEPT udp -- * * 0.0.0.0/0 0.0.0.0/0 udp spts:67:68 dpts:67:68
0 0 ACCEPT udp -- * * 0.0.0.0/0 0.0.0.0/0 udp spt:427
0 0 ACCEPT tcp -- * * 0.0.0.0/0 0.0.0.0/0 tcp spt:427 state NEW
2071 157K ACCEPT udp -- * * 0.0.0.0/0 0.0.0.0/0 udp dpt:123
0 0 ACCEPT tcp -- * * 0.0.0.0/0 0.0.0.0/0 tcp dpts:2050:2250 state NEW
573 1315K ACCEPT udp -- * * 0.0.0.0/0 0.0.0.0/0 udp dpts:2050:2250 state NEW
0 0 ACCEPT tcp -- * * 0.0.0.0/0 0.0.0.0/0 tcp dpts:8042:8045 state NEW
132K 33M ACCEPT udp -- * * 0.0.0.0/0 0.0.0.0/0 udp dpts:8042:8045 state NEW
0 0 ACCEPT tcp -- * * 0.0.0.0/0 0.0.0.0/0 tcp dpt:443 state NEW
0 0 ACCEPT tcp -- * * 0.0.0.0/0 0.0.0.0/0 tcp dpt:902 state NEW
13288 1249K ACCEPT udp -- * * 0.0.0.0/0 0.0.0.0/0 udp dpt:902 state NEW
0 0 ACCEPT tcp -- * * 0.0.0.0/0 0.0.0.0/0 tcp dpt:80
0 0 ACCEPT tcp -- * * 0.0.0.0/0 0.0.0.0/0 tcp dpts:2500:3000
6 388 REJECT all -- * * 0.0.0.0/0 0.0.0.0/0 reject-with icmp-port-
unreachable

Incoming and outgoing ports blocked by default.
Enabled services: CIMSLP ntpClient aam VCB CIMHttpsServer vpxHeartbeats CIMHttpServer faultTolerance sshServer

Opened ports:
VMX : port 2500:3000 tcp.in tcp.out
Added Iprules:

[root@esxi ~]#
```

7.3 Deploy and Administer vShield Zones

Knowledge

- Identify vShield Zones components
- Identify the four CLI command modes

Skills and Abilities

- Configure vShield Zones
- Backup and restore vShield Manager Data
- Backup CLI Configuration
- Create/Delete Layer 2/3/4 firewall rules using VM Wall
- Install/Uninstall a vShield manually and from template
- Configure vShield Manager plug-in capability
- Configure VM Flow charts
- Update vShield Zones
- Add/Edit/Delete User Accounts
- Assign rights to a user
- Add/Delete Application-Port Pair mapping
- Execute/Schedule Execution of virtual machine discovery
- Utilize vShield Zones CLI commands to configure and monitor vShield Zones
- Analyze traffic using VM Flow to determine root cause of network related issues

Tools

- Product Documentation
 - [Introduction to vShield Zones](#)
 - [vShield Zones QuickStart Guide](#)
 - [vShield Zones Administration Guide](#)
- vShield Manager
- vShield CLI
- vSphere Client
- [Rodos's musings on vShield Zones](#)
- Eric Sieberts three part series on vShield Zones (part [one](#), [two](#), and [three](#))
- Dave Convery's post on [vShield Zones' shortcomings](#)
- VMworld 2009 sessions – [session one](#), [session two](#)

vShield Zones is basically a firewall framework to protect your VMs without required external or hardware based firewalls. It requires Advanced or higher licencing. For study I'd suggest going through Eric Siebert's blogposts to start with (they cover real world issues) and then getting stuck into the official docs above – they cover everything on the blueprint. There's quite a bit to learn making this is one of the larger objectives on the VCAP-DCA blueprint.

vShield Zones is NOT the same as vShield App, Edge, and Endpoint so make sure you download the right version. The VCAP-DCA exam only covers v1.0 of vShield Zones (not the most recent v4.1) and doesn't cover the more feature rich vShield App Suite. See [VMware's product page](#) for more details.

7.3.1 Installing, configuring and updating vShield Zones

Deployed as an appliance with two components;

- Setup the vShield Manager appliance
 - Deploy the vShield Manager from OVF
 - Create a port group on the vSwitch which hosts your VM traffic, named vsmgmt and amend the vNIC on the vShield Manager VM to use this network.
 - Power up the VM, login with 'admin' and 'default', then run 'setup' to configure the server.
 - Allocate IP details
 - Upgrade VMtools (you can use the 'Automatic' option – being Linux based no reboot is required)
- Initial install of the vShield Agent
 - Deploy from OVF and then convert to a template. This simply gets the agent ready for deployment.

If you're wondering whether VMtools make a significant difference to this customised Linux appliance see (the pointless) [VMwareKB1011501](http://www.vmware.com/kb/1011501)! You can also find out [what's new in vShield Zones 1.0 Update 1](#).

Deploying/uninstalling an agent from template

Deploying a template based agent

- In vShield Manager, choose the host and click 'Install vShield' tab.
- Complete the fields giving a unique name and IP to each agent;

The screenshot shows the vShield Manager web interface. The main configuration area is titled 'Install vShield'. It includes sections for selecting a vShield to install, specifying vShield configuration (management port, IP address, default gateway, and vSwitch), and selecting a vSwitch to protect. Below this is a summary table of vShield instances.

Name	Target	Status
Reconfigure virtual ma...	labsvr4-vShield	Completed
Reconfigure virtual ma...	vShield Manag...	Completed
Reconfigure virtual ma...	labsvr4-vShield	Completed
Copy file	Systemfiles1 iS...	Completed
Copy file	Systemfiles1 iS...	Completed
Reconfigure virtual ma...	vShield Manag...	Completed
Reconfigure virtual ma...	vShield Manag...	Completed
Reconfigure virtual ma...	vShield Manag...	Completed
Add port group	192.168.8.14	Completed
Add port group	192.168.8.14	Completed
Add port group	192.168.8.14	Completed
Add virtual switch	192.168.8.14	Completed
Clone virtual machine	vShield	Completed

- Wait for the installation to complete. A new vSwitch (and associated portgroups) will be created (named as per the original vSwitch but with a '_VS' postfix).
- Upgrade VMtools in the vShield appliance once configuration has completed (you can use the 'Automatic option' – being Linux based no reboot is required)
- If the vShield appliance is in an HA/DRS cluster;
 - Go to HA settings and set 'Disabled' under Virtual Machine Options
 - Go to DRS settings and set 'Disabled' under Virtual Machine Options
 - Edit the .VPX to prevent vMotion warnings due to the 'internal only' vSwitch (see p13 of the vShield Quick Start guide)

Uninstalling a template based agent

- Simply select the vShield agent in the inventory and click on the 'Uninstall' tab. You'll be presented with a summary of actions to be carried out.
- Click Uninstall to confirm.

Deploying an agent manually

If you're using a VDS you'll need to deploy the vShield agent manually. I guess this might also be useful in case the automated deployment fails for any reason.

For a standard vSwitch;

- On your chosen host, create a new vSwitch (with a name which reflects it's a vShield vSwitch)
- On the same host create two new portgroups (three if you want a separate mgmt portgroup);
 - Unprotected on the original vSwitch (with the physical NIC)
 - Protected on the new vSwitch_VS you created in step 1
- Install the agent (Deploy from OVF, power on VM, install VMtools) to your chosen host
- Assign the three vShield agent vNICs to the three portgroups;
 - Management portgroup
 - Protected (vNIC2)
 - Unprotected (vNIC3)
- Configure the vShield agent via the CLI
 - Login as admin/default
 - Enable the interfaces using 'no shutdown' (as per Cisco devices). The interfaces are referred to as 'u0', and 'p0'.
- Add the vShield agent to the vShield Manager
 - Settings & Reports -> Manual Install, complete the fields (VM name, IP)
- Move the VMs to the new vSwitch.

For a dvSwitch the process is even more convoluted. I'm not going to cover it here – if it comes up in the exam refer to pages 36-41 in the vShield Administration Guide. I hope it doesn't!

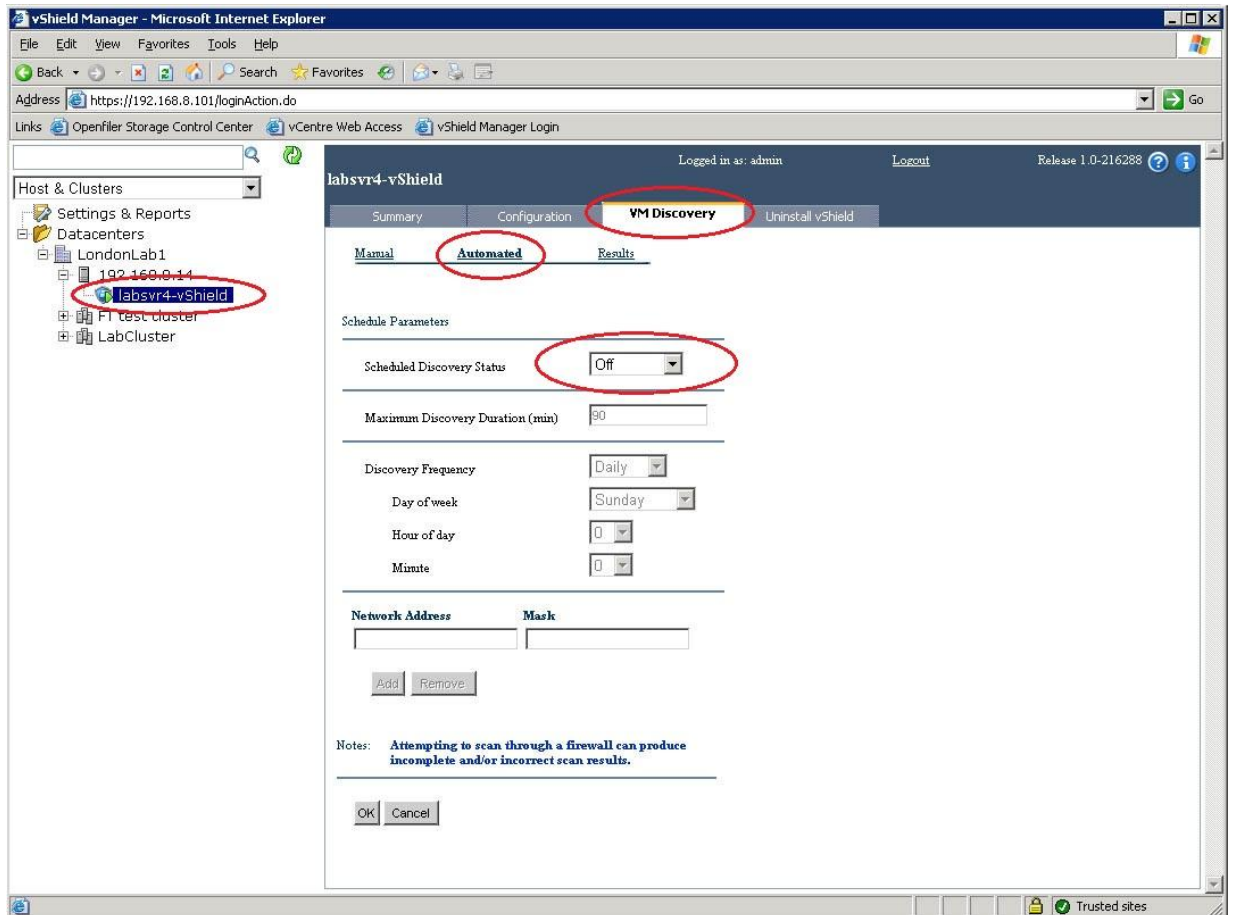
Uninstalling a manually installed agent

- In vSphere move the vNIC assignments for each VM to the unprotected vSwitch

- In vShield Manager go to Settings & Reports -> Configuration -> Manual Install
- Select the agent and click Remove
- In vSphere remove the extraneous portgroups

Execute/schedule discovery of virtual machines

- In vShield Manager, select the agent you want to configure then click on 'VM Discovery' tab.
- Select Automated and then set the 'Scheduled Discovery Status' to 'Continuous'. Click OK.



Updating vShield Zones

- Updated via offline bundles (not integrated with VUM)
- No notifications when updates are available, and I couldn't find any updates (though not entirely sure where to look). Not the easiest to use then!

- Very little useful detail in official documentation

System	System Software	Application Software	Decoder Software
Manager	192.168.8.101	-	11.0-1.0-216288
labsvr4-vShield	192.168.8.110	-	R-2137

Backup and restore

It's recommended to back up your system configuration regularly (which consists of configuration details, audit logs and event history). Simply go to Settings & Reports -> Configuration -> Backups and configure the appropriate fields. Decisions required;

- FTP or SFTP
- One off or scheduled

NOTE: I couldn't find anything in the official docs about how to restore from one of these backups...

Add users/groups

Users and groups exist both in the vShield Manager, the vShield Manager CLI (not the same accounts as the GUI) and also separately in every vShield agent. This means managing your users can be quite a chore (there's no AD integration either).

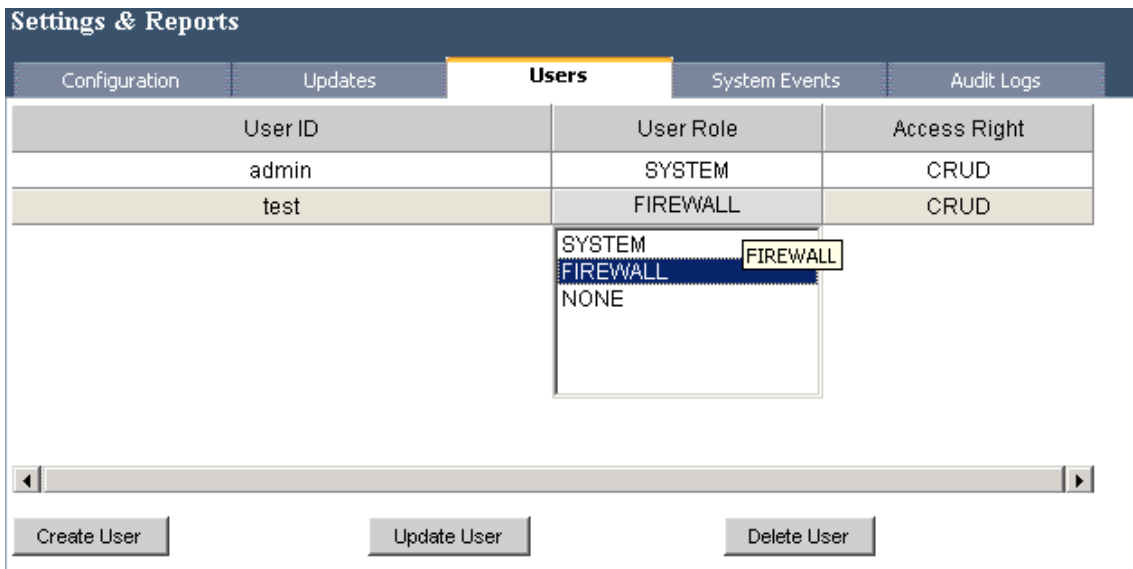
Possible rights (who came up with these?);

- R = read only
- CRUD = read and write

Resources (which can have the above rights granted to them);

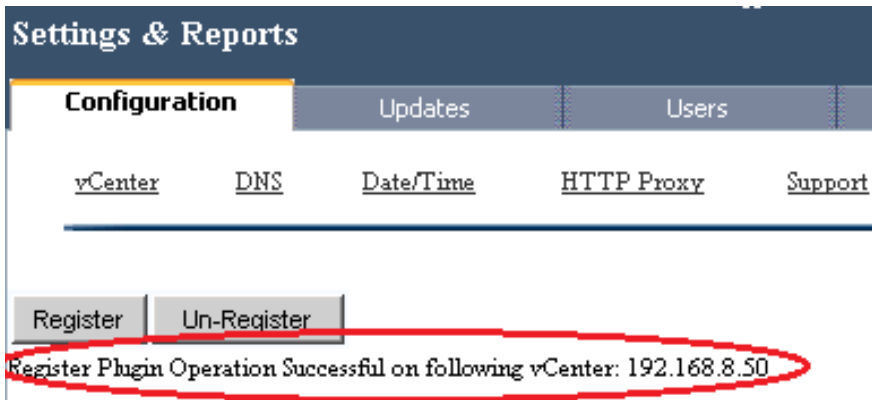
- System
- Firewall

- None



vShield Manager plugin

This isn't real integration with the VI client, it simply allows you to open the vShield Manager webpage within the VI client. Under Settings & Reports -> Configuration -> vSphere plugin click Register. After registering the plugin it'll show up in the Solutions section of the VI client. All rather pointless really!



7.3.2 Using vShield Zones

VM Wall

This is the actual firewall component.

- Rules are hierarchical and can be specified at a datacentre or cluster level

- The default rules (both layer4 and layer2/3) ALLOW all traffic. The default rules can't be added to or deleted, but they can be changed to DENY instead.

Source (A.B.C.D/nn)	Destination (A.B.C.D/nn)	Protocol	Action	Log	Notes
DataCenter Rules					
Default Rules					
ANY	ANY	ARP	ALLOW	<input type="checkbox"/>	
ANY	ANY	OTHER IPv4	ALLOW	<input type="checkbox"/>	
ANY	ANY	OTHER LAYER 3	ALLOW	<input type="checkbox"/>	

- Layer2/3 rules (TCP, IP, ARP) can ONLY be configured at a datacentre level.
- As is typical for firewalls, all traffic is matched against the rules from top to bottom of the table – the first rule which matches is enforced.
- The two common configurations;
 - Allow all traffic but deny specific traffic as defined in the VM Wall
 - Deny all traffic but allow specific traffic as defined in the VM Wall
- The screenshot below shows a layer 4 rule I created which DENIES access to RDP (on port 3389 by definition) from any VM outside my cluster;

Source (A.B.C.D/nn)	Source Port	Destination (A.B.C.D/nn)	Destination Application	Destination Port	Protocol	Action	Log	Notes
Data Center High Precedence Rules								
ANY	ANY	Outside FT test cluster (LondonLab1)	RDP	3389	TCP	DENY	<input checked="" type="checkbox"/>	
Rules below this level have lower precedence than the cluster level rules								
Default Rules								
ANY	DHCP-Client	ANY	DHCP-Server	67	UDP	ALLOW	<input type="checkbox"/>	
ANY	DHCP-Server	ANY	DHCP-Client	68	UDP	ALLOW	<input type="checkbox"/>	
ANY	ANY	ANY	-	ANY	TCP	ALLOW	<input type="checkbox"/>	
ANY	ANY	ANY	-	ANY	UDP	ALLOW	<input type="checkbox"/>	

Add application-port pair mappings

vShield Zones ships with definitions for many well known protocols (for example RDP on TCP port 3389). If you need to create new mappings (for example VNC isn't present) you can;

- Select either datacentre or cluster node
- Go to the VM Flow tab, then click 'Edit Mappings'
- Click Add at the bottom of the screen. Fill in the columns accordingly.

LondonLab1

Summary **VM Flow** VM Wall

Hide Port Mappings

Application	Port	Protocol	Resource	Description
NBDG-Unicast	138	UDP	ANY	
NBDG-Broadcast	138	UDP	ANY	
NBSS	139	TCP	ANY	
IMAP	143	TCP	ANY	
SNMP	161	UDP	ANY	
LDAP	389	TCP	ANY	
LDAP	389	UDP	ANY	
HTTPS	443	TCP	ANY	
MS-DG	445	TCP	ANY	
MS-DS	445	UDP	ANY	
ISAKMP	500	UDP	ANY	
LDAP over SSL	636	TCP	ANY	
MS-SQL-S	1433	TCP	ANY	
MS-SQL-M	1434	UDP	ANY	
ORACLE-TNS	1521	TCP	ANY	
ORACLE-XDB-FTP	2100	TCP	ANY	
Windows Global Catalog	3268	TCP	ANY	
Windows Global Catalog over	3269	TCP	ANY	
RDP	3389	TCP	ANY	
ORACLE-HTTP	7777	TCP	ANY	
New	Port	TCP	ANY	
ORACLE-FORM-SERVICES	9000	TCP	ANY	

Add Delete

VM Flow

VM Flow is your monitoring tool to see traffic flows and the impact of any rules you've defined. There are two available views, graphical (real time?) 'charts' and 'reports' which are a tabular table.

Reports are generally more useful as they let you see actual metrics. The screenshot below clearly shows the three blocked RDP attempts I made after creating a rule to deny RDP (port 3389);

LondonLab1

Summary **VM Flow** VM Wall

Start Date: End Date:
 02/23/2011 03/02/2011 Update Report Show Chart

Application	Sessions	Packets	Bytes	VMWall
BLOCKED	0	3	144	
TCP	0	3	144	
INCOMING	0	3	144	
UNCATEGORIZED	0	3	144	
RDP	0	3	144	
zcgtestdc01 (192.168.213.11)	0	3	144	
192.168.213.11	0	3	144	
UNCATEGORIZED	0	0	0	
OUTGOING	0	0	0	
INTRA	0	0	0	
INTRA_HOST	0	0	0	
ALLOWED	1202	156,375	30,331,513	
TCP	153	154,133	30,184,376	
UDP	1047	1,099	141,785	
DYNAMIC_TCP	2	14	592	
ICMP	0	76	4,560	
OTHER-IPV4	0	5	200	
ARP	0	1,048	0	

You can also create VM Wall rules directly from the VM Flow reports. Simply drill down far enough and click on the VM Wall radio button – it'll automatically switch to the VM Wall screen with a rule prepopulated.

You can also look at the logs directly although if there's significant traffic in your network this may well become hard to read pretty quickly;

- Select the vShield agent in question (one reason NOT to use this method – it won't always be easy to know which host)
- Click Configuration -> Show Status then click on Show Logs at the bottom of the screen. In the screenshot below you can again see my blocked RDP attempts (in a slightly harder to read format).

The screenshot shows the vShield Configuration interface. The 'System Status' tab is selected, showing system utilization (CPU 1%, Memory 38%, Base Software Storage 50%, Configuration Storage 2%) and a table of segments with their link statuses. Below this, the 'VM Wall Log' section is visible, containing several log entries for blocked RDP attempts. The log entries are as follows:

```

Mar 2 15:05:11 labrv4-vShield [64949.905690] VMWALL: (1013-DROPP) IN=>OUT=ip0 SRC=192.168.213.11 DST=192.168.213.150 LEN=48 TOS=0x00 PREC=0x00 TTL=128 ID=426 DF PROTO=TCP SPT=1171 DPT=3389 WINDOW=65535 RES=0x00 SYN URGP=0
Mar 2 15:05:14 labrv4-vShield [64952.897382] VMWALL: (1013-DROPP) IN=>OUT=ip0 SRC=192.168.213.11 DST=192.168.213.150 LEN=48 TOS=0x00 PREC=0x00 TTL=128 ID=435 DF PROTO=TCP SPT=1171 DPT=3389 WINDOW=65535 RES=0x00 SYN URGP=0
Mar 2 15:05:20 labrv4-vShield [64958.934465] VMWALL: (1013-DROPP) IN=>OUT=ip0 SRC=192.168.213.11 DST=192.168.213.150 LEN=48 TOS=0x00 PREC=0x00 TTL=128 ID=448 DF PROTO=TCP SPT=1171 DPT=3389 WINDOW=65535 RES=0x00 SYN URGP=0
  
```


7.3.3 Administering vShield Zones at the command line

Identify the four CLI modes

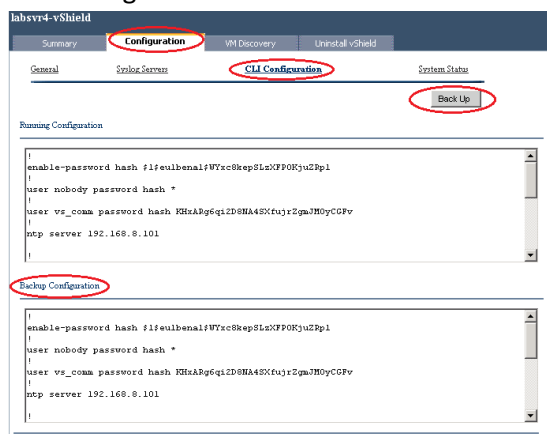
Administering either vShield Manager or a vShield agent at the CLI is much like (insert vendor or choice: Cisco, Vyatta, Netapp) administration. There are four command modes which give access to different commands and tab completion can be used to effectively navigate;

- Basic – this is a read only mode (with the exception of the ‘setup’ wizard)
- Privileged
 - Accessed via ‘enable’ (password is optional)
 - Typical operations: enable an interface (‘no shutdown’), reboot,
 - You must use ‘write’ to make any changes persistent
- Configuration
 - Accessed via ‘config terminal’. Familiar eh? ☺
 - Typical operations: clear vmwall rules, copy running-config startup-config, set the ‘enable’ password, add/edit/delete user accounts
- Interface configuration.
 - Accessed via ‘interface <mgmt. | p0 | u0>’
 - Typical operations: set an IP address

Backup CLI configuration

This is shockingly basic. Just like a Cisco device the configuration is stored in a ‘running config’ which you can simply cut and paste as a backup.

- Select the vShield agent in the vShield Manager Inventory
- Click Configuration -> CLI Configuration
- Click ‘Back up’. All this does is copy the contents of the top window to the bottom window. In the event you need to restore the configuration you would simply cut and paste the configuration from the bottom window into the CLI (at the configuration prompt) to restore the settings.



NOTE: There is no date/timestamp or even comments field for the backup window, so it'd be easy to forget when the backup was taken and what the state was at the time. I'd have thought cut and pasting this to a Notepad document would provide a better service than this built in backup!

8 Scripting and Automation

8.1 Execute VMware Cmdlets and Customize Scripts Using PowerCLI

Knowledge

- Identify vSpherePowerCLI requirements
- Identify Cmdlet concepts
- Identify environment variables usage

Skills and Abilities

- Install vSpherePowerCLI
- Install Update Manager PowerShell Library
- Use basic and advanced Cmdlets to manage VMs and ESX Hosts
- Use Web Service Access Cmdlets
- Use Datastore and Inventory Providers
- Use VMRun to execute commands in a guest OS
- Given a sample script, modify the script to perform a given action

Tools & learning resources

- vSpherePowerCLI Installation Guide (12 pages)
- vSpherePowerCLI Administration Guide (30 pages. Work through all examples)
- vSpherePowerCLI for Update Manager Admin Guide (12 Pages)
- The VMWare course - [VMware vSphere: Automation with vSpherePowerCLI \[V4\]](#)
 - Module 1: Course Introduction
 - Module 2: Introduction to vSpherePowerCLI
 - Define the main vSpherePowerCLI object
 - Define the main commandlets in vSpherePowerCLI
 - Connect to a vSphere infrastructure
 - Get help for commandlets
 - Module 3: Automating ESX Host Configuration
 - Automate configuration of virtual switches on ESX hosts
 - Automate configuration of datastores on ESX hosts
 - Module 4: Virtual Machine Provisioning, Configuration, and Protection
 - Automate creation of virtual machines
 - Change virtual machine settings programmatically
 - Run vSpherePowerCLI scripts in virtual machines
 - Automate virtual machine protection
 - Module 5: Automating Cluster Operations
 - Automate virtual machine storage migration
 - Automate VMware vMotion™
 - Create a VMware Distributed Resource Scheduler/VMware High Availability cluster
 - Automate cluster configuration
 - Automate resource pool creation and configuration
 - Module 6: Automating Reporting
 - Automate reporting about virtual machines
 - Automate reporting about ESX hosts
 - Automate reporting about clusters
- [Using VMRun to control Virtual Machines](#) (VMware white paper, 19 pages)

- [Train signal videos](#)
- [vSphere SDK Reference Guide](#)
- [Powershell Community Toolbar](#) – provides a great list of blogs and resources (tutorials etc)
- [VMware PowerCLI communities](#) – ask questions here and someone will answer
- Hal Rottenberg's book – [Managing VMware Infrastructure with Windows Powershell](#)
- Icamasoft's [PowerCLI reference card](#)
- [Managing vSphere with PowerCLI \(VM2241, VMworld '09\)](#)

8.1.1 Installing PowerCLI

Requires multiple products (not all VMware).

Client requirements: WinXP SP2, Win2k3 or greater, 32 or 64 bit

Server requirements: ESX or ESXi v3.0, vCentre 2.01 (or greater)

.NET framework v2.0 SP1 (or greater)

Powershell v1 or v2

PowerCLI

Note: if you're using the free edition of ESXi it's read only to PowerCLI.

You can also install PowerGUI for a graphical shell, or Visioncore's Virtualisation EcoShell (VESI). Both come with script editors offering syntax highlighting and debugging facilities. Powershell v2 comes with the Powershell integrated Scripting Environment (ISE) which offers similar functionality. All three are free. The VESI can import pre-supplied powerpacks such as AI Renouf's Community Powerpack.

Runs as a webservice, so needs access from the management workstation on ports 80 and 443 to the ESX hosts or vCentre. Scripts can be run against either hosts or the vCentre server.

PowerCLI is a group of .NET classes bundled into a snapin to Powershell.

Set-ExecutionPolicyRemoteSigned

By default Powershell restricts script execution. Use this command to allow local scripts but require signing for remote scripts.

Add-PSSnapInvmware.vimautomation.core
(or *Add-PSSnapinvmware**)

Get-PSSnapIn –registered
\$host.version

To confirm which snapins are loaded

Use at the Powershell prompt to check the installed PS version

8.1.2 Installing Update Manager Powershell library

Installing the Update Manager cmdlets is a separate installation, adds a second snapin (VMware.VUMAutomation). Same requirements for installation as PowerCLI.

Update Manager cmdlets are only supported on vSphere 4.0u1 or newer.

8.1.3 Cmdlet concepts

Gives full access to the VI SDK – anything you can do in VI client (and things you can't do in client!). For instance you can move a template without converting through to VM.

A cmdlet provides a simplified interface into the underlying .NET objects. As of vSphere 4.0 Update1 there are over 200 cmdlets. Cmdlets are referred to as Automation Objects as opposed to the more comprehensive objects in the SDK (known as Managed or View objects).

The vSphere SDK has two main types of objects;

- Managed objects provide both methods and properties
- Data objects provide only properties

Sometimes a cmdlet won't provide all the functionality you need or there won't be a cmdlet at all – in these cases you can use the Get-View cmdlet to access the API directly. Combined with 'cmdlet binding' (a feature of PowerShell v2) this allow you to write your own cmdlets. A collection of advanced functions using cmdlet binding can be saved in a .PSM1 file and imported using the *Import-Module* cmdlet.

Cmdlets can be used both interactively at the command line and via scripts.

8.1.4 Environment variables

Variables in Powershell are loosely typed and start with a '\$' symbol.

\$PROFILE points to the current profile although properties of this object contain paths to all four possible profiles;

- AllUsers all hosts (\$profile or \$profile.AllUsersAllHosts)
- alluserscurrent host (\$profile.AllUsersCurrentHost)
- this user all shells (\$profile.CurrentUserAllHosts)
- this user PowerCLI (\$profile.CurrentUserCurrentHost)

Variable scope - global variables, script variables, Powershell sessions

CD env:-provides access to machine environment variables (similar to SET in MS-DOS). To access use the syntax *\$env:<variable>* for example *\$env:path*

CD variable:- provides a list of Powershell values such as \$WarningPreference

\$DefaultVIServer– points to the currently connected ESX host or vCentre server. Can be an array pointing to multiple hosts (vSphere 4.0U1 and newer). A useful property is \$DefaultVIServer.IsConnected to check if connection is active.

8.1.5 Web Service Access cmdlets

These are the Get-View (get the complex version given a simple object) and Get-VIObjectbyVIView (get the simple view given a complex object)

Allow access to the full vSphere API. More complex but more powerful and often faster.

Can browse the API using the Managed Object Browser (MOB) - <http://vCentre/mob> or <http://esxhost/mob>

Typical operations that used to (or still do) require Get-View;

- converting a VM to a template (rather than cloning to template)
- powering down an ESX host (as of vSphere 4.0U1 there is now a Stop-VMHostcmdlet)
- Enabling vMotion on a host (can now be done via Set-VMHostNetworkAdapter – vMotionEnabled \$true)
- ...lots of queries where speed is an issue. Get-View tends to be quicker!

8.1.6 Datastore and Inventory providers

'Powershell drives' are navigational structures much like a directory tree in a filesystem. Can be written by third parties – common datastore providers are;

- vmstore: The DatastorePSDrive. Lets you navigate the VI Datastores.
- vi: The Inventory PSDrive. Lets you navigate the VI folder structure.
- env: Environment variables

- variable: Powershell variables
- HKCU: The current user's Windows registry key
- AD: Active Directory (W2k8 R2 onwards)

cmdlets: Get-PSDrive, New-PSDrive, Remove-PSDrive

The PowerCLI Admin Guide states that operations in Datastore provider are case sensitive but in my tests case was irrelevant.

When using the PSDrives, the objects can be treated just like files;

- Del <vm> Delete a VM
- Ren<vm> Rename a VM
- Dir<vm> | Start-VM Start all the VMs in the current directory

One exception to this is copying files to or from a datastore. To do this use the Copy-DatastoreItemcmdlet;

- Copy-DatastoreItem<src><dest>
Copy-DatastoreItemvmstore:\MyDC\NFSSMount\VM c:\VMBackup*

Very slow!

8.1.7 Powershellcmdlets

- Get-Task, Wait-Task
- Get-Content
- Out-File
- Format-Table, Format-List, Format-Custom (shows everything, particularly useful with Get-View)
- Select-Object, Where-Object
- Help-Object
- Using –whatif and –confirm parameters

8.1.8 Basic PowerCLIcmdlets(used in PowerCLI admin guide)

- Connect-VIServer
- Get-VM, New-VM, Set-VM, Stop-VM, Start-VM, Move-VM (vMotion and svMotion). *NOTE: Set-VM – Snapshot is used to revert a VM to a previous snapshot.*
- Get-VMHost, Add-VMHost, Remove-VMHost, Restart-VMHost,Set-VMHost (used to put host in maintenance mode)
- Get-ResourcePool, Set-ResourcePool, Remove-ResourcePool
- Get-Datastore, Set-Datastore, Remove-Datastore
- Get-DRSRecommendation, Apply-DRSRecommendation
- Get-VMGuest, Suspend-VMGuest, Shutdown-VMGuest, Restart-VMGuest*NOTE: no Start-VMGuest, use Start-VM instead*

8.1.9 AdvancedPowerCLIcmdlets(used in PowerCLI admin guide)

- Get-Folder, New-Folder, Remove-Folder, Move-Folder, Set-Folder. *NOTE: What if there's a blue folder and yellow folder with the same name? How to distinguish? A blue folder can only contain VMs and templates whereas a yellow folder can contain datacenters, hosts, VMs, and clusters. If you check the folder's properties you'll see anIsChildTypeVM property which can be used to distinguish the two types. You can also use the Inventory PSDrive to see the base 'vm' or 'host' folder which corresponds to the VM and Hosts views respectively.*
- Get-Cluster, Set-Cluster (used to enable HA, DRS etc), New-Cluster

- Get-Template, Set-Template, New-Template. *NOTE: Use Set-Template-ToVM to convert a template to a VM rather than clone. To convert a VM to a template rather than clone you have to use 'Get-VM <VMname> | Get-View | \$_.MarkAsTemplate()' (a method exposed by the full object).*
- Get-Task, Stop-Task, Wait-Task (VI tasks such as moving a VM, adding a folder etc)
- New-Harddisk, Remove-Harddisk
- New-Snapshot, Get-Snapshot, Remove-Snapshot
- Get-VMHostAdvancedConfiguration, Set-VMHostAdvancedConfiguration (works with hashtables)
- New-VMHostProfile, Get-VMHostProfile, Apply-VMHostProfile, Test-VMHostProfileCompliance, Export-VMHostProfile, Import-VMHostProfile, Remove-VMHostProfile
- Get-VMHostNetwork, Set-VMHostNetwork (used to get physical NICs in a host, SC configetc)
- Get-VirtualSwitch, Set-VirtualSwitch, Remove-VirtualSwitch
- New-StatsInterval, Get-Stat, Get-StatType
- Get-NICTeamingPolicy, Set-NICTeamingPolicy – used to set teaming policy, not exposed via other cmdlets

8.1.10 Update Manager cmdlets

- Get-Baseline, Attach-Baseline, Detach-Baseline, Remove-Baseline
- Get-Patch, Download-Patch, Stage-Patch
- Get-PatchBaseline, New-PatchBaseline, Set-PatchBaseline
- Remediate-Inventory, Scan-Inventory
- Get-Compliance

8.1.11 Examples

- Get a list of available commands containing the keyword 'Guest' in the VMware snapins
Get-Command -Module vmware -Name *Guest**
- Reverting a VM to a snapshot
Get-VM <myVM> | Set-VM -snapshot <snapshot name> -confirm:\$false
- Creating a 'blue' VM folder
Get-Datacenter<myDC> | Get-Folder vm | New-Folder <myNewFolder>
- Converting a VM to a template
Get-VM <myVM> | Get-View | % {\$_ .MarkAsTemplate() }
- Converting a template to a VM
Get-Template <myTemplate> | Set-Template -toVM
- Putting a host into maintenance mode
Get-VMHost<myHost> | Set-VMHost -state Maintenance
- vMotioning a VM
Get-VM <vm> | Move-VM -Destination <new host>
- Storage vMotion
Get-VM <vm> | Move-VM -Datastore<new datastore>
- Show the possible migration settings for a host
*Get-VMHostAdvancedConfiguration -vmhost<vmhost> -Name Migrate**
- Copy the advanced 'Migration' settings from one host to another
Set-VMHostAdvancedConfiguration -vmhost<destHost> -Hashtable (Get-VMHostAdvancedConfiguration -vmhost<srcHost> -name Migrate)*
- Get the physical adaptors in a given ESX host
Get-VMHost<host> | Get-VMHostNetwork | Select PhysicalNic,Name

Hashtables

Searching a hashtable (key/value pairs, such as those returned by Get-VMHostAdvancedConfiguration) is slightly different from searching scalar values. The code below will only search the keys, not the assigned values. It also doesn't accept wildcards;

```
$hashtable = Get-VMHostAdvancedConfiguration -vmhost<vmhost>  
$hashtable.ContainsKey("Migrate.NetTimeout")      Returns TRUE if the key exists  
$hashtable.GetEnumerator() | Sort Name              Sorts the hashtable
```

Hashtables are used with the Filter parameter for Get-View.

8.1.12 Running tasks inside the guest OS

While the VI API (which PowerCLI gives access to) is very powerful there are some operations it can't do, namely interactions with the guest OS running *inside* a VM (both Windows and Linux). This is where the VIX API comes in, and it's available via the VMRun.exe command and a couple of associated PowerCLI cmdlets.

[What is the VIX and why use it?](#) (VMware blog)

[VMware VIX APIs - Managing and Automating Guest OS](#) (48 min video)

PowerCLI installs v1.6.2 of the VIX API automatically but you can download alternative versions of the software for Windows or Linux [here](#) (PowerCLI only supports v1.6.2 although the latest is 1.10). The latest version supports the following products;

- vSphere 4.1
- Fusion 3.1
- VMServer 2.0
- VMware Workstation 7.1

The VIX API lets you do a few operations that would be impossible with the VI API;

- Run a batch file inside Windows (MS-DOS .BAT files) or Linux (Bash scripts)
- Run a program inside a guest OS
- Copy files to/from a guest (NOTE: the guest does NOT need network access configured so long as the VMTools are running)
- List (and optionally kill) processes in the guest OS

The following cmdlets leverage the VIX API and are only available in vSphere 4.0U1 onwards;

- Invoke-VMScript (-ScriptType can be Powershell (default for Win VMs), Bat or Bash (default for Linux).
- Copy-VMGuestFile
- Get-VMGuestRoute, Set-VMGuestRoute, New-VMGuestRoute, Remove-VMGuestRoute
- Get-VMGuestNetworkInterface, Set-VMGuestNetworkInterface

NOTE: All the cmdlets which utilise the VIX API require authentication on both the host and the guest.

Requirements on the client;

- VMware VIXv1.6.2 must be installed (NOTE: PowerCLI *should* install this version during install but if you run VMware Workstation or Server they may have installed a newer version. You'll need to install this older version if you want the PowerCLI cmdlets to work)
- PowerCLI installed and working

- Only 32bit version of PowerCLI is supported for Invoke-VMScript and Copy-VMGuestFile (but x86 version can be run on an x64 client)

Requirements on the guest

- Must know valid credentials for the ESX server hosting the VM
- Must know valid credentials for the guest OS (local or domain is OK for Windows VMs)
- VM must be powered on
- Help for Invoke-VMScript says that Powershell must be installed, but running against a W2k3 server worked fine without. You may also need to reboot the VM after installing Powershell (according to cmdlet help)
- VMTools should be up to date and running in the guest OS (NOTE. Running against an old version can cause the console to lockup according to the vSphere 4.0u1 release notes)
- The underlying ESX host must be v3.5U2 or greater.

Examples

- *Vmrun -T server -h 192.168.0.60 -u<username> -p <password in plain text!><command>*

Some of the more common commands are listed below (see the VMware docs for all options);
start | stop | restart | suspend | register | unregister | runProgramInGuest | runScriptInGuest

- Copy a file from a guest OS to a local file (using VMTools rather than VM network)
Copy-VMGuestFile -Source c:\test.txt -Destination c:\ -GuestToLocal -VM (Get-VM <vm>) -HostUser root -HostPassword<password> -GuestUser<username> -GuestPassword<password>
- Run a script inside the guest OS
Invoke-VMScript -vm (Get-VM <vm>) -ScriptText "example.bat" -HostUser root -HostPassword<password> -GuestUser<username> -GuestPassword<password> -ScriptType Bat

Some [good examples courtesy of A. Mikkelson](#)

Miscellaneous

VMware Guest Console (available via [VMware Labs](#)) uses VIX API to offer a GUI for the above tasks

[Demo video of using Invoke-VMScript from Carter Shanklin](#) (PowerCLI Product Manager)

You can disable guest operations on a per VM or per host basis but this breaks functionality in VMware Update Manager and potentially other tools too;

- Set "*guest.commands.enable = FALSE*" in the .VMX file
- Set "*guest.commands.enable = FALSE*" in the host-wide configfile

8.2 Administer vCentre Orchestrator

Knowledge

- Identify vCenter Orchestrator requirements
- Identify default Orchestrator plug-ins

Skills and Abilities

- Install and Configure vCenter Orchestrator
 - Configure vCenter Orchestrator database
 - Configure vCenter Orchestrator LDAP connection
 - Configure vCenter Orchestrator vCenter server connections
- Run a Workflow
- Administer Actions, Tasks, Workflows and Policies
- Administer Packages
- Identify appropriate Workflow for a given management activity

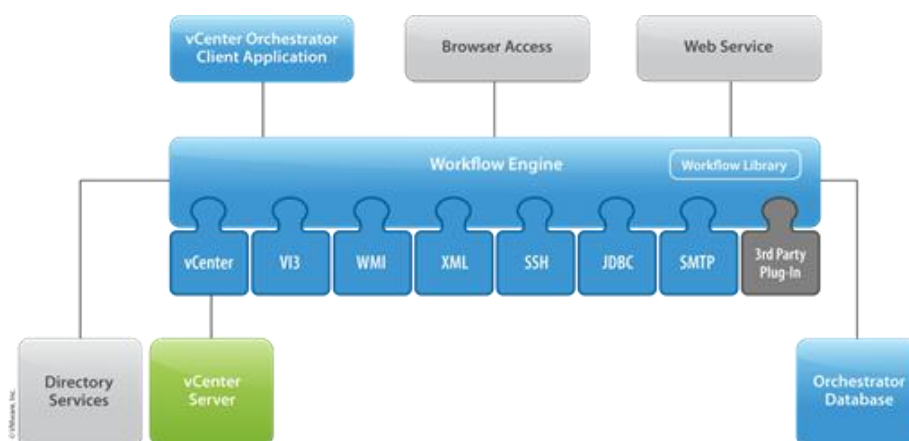
Tools & learning resources

- [Product documentation](#)
 - [vCenter Orchestrator Installation and Configuration Guide](#)
 - [vCenter Orchestrator Administration Guide](#)
 - vCenter Orchestrator Web Configuration
- vCenter Orchestrator Client
- [vCenter Orchestrator homepage](#)
- [vCenter Orchestrator team blog](#), especially their [Learn vCO series](#)
- [Communities roundtable podcast](#) #91
- [An Introduction to VMware Orchestrator](#) (VM5361, VMworld '09)
- MA8030 Saving Time with VMware Orchestrator (VMworld '10 session, sign in required)
- vCO APIs Powerpoint presentation (from [@heyitspablo](#))
- The vCO posts at the ['mighty virtualisation' blog](#)

8.2.1 Orchestrator Requirements

Architecture

vCO is an automation platform which integrates with vCenter, an LDAP directory (for delegation of permissions) and offers a plugin architecture to facilitate automation over a wider ecosystem. It consists of a client, a server, and database components and comes with some plugins ready to go;



Server Requirements

- Included with vCentre Standard (Foundation & Essentials plus are read only workflows)
- Recommended - 4GB RAM, 2 vCPUs, 5GB disk. vCO server and database should be on separate servers.
- Windows 2003 (x64 only) or Windows 2008 (x32 or x64) - [VMware compatibility matrix](#) (p17-18)
- *Database support:* SQL 2005 (32 bit), SQL 2008 (64 bit), Oracle 10gR2 (SQL Express not supported but works). vCO v4.1 supports Oracle 11g.
- *LDAP support:* Active Directory 2003+, Novell eDirectory, Java Directory Server Enterprise Edition

Client requirements

- There are two clients – one bundled automatically when you install vCenter and the standalone client
 - The bundled client (both v4.0 and v4.1) won't run on Windows 2003 32bit. Not an issue for 4.1 as vCentre requires 64 bit, but could be issue if your v4.0 vCentre server is already 32bit. Check this [post in VMware communities](#).
 - The standalone client will only run on x32 Windows.
- Check compatibility matrix (page 19);
http://www.vmware.com/pdf/vsphere4/r40/vsp_compatibility_matrix.pdf
- NOTE: End users (rather than Orchestrator administrators) access Orchestrator through a webclient, so browser compatibility is all that's required (IE7+, FF 3+)

General considerations

- Orchestrator (client and server) are installed automatically with vCentre but are NOT preconfigured.
- A 4.1 vCO can run against an older v4.0 vCentre but 4.1 specific workflows may not work
- Maximums
 - 10 - vCenter servers per vCO
 - 300 – ESX/ESXi hosts per vCO
 - 15000 – connected VMs
 - 150 – simultaneous workflows
- Works OK with vCenter Linked mode

8.2.2 Install and configure Orchestrator

Read the [vCO blog entry on installing vCO](#) and [watch the video](#)

- Create Orchestrator database (step 1 of 2)
 - Check dynamic port setting (1190 if using out of the box vCentre setup)
 - Create new database
- Start VMware Orchestrator Configuration service (set to Manual)
- Login to webpage (<http://server:8282>) using default vmware/vmware credentials
- Configure networking
 - IP Address (leave ports on defaults)
 - Import SSL certificate from integration services (vCentre, LDAP). This ensures that vCO uses secure communication when talking to these services.
- Configure LDAP
 - Define DCs
 - Base LDAP path for Users
 - Base LDAP path for Groups
 - LDAP path to vCO Administrators group
 - Test Login
- Configure database connection then create database tables (step 2 of 2)

- Configure DB type (SQL or Oracle)
- Port (SQL default 1433, vCenter default 1190)
- Username/password (domain credentials)
- Database name
- SQL Instance
- Windows domain
- Create database tables - click on 'Install the Database'
- Configure certificate services for the vCO server itself (used to sign packages for third parties etc)
 - Create self signed certificate (if you don't already have one)
- Licence the server
 - Can use vCenter licence (specify host, username/password)
- Add plugins
 - Enter credentials (need to be member of vCO Admin group)
- Configure mail services (to allow user notifications etc)
 - Enter domain username (user@domain format)
- Restart the vCO *configuration* server (under startup options)
- Configure integration with Virtual Center
 - Login and go to the Virtual Center tab
 - Add your vCentre host (credentials DON'T user @domain.com)
 - Use 'share a single session' – this is recommended best practice (vCO team blog). Determines how many connections are made from the vCO server to vCenter.
- Install vCO as a service
- Start vCO
 - Best practice is to now backup the configuration
- COMPLETE!

Orchestrator uses javascript as it's language of choice but from 4.0U2 you can configure Orchestrator to allow access to local OS commands (such as call a Powershell script on the same host as the vCO server). See this [post from the VMware communities](#) (or the Orchestrator Admin Guide p66). Can also call vbscript – see this [thread](#).

C:\Program Files\VMware\Infrastructure\Orchestrator	install path
vCenter_Orchestrator_InstallLog.log	Install logfile (created in above path)
VMware vCenter Orchestrator server	new service created during config

8.2.3 Orchestrator plugins

Default plugins

The default plugins can be seen via the vCO configuration webpage on the Plugins tab (and also on the architecture diagram at the start of these notes). Each plugin populates vCO with predefined workflows as described below;

- Common enumerated types
- SSH – allow you to issue remote command and file transfer sessions using either passwords or PKI
- Net – provides FTP, POP3 and Telnet functionality.
- Database – provides JDBC connectivity to a wide range of databases
- vCO WebOperator – adds a default webview for end users
- vCO Library – workflows to automate vCO itself (create workflow, create task etc)
- Mail – send and receive SMTP emails in your workflows
- XML – an XML parser to you can use XML import/export in your workflows
- vCenter Server – access to the vCenter API allowing you to automate your virtual infrastructure

Other plugins

- VMware Lifecycle manager is implemented as a plugin for vCO. It's contains workflows for managing VM lifecycle.
- VMware's new Cloud Director is built on top of vCO
- Netapp are working [on various plugins to automate storage provisioning](#) with Cloud Director.
- The VMware VIX API has also been exposed via a plugin. See the [VMware Labs](#) for details.
- There is a Microsoft plugin which allows use of WMI and querying AD.

NOTE: Some plugins require licences (although the built-in ones do not). These licences can be installed via the Plugins tab.

8.2.4 Administer Actions, Tasks, Workflows and Policies

This is a tricky objective – understanding how to *create* workflows, actions and policies is really a development task and beyond the scope of most administrators (according to VMware) so presumably you only need to know how to use the existing ones. The vCO team blog has a worked example for creating a provisioning portal (which was used at the blueprint for the Orchestrator lab at VMworld 2010) so I've based my studies on the functionality covered in those tutorials.

Actions (discrete steps such as creating a folder, delete a file etc)

- Used as building blocks for workflows. Actions consist of javascript code.
- Create, Duplicate, Export, Import, Move. If you move an Action to a different module you may have to use the 'Find Elements that Use This Element' function to avoid breaking existing functionality.

Workflows (*a series of actions and decisions which run sequentially until they reach an endpoint*)

- User Permissions – View, Inspect (see schema and scripts), Execute, Edit, Admin. Cumulative.
- Credentials - For the full client or webviews the workflow runs under the context of the user who runs the workflow (Admin guide, p30). Nested workflows run under the context of the calling workflow. A scheduled workflow can be set to run as an alternative user.
- Attributes – static data or dynamically generated data used within the workflow (doesn't involve the end user)
- Parameters – data passed into or out of a workflow. Typically user defined but can be passed in via another workflow
- Orchestrator comes with a built-in set of workflows which you can use 'as is' or as a base for your own custom workflows. These built-in workflows are read only – you have to duplicate and then modify if required.
- Workflows can be nested.

Tasks

- There's no mention of tasks in either the install or admin guides, so I guess this refers to scheduled Workflows?

Policies

- Event triggers based on activity of the system.
- There are no policies or policy templates by default
- NOTE: Policies are deprecated as of vCO v4.1.

8.2.5 Running Workflows

- Workflows can be run in various ways – via the vCO client (either interactively or via the scheduler) or via a webview (again interactively or via a schedule).

- 'weboperator' is the default webview. Not enabled by default (right click, Publish)
- You may need to specify input values (depending on the workflow)
- NOTE: You can only run VMware workflows once the vCenter plugin is enabled.

You can check progress of a Workflow via the Schema tab or the Events tab for a given workflow.

8.2.6 Administer Packages

- Packages contain workflows, policies, web views, resources and actions.
- Export packages to migrate content to another Orchestrator server.
- Import packages from third parties to extend the functionality of Orchestrator. For example there are no built in workflows to manipulate virtual switches, but you can use [downloadable workflows on the Orchestrator communities page](#) to add this functionality.
- You can apply DRM to packages to dictate what others can do with the contents.

8.2.7 Identify appropriate Workflow for a given management activity

You can search existing Workflows by name using the Orchestrator API search (Tools -> API Explorer)

There is a workflow for exporting Orchestrator logs for diagnostic purposes

8.3 Administer vSphere using the vMA

Knowledge

- Identify vMA prerequisites
- Identify vMA specific commands
- Determine when vMA is needed

Skills and Abilities

- Install and Configure vMA
- Add/Remove target servers
- Perform updates to the vMA
- Use vmkfstools to manage VMFS datastores
- Use vmware-cmd to manage VMs
- Use esxcli to manage Storage Multipathing
- Troubleshoot common vMA errors and conditions

Tools & learning resources

- [vMA Homepage at VMware.com](#)
- vSphere Management Assistant Guide
- vSphere Command-Line Interface Installation and Scripting Guide
- vSphere Management Assistant
 - vifp
 - vima-update
- vSphere CLI
 - vicfg-*
 - vmkfstools
 - esxcli
 - vmware-cmd
- [vFail.net's VCAP Study guide for vMA, section 8.3](#)
- [Bridge the ESX/ESXi Management Gap with VMware vMA \(MA6580, VMworld 2010\)](#)
- VMworld 2009 session [TA2659- Managing ESX in a COS-less world](#)
- [Using AD authentication with vMA \(Geeksilver's blog\)](#)
- [Using AD authentication with vMA \(VirtualGhetto blog\)](#)
- [Why you should upgrade vMA to v4.1 \(VirtualGhetto blog\)](#)
- [vMA configuration examples](#)
- [Using vMA as a PXE boot server](#)
- [The vMA community forums](#)

Main uses:

- syslog server (vlogger component)
- centralised scripting repository
- replacement for ESX service console (scripts and third party plugins)
 - easier to port service console scripts rather than converting to PowerCLI
 - scripts may need amending (new authentication methods etc)
 - facilitates move to ESXi

8.3.1 Prerequisites, installation and updating

Prerequisites;

- ESX host must support 64-bit VMs (Intel EM64T and vT technology)
- ESX 3.5U2 onwards, vCentre 4.0 onwards (2.5 NOT supported)
- 512MB, 1vCPU, 5GB+ disk space
- Two deployment methods;
 - Import OVF directly from VMware
 - Download ZIP file, extract, then Import from local OVF

8.3.1.1 Installation

Installed components;

- x64 Red-Hat compatible Linux distribution (CentOS 5.3 in latest 4.1 release)
- vSphere CLI + vi-fastpass component (not supplied with standard vCLI installation)
- Syslog component (vilogger)
- SMTP server (for monitoring vMA itself), vSphere SDK for Perl, Java JRE 1.5
- Sample scripts in /usr/bin/

During install you'll need to provide network details (DHCP or static) as per any appliance along with a password for the vi-admin account.

NOTE: If you're installing the appliance via .OVF and you're behind a firewall the deployment will fail. You need to use a command line utility OVFTOOL (separate [download](#)) to specify a proxy;

```
ovftool --proxy=user:pass@proxy.example.com http://external-site.com/ovf/package.ovf
```

8.3.1.2 Upgrading vMA

Use VIMA-UPDATE

- Find available updates;

```
[vi-admin@vma2 ~]# sudo vima-update scan
Applicable bulletins with updates are listed.
Bulletin ID ---Date--- -----Summary-----
VIMA410-GA  2010-07-13  VIMA 4.1 GA update
[vi-admin@vma2 ~]# _
```

- Show detailed information about updates (inc packages updated etc);

```
[vi-admin@vma2 ~]$ sudo vima-update info
Password:
Id - VIMA410-GA
Releasedate - 2010-07-13T00:00:00-08:00
Vendor - VMware, Inc.
Summary - VIMA 4.1 GA update
Severity - critical
Category - critical
Installdate -
Description - This upgrade patch updates vMA 4.0 appliance to vMA 4.1 GA. See
more details at http://www.vmware.com/support/developer/vima/
Kburl - http://kb.vmware.com/kb/
Contact - http://www.vmware.com/support/contacts/
List of constituent VIBs:
rpm_python_2.4.3-24.el5_3.6@x86_64
rpm_lsof_4.78-3@x86_64
rpm_slang_2.0.6-4.el5@x86_64
rpm_nss_ldap_253-17.el5@x86_64
rpm_net-tools_1.60-78.el5@x86_64
rpm_bind-libs_30:9.3.6-4.P1.el5_4.1@i386
rpm_becrypt_4.1.2-10.1.1@x86_64
rpm_mailcap_2.1.23-1.fc6@noarch
rpm_chkconfig_1.3.30.1-2@x86_64
```

- `sudo vima-update update` OR `sudo vima-update -b <bulletin>`

NOTE: You can also update vMA from a local repository (all versions) or an offline bundle (4.1 onwards). See this [VMware thread](#) for details.

NOTE: If you're behind a proxy you'll need to configure it in the vimaupdate.conf config file

NOTE: You can't upgrade from vMA 1.0 (also known as VIMA) to vMA 4.0.

You can check the installed vMA version via the appliance status screen in the VI client (only available when deployed via http (not local) OVF, not updated when using vima-update). You can also check version info via the main config file, or via the splash screen on the appliances console session, or by using `'cat /etc/vma-release'`.

8.3.2 Configuring vMA

vMA commands;

- vima-update
- vifp
- vifptarget (vifpinit in v4.0)
- vilogger
- domainjoin-cli (vSphere 4.1 onwards)

Configure NTP - use chkconfig to enable the service, edit /etc/ntp.conf and /etc/ntp/step-tickers to set servers to use.

Configure timezone to use UTC (ESXi uses UTC)- details here

Optionally configure keyboard

/etc/vma-release	vMA version information
/etc/vmware/viconfig/vilogdefaults.xml	Main config file vSphere v4.0
/etc/vmware/vMA/vMA.conf	Main config file vSphere v4.1
/etc/vmware/esxupdate/vimaupdate.conf	vima-update config file
/home/vi-admin/.vmware/credstore/vicredentials.xml	Credential store

vifs doesn't support wildcards or the concept of a working directory (hence absolute paths must be used).
vifs works much like the datastore provider in PowerCLI – see section 8.1

Miscellaneous

- Always use FQDN to specify servers
- vma-help – useful information on vMA and vCLI commands
- It's perfectly possible to have multiple vMA appliances and for the same servers to be registered via vifp on them all. Likewise you can have multiple vMA's all collecting logs (via vilogger) from the same hosts.

8.3.3 Authentication using Fastpass

This allows unattended authentication for ESX hosts and vCentre. You can then to run tasks against them without further authentication (cron jobs, scripts etc). It's a two step process;

1. Setup the servers you want to authenticate with using VIFP
2. Enable the connections using VIFPINIT (for v4.0) or VIFPTARGET (for v4.1)

NOTE: the vifp connections (step 1) are persistent across reboots, but step 2 will need repeating after every logout/reboot.

Step 1 – Enable authentication for the servers

Adding servers to the Fastpass component;

```
[vi-admin@zcgprvma01 ~]$ vifp addserver --server zcgprvsh02.mfl.co.uk  
root@zcgprvsh02.mfl.co.uk's password:  
[vi-admin@zcgprvma01 ~]$
```

List servers currently configured with Fastpass;

```
[vi-admin@zcgprvma01 ~]$ vifp listservers -l  
zcgprvsh01.mfl.co.uk    ESX    fpauth  
zcgprvsh02.mfl.co.uk    ESX    fpauth
```

Remove servers from Fastpass;

```
[vi-admin@zcgprvma01 ~]$ vifp removeserver --server zcgprvsh02.mfl.co.uk  
[vi-admin@zcgprvma01 ~]$ vifp listservers -l  
zcgprvsh01.mfl.co.uk    ESX    fpauth
```

Reconfigure the authentication parameters;

```
[vi-admin@zcgprvma01 ~]$ vifp reconfigure zcgprvsh01.mfl.co.uk --username vmware --password Virt  
Error: Insufficient permissions for vmware@zcgprvsh01.mfl.co.uk
```

NOTE: When you add a new ESX/i host using Fastpass two users are created locally on the host;

- vi-admin (administrator access)
- vi-user (read only)

These users have no shell defined in /etc/passwd so can't be used interactively on the host.

Step 2 – set target servers

Set the default target server;

Show the current target;

Remove a target server;

Miscellaneous

- The internal password used within the Fastpass system is rotated every week by default.
- Only ESX 3.5U2 onwards is supported.
- 'vifp recoverserver' can be used to fix a corrupted credential store. See Troubleshooting section.

8.3.4 Authentication using AD (vMA v4.1 onwards)

Prerequisites;

- Join vMA to a domain using `domainjoin-cli`. If you want to execute commands against ESX hosts directly the hosts also need to be joined to the domain for AD authentication to work.

```
[vi-admin@vma2 ~]$ sudo domainjoin-cli join thefunkysite.com egrigson
Password:
Joining to AD Domain:  thefunkysite.com
With Computer DNS Name: vma2.thefunkysite.com

egrigson@THEFUNKYSITE.COM's password:
Warning: Unknown pam module
The likewise PAM module cannot be configured for the wbem service. This service
support and include a copy of /etc/pam.conf or /etc/pam.d.

Warning: A resumable error occurred while processing a module
Even though the configuration of 'pam' was executed, the configuration did not

SUCCESS
[vi-admin@vma2 ~]$ █
```

Once the vMA appliance is joined to a domain you can login using your domain credentials (domain\user) instead of vi-admin.

NOTE: If you join the vMA appliance or ESX hosts to your domain they need to be licenced according to the usual Microsoft policy.

8.3.5 Using VILogger (syslog server)

Two step process;

1. Enable authentication for the hosts you want to collect logs from using Fastpass.
2. Configure and enable logging (all servers, per server or per logfile)

Turn on logging for all vifp enabled hosts;

Enable logging for a specific host;

```
[vi-admin@vma2 ~]$ vilogger enable --server vcentre.thefunkysite.com

Target Server: vcentre.thefunkysite.com
vpxd          ... Enabled
```

Display the logging settings;

```
[vi-admin@vma2 ~]$ vilogger list

Target Server: vcentre.thefunkysite.com
Log           State      Status           CollectionPeriod NumRotation
(Seconds)
vpxd          Enabled    Collecting       10              5
```

Turn off logging for a specific host;

```
[vi-admin@vma2 ~]$ vilogger disable --server vcentre.thefunkysite.com  
  
Target Server: vcentre.thefunkysite.com  
vpxd ... Disabled
```

Update logging configuration;

`vilogger updatepolicy`

Miscellaneous

- With ESXi prior to 4.1 the vpxa.log files were not sent by default to a syslog server. See [VMware KB1017658](#) for details.
- `service vmware-vilogd restart`
- Timestamps – uses UTC on ESXi
- `tail -f /var/log/vmware/<FQDN of Host>/vpxa.log`
- `vim-cmd` is NOT available as it is with ESX/ESXi hosts. This means you can't enable lockdown mode for example.
- Lockdown mode breaks the fastpass authentication.

8.3.6 Troubleshooting vMA

- Enable DNS
 - `edit /etc/resolv.conf`, restart networking.
 - Edit `/etc/sysconfig/network` to include domain
- Customizing timezone and keyboard settings (timestamps important when used as syslog) – [VMware KB1007551](#)
- Resolving credential store corruption – [VMware KB1010178](#)
- Whenever the Fastpass authentication is performed an event is logged in vCentre. If the Fastpass authentication is incorrectly configured it will try to authentication continuously, generating up to five failure events per second. This can quickly fill the vCentre database as Event data is retained for ??? by default. This can be changed via the Administration -> vCentre Server Settings -> Database Retention Settings in vCentre
- When decommissioning a vMA appliance you should clear down the VIFP server list. This will remove the local vi-admin and vi-user accounts on the ESX/i hosts.

9 Advanced installations

9.1 Install ESX server with custom settings

Knowledge

- Identify Service Console memory defaults and maximums
- Identify default and optional ESX partitions

Skills and Abilities

- Configure optional ESX partitions during installation
- Install/uninstall custom drivers
- Configure advanced bootloader options
- Configure kernel options
- Given a scenario, determine when to customize a configuration

Tools & learning resources

- Product Documentation
 - [ESX and vCenter Server Installation Guide](#)
- vSphere CLI
 - [vicfg-advcfg](#)
 - [vicfg-module](#)
- vFail.net's [study notes](#)
- [Notes on partition sizes at Yellow Bricks](#)

While the blueprint only refers to installing ESX (not ESXi) I've covered both in anticipation of the VCAP-DCA labs going to 4.1.

9.1.1 When to use a customised installation

There are plenty of reasons to use some advanced installations;

- Your hardware isn't supported in the 'out of the box' setup so you need custom drivers
- You want to streamline the deployment process by building a custom install CD, including some post configuration steps, or utilising PXE boot etc
- You want to gain maximum performance from every host, which means performance and configuration tweaks (vmKernel parameters, service console memory settings etc)

9.1.2 Installing ESX/ESXi

- Interactive installations can be done via the GUI or text mode.
- The installer can be located on CD/DVD, USB flash or via a PXE boot. While PXE is typically used for scripted builds it can be used as a source of installation files for an interactive build.
- Scripted methods (PXE boot using HTTP, FTP, NFS are covered in section 9.2.
NOTE: Scripted installs of ESXi were only added to v4.1 – prior to that only ESX classic could be scripted.
- USB flash devices can be used in two ways;
 1. You can run ESXi (but NOT ESX classic) directly from an attached USB drive (provided your BIOS sees it as bootable). To do this simply choose the USB drive as the install point when prompted by the ESX installer.

2. You can use the USB drive as the source for installation media and scripts
- To install a virtual ESX host on ESX (for lab testing) you need some specific configuration tweaks – see the [article at vCritical](#) for full details.
 - For 64 bit guests you must have a 64 bit chip with Intel-VT support enabled or an AMD chip of revision E or later. [Wikipedia has details](#) and you can check using [VMware's CPU Identification Utility](#). You cannot run nested 64 bit VMs.
 - Boot from SAN is now supported for ESXi (4.1 onwards). This includes iSCSI and FCoE for a limited set of supported adapters.
 - Both ESX and ESXi v4.0 will erase all local partitions by default, including existing ESX installs and VMFS partitions (if you're upgrading an older ESX version for example).

9.1.2.1 Service Console memory defaults and maximums

Default is 300MB for servers with up to 8GB RAM then scales towards the maximum of 800MB for a host with more than 128GB RAM. VMware recommend you don't change manually, although it is generally recommended practice to size your SWAP partition to 1600MB so that you don't need to rejig partitions when increasing the memory size at a later date.

This is all [documented on Duncan Epping's Yellow Bricks site](#).

9.1.2.2 Configure optional partitions during install

- Only ESX classic has a user definable partition layout – ESXi uses an automatic partitioning scheme which the user can't change. See this [blogpost at Jason Boche's site](#), and another [blog article at Geeksilver's site](#) (and [one more](#) for luck).
- From the installation guide - *"You cannot define the sizes of the /boot, vmkcore, and /vmfs partitions when you use the graphical or text installation modes. You can define these partition sizes when you do scripted installations"*. See the [ESX and vCenter Server Installation Guide](#) (chapter 6) for full details.
- NOTE: the '/boot' partition used to be only 100MB (with ESX 3.5 and prior) but has now been increased (to 1.25GB) to allow a future upgrade from ESX to ESXi

Mount point	Default size	Recommended size	Filesystem type	Physical location
/boot	1100MB	1100MB	Ext3	Physical partition
vmkcore	100MB	100MB	vmkcore	Physical partition
n/a	1.2GB+	1.2GB+	VMFS	Physical partition which fills any remaining space on the physical disk
The following are all in the VMFS volume created during install				
/	5GB+ (various depending on disk size)	Leave as default	Ext3	VMDK in the VMFS volume
esxconsole.vmdk	1.2GB	Leave as default	Ext3?	VMDK in the VMFS volume
swap	600MB	1600MB	swap	VMDK in the VMFS volume
/var	2GB (/var/log)	5GB	ext3	
/opt	Optional	2GB	ext3	
/home	Optional	512MB	ext3	
/tmp	Optional	1GB	ext3	
/usr	Optional	None specified	ext3	

9.1.2.3 Configure advanced bootloader options

- Can be used to set a password on the bootloader (which is then requested if you want to change kernel parameters at boot time. See [this article on securing ESX](#) for details)
- By default the GRUB bootloader is installed in the MBR (master boot record). Some legacy hardware stores BIOS info in the MBR so in these cases you have to install GRUB on the first partition of the disk instead.
- Allows you to specify kernel parameters, which are written to the GRUB.CONF file to ensure they're persistent across reboots. I couldn't find much documentation on how these are used with ESX – Eric Sloof has a post around [possible kernel parameters with ESX3i](#).
NOTE: One use is with interleaved NUMA nodes as described in [VMwareKB1021454](#).

9.1.2.4 Configure kernel parameters

Typically used with scripted installs (see section 9.2).

9.1.2.5 Install/uninstall custom drivers

There are occasions when required drivers are not included in the ESX or ESXi builds (for example the HP 375T quad port 10GB NIC which requires custom drivers for the HP BL460c G6 and G7 blades). With vSphere you can now add drivers either during installation (ESX only) or post install (both ESX and ESXi) although the procedure is different for each;

At install time

- Start the installation as usual – this can be interactive or scripted
- When prompted, reply Yes to 'install additional drivers?'. You can either embed them in a custom ESX .ISO or provide a separate CD.
- NOTE: You can't add custom drivers when using a PXE install (see chapter 3 of install guide)

Post-install

- Check [online HCL](#) to determine the driver required. Provides link to .ISO.
- Download drivers from VMware, check signature using `md5sum <drivers.ISO>`
- Use `esxupdate` for ESX classic hosts or `vihostupdate` (vMA, vCLI etc) for both ESX and ESXi hosts
 - `esxupdate -bundle <drivers-file.zip> update`
 - `vihostupdate [options]—install -bundle <drivers-file.zip>`
- Typically a reboot of the host is required (this is normally indicated in the release notes). Some bundles will require the host to be in maintenance mode before application – this will be indicated where necessary.

Instructions for this are summarised in this [VMware post](#), or you can read more in the [VMware Patch Management Guide](#).

Querying existing drivers;

- `esxupdate query -vib-view` (used to get the driver name)
- `ethtool -l vmnicX` (shows the driver version in use. This only works for ESX. For an ESXi version refer to [VMware KB 1027206](#))

Uninstalling a driver uses a similar syntax;

- `esxupdate --bundle <driver filename> remove` (for example `esxupdate --bundle tg3 remove`)

NOTE: Some driver updates are provided in a metadata file rather than a bundle. This is simply an XML file which point to the actual bundle to use. When using metadata use `--metadata` not `--bundle` in the command.

9.1.3 Post install configuration

Despite the functionality offered during the install there is often further configuration required. The tools available to do this vary depending on the deployment scenario;

- PowerCLI – can't run natively during scripted installs but can be used via a server based component which 'listens' for a completed build and finishes the post build tasks. Links?
- vCLI – any vCLI commands can be included in %pre and %post
- Host profiles (see chapter 5, Operations Maintenance)

vicfg-module

This is used to configure advanced VMkernel options. Typical uses;

- to set queue depth for HBAs ([VMware KB1267](#))
- enabling Netqueue (see blueprint section 2.1 on Networking for details)
- for fixing faulty drivers (see [VMware KB1029070](#)). This will also be relevant to section 6.4, troubleshooting storage performance and connectivity.

Syntax

```
vicfg-module [<connection_options>]
  [--get-options <module_name> |
  --help |
  --list |
  --set-options "<option> <value>" <module_name> |
  --vhost <esx_host> ]
```

NOTE: Setting options via `esxcfg-module` is NOT cumulative. If you only specify one option in the command it will clear any other previously set parameters.

Examples

vicfg-module --get-options

Shows the enabled options for a module. NOTE: This doesn't show the available options, only the enabled ones. To get a list of possible values use the older `vmkload_mod -s <module>`.

vicfg-module -d vmfs2

Disable the module, preventing it from reloading after a reboot. Can be used in conjunction with `-u`.

vicfg-module -u vmfs2

Unload the module immediately. Would potentially be re-enabled at reboot unless the `'-d'` option was also used.

vicfg-advcfg

This is a vCLI command to configure advanced parameters on ESX and ESXi hosts, equivalent to the host Configuration/Advanced settings you'd configure with the VI client. Typical uses;

- enable or disable CIM providers (ESXi)
- configure a host during a scripted build – set NFS options as per [this blogpost from Xtravirt](#)

Syntax

```
vicfg-advcfg <connection_options>
```

```
[--default <value> |  
--get <path> |  
--get-kernel <boot_parameter> |  
--quiet |  
--help |  
--set <value> <option> |  
--set 0|1 UserVars.CIMEnabled |  
--set 0|1 UserVars.CIMOEMProvidersEnabled |  
--set 0|1 UserVars.CIMCustomProvidersEnabled |  
--set-kernel <value> <boot_parameter> |  
--set-message <message> |  
--list  
--vihost <esx_target>]
```

Examples

```
vicfg-advcfg --server esx01.vExperienced.co.uk --get LVMDisallowSnapshotLun
```

```
vicfg-advcfg --server esx01.vExperienced.co.uk --set 2 LVMDisallowSnapshotLun
```

```
vicfg-advcfg --server esx01.vExperienced.co.uk --get-kernel
```

```
vicfg-advcfg --server esx01.vExperienced.co.uk --set-kernel
```

Links to find out more;

http://www.vm-help.com/esx/esx3i/esx_3i_rccli/vicfg-advcfg.php

<http://it-john.com/home/technology/vmware/esxcfg-advcfg/>

9.1.4 Further reading

Custom ESX CD/DVD – if you want to have the flexibility of a scripted install combined with a custom ESX CD there are various sites with instructions ([here](#) and [here](#)).

9.2 Plan and execute scripted installations

Knowledge

- Identify default installation scripts
- Identify boot options for scripted installation

Skills and Abilities

- Perform a scripted ESX Host installation
- Perform a scripted ESXi Host installation
- Configure media repository
- Edit installation script parameters
- Configure pre/post script tasks
- Evaluate use cases for scripted installation

Tools & learning resources

- [ESX and vCenter Server Installation Guide](#)
- [ESXi and vCenter Server Installation Guide](#)
- ks-first-safe.cfg
- ks-first.cfg

The blueprint for this section seems to refer mainly to ESX but I've described both ESX and ESXi on the assumption the lab environment used for the exams will move to v4.1 sooner rather than later.

NOTE: Weasel is VMware's scripted installer. It's similar to Kickstart as used with Linux, but not identical.

A summary for a scripted install;

- Decide where to load the bootloader from
- Configure a media repository to hold your source files and scripts
- Create an install script (either from scratch or from a previously built host)
- Perform the scripted install

9.2.1 Use cases for scripted installations

Reasons to use a scripted install;

- Reduce deployment time
- Ensure consistency, reduce human error
- Remove need for local media (when using PXE boot. Very useful for blade and remote environments)
- Delegate installations to junior staff who don't know how to configure ESX

Along with knowing why you might use a scripted install in the first place you should consider the various types of scripted install and when to use each one. Factors to consider;

- **Maintainability.** Over time you'll want to update your install for new releases of ESX, patches, post install steps etc. While a custom CD has the least dependencies it's harder to maintain compared to a network media repository.
- **Dependencies.** I created an NFS based install only to find that most of the time the host's physical networking hasn't been completed when we want to build the OS, rendering this method useless. I had to convert it to a custom CD instead which was mounted via ILO (it was a blade environment).

Another example is USB flash – it's easier than CD to amend/update but won't be much use for remote installs.

9.2.2 Boot options

You can boot from the following (regardless of where your media repository resides);

- CD/DVD
- USB Flash
- PXE/DHCP (boot from network)

To prepare a USB device as a bootable ESX installation source you'll need to run SYSLINUX on it – see page 25 of the [ESXi and vCenter installation guide](#) or this [blog article at Ubiquitous Talk](#).

gPXE is an open source implementation of the PXE standard. LINUXPXE is largely similar, although gPXE improves the process by allowing some files to be transferred via HTTP rather than TFTP. This can be more reliable when network traffic is heavy as HTTP is a more reliable protocol compared to TFTP.

To use PXE you'll need a PXE server (most modern ones also include support for gPXE), DHCP server, TFTP server and possibly a webserver (if you use gPXE) though they all commonly reside on the same server instance. There are various choices if you don't already have these services setup;

- add DHCP, TFTP etc to the vMA (see this [blog post](#))
- download an appliance from the VMware Marketplace (such as [V-PXE server](#) – thanks Simon Long!)
- Use the [VMware labs 'stateless ESX' project](#) (with another [great post from Simon Long...](#))

For full details of preparing for a PXE based boot see the [ESXi and vCenter installation guide](#) installation guide p29-33. You can also use custom appliances such as the UDA and EDA – see the links at the end of this article for details.

9.2.3 Configure media repository

Media repository (including installation script) can reside on;

- CD/DVD
- USB Flash (ESXi only)
- HTTP/HTTPS (an [example for setting up HTTP media repository on VirtualKenneth's site](#))
- NFS
- FTP

NOTE: A scripted install to USB devices isn't currently supported (see [VMware article](#) under Installation and Scripting) but it can be used as a repository for the source files.

NOTE: You can't use HTTPS with a proxy server

NOTE; A media repository must contain the entire contents of the ESX/ESXi DVD.

NOTE: To prepare a USB stick as a media repository format a partition (with FAT32) and copy on the ESXi DVD contents.

NOTE: There are various files you can customise in the ESX bootloader – see this [blogpost by Mike LaSpina](#) for details.

NOTE: It doesn't seem easy to 'slipstream' drivers to a scripted install. There's a [good post by Patrick Van Beek](#).

9.2.4 Default installation scripts

There are scripts provided by VMware which you can use with no customisation – just boot from the DVD and choose the relevant option;

For ESX;

- ks-first.cfg – Installs ESX to the first detected hard disk. NOTE: all existing installs (including existing VMFS volumes) will be overwritten.
- ks-first-safe.cfg – Installs ESX to the first detected hard disk, but preserves existing VMFS volumes.

For ESXi;

- ks.cfg. Install ESX on the first detected hard disk. NOTE: all existing installs (including existing VMFS volumes) will be overwritten.

With all three scripts the default root password is 'mypassword'.

NOTE: These default installation scripts can be found at /etc/vmware/weasel/ks.cfg (although these are embedded in the ienvron.gz file – see here for [details of extracting it](#)). For ESX they're embedded in the initrd.img file - see [VMware KB 1018990](#) for details of how to extract them.

9.2.5 Creating a custom installation script (installation script parameters)

The installation script is where you specify the final configuration of your host. It's stored in the media repository and referenced in the command line issued to the bootloader (see next section).

Key script parameters (there are plenty of others – see the install guide);

- install
- partition
- network
- rootpw
- %pre and %post sections

NOTE: Rather than creating a script from scratch you can configure a baseline host and then copy the ks.cfg file from /root /ks.cfg (ESX classic only. I couldn't find any equivalent on an ESXi host). This will only capture settings applied during installation of the original host – not settings applied via Host Profiles or manual configuration for example.

NOTE: The bootloader and installation script parameters are NOT the same as Kickstart for Linux. Differences include using VMKNIC with the ksdevice parameter instead of ETH plus the partition command is totally different. See p45-48 of the [ESXi and vCenter installation guide](#) for details.

NOTE: As network configuration (hostname, IP etc) is in the install script you need a script file per host. As an alternative you can use the %pre and %post sections to automate this yourself.

NOTE: The acceptable script parameters changed from ESX v3 to v4 (details in [VMware KB 1010212](#))

Example script for ESX classic;

Specify NFS as the media repository

```
install url nfs://10.1.118.91/vol/v_global_nfs_install/install/esx/sourceRPMs/4_0u1/
```

```
rootpw --iscrypted $1$E/cRXY09$k6UP6IdzPEHdBRdD09Yq50  
timezone --utc 'Europe/London'
```

Specify network settings

```
network --addvmportgroup=false --device=vmnic0 --vlanid=118 --bootproto=static --
ip=10.1.118.165 --netmask=255.255.255.0 --gateway=10.1.118.1 --
nameserver=10.1.100.15,10.2.100.15 --hostname=zcgprvsh06.test.co.uk
```

Specify partition layout

```
part '/boot' --fstype=ext3 --size=1100 --ondisk=mpx.vmhba0:C0:T0:L0
part 'none' --fstype=vmkcore --size=110 --ondisk=mpx.vmhba0:C0:T0:L0
part 'zcgprvsh06-localstorage' --fstype=vmfs3 --size=17967 --grow --
ondisk=mpx.vmhba0:C0:T0:L0
virtualdisk 'esxconsole' --size=16967 --onvmfs='zcgprvsh06-localstorage'
part 'swap' --fstype=swap --size=1600 --onvirtualdisk='esxconsole'
part '/var' --fstype=ext3 --size=4096 --maxsize=4096 --grow --onvirtualdisk='esxconsole'
part '/opt' --fstype=ext3 --size=2048 --maxsize=2048 --grow --onvirtualdisk='esxconsole'
part '/home' --fstype=ext3 --size=2048 --maxsize=2048 --grow --onvirtualdisk='esxconsole'
part '/tmp' --fstype=ext3 --size=2048 --maxsize=2048 --grow --onvirtualdisk='esxconsole'
part '/' --fstype=ext3 --size=5120 --maxsize=5120 --grow --onvirtualdisk='esxconsole'
```

Specify post install instructions

```
%post --interpreter=bash
```

Create a vSwitch2 with a port group of Production using vmnic1

```
esxcfg-vswitch -a vSwitch1
#esxcfg-vswitch -a vSwitch1 -m 9000
esxcfg-vswitch -A Production vSwitch1
esxcfg-vswitch -L vmnic1 vSwitch1
```

Example script for ESXi;

```
# accept the EULA, prevent it from stopping the install
vmaccepteula
```

```
# set the root password
```

```
rootpw mypassword
```

```
# use the first disk, always overwrite
```

```
autopart --firstdisk=local --overwritevmfs
```

```
# install from local media
```

```
install cdrom
```

```
# basic networking
```

```
network --bootproto=static --device=vmnic0 --ip=10.1.118.165 --
netmask=255.255.255.0 --gateway=192.168.118.1 --hostname=esx-01 --
vlanid=118
```

```
# reboot at the end
```

```
reboot
```

9.2.6 Perform a scripted ESX/ESXi host installation

So you've decided how to boot the installer and prepared a media repository (along with an installation script, ks.cfg for example). The next step is starting the installation itself. You boot from your boot source and at the prompt you can pass the following parameters to the installer.

NOTE: You must press TAB to edit the boot options (ESXi);

ip=<IP ADDRESS>	Sets the IP address for the duration of the install
netmask=<NETMASK>	Sets the netmask for the duration of the install
vlan=<vlan ID>	
gateway=<IP address>	Set the default gateway used during the install.
ks=<device>:<path>	Used to specify the location of the installation script file (CD, HTTP, NFS etc)
Ksdevice=<device>	The network device to use

NOTE: The bootloader and installation script parameters are NOT the same as Kickstart for Linux. Differences include using VMKNIC with the ksdevice parameter instead of ETH plus the partition command is totally different. See p45-48 of the [ESXi and vCenter installation guide](#) for details.

Example using HTTP (for ESX classic);

```
initrd=initrd.img mem=512m ksdevice=vmnic3 ip=192.168.1.123 netmask=255.255.255.0 gateway=192.168.1.1 ks=http://192.168.1.10/install/ks.cfg quiet
```

Example using NFS (for ESX classic);

```
initrd=initrd.img mem=512m ksdevice=vmnic1 ip=192.168.1.123 netmask=255.255.255.0 gateway=192.168.1.1 ks=nfs://192.168.1.10/nfs/install/ks.cfg quiet
```

Example using a local DVD (for ESX classic);

```
initrd=initrd.img mem=512m ksdevice=vmnic1 ip=192.168.1.123 netmask=255.255.255.0 gateway=192.168.1.1 ks=cdrom:/install/ks.cfg
```

Example using USB (for ESXi);

```
mboot.c32 vmkboot.gz ks=usb://esxi1.cfg --- vmkernel.gz --- sys.vgz --- cim.vgz --- ienviron.vgz --- install.vgz
```

Example using NFS (for ESXi);

```
mboot.c32 vmkboot.gz ks=nfs://192.168.0.10/nfs/install/config/esxi/esxi1.cfg --- vmkernel.gz --- sys.vgz --- cim.vgz --- ienviron.vgz --- install.vgz
```

These command lines are quite long and making errors is easy. You can create custom menu entries with these entries predefined – see this [blogpost by Mike LaSpina](#) or [get-admin's blog post](#) for details.

NOTE:

- ESXi only supports scripted installs from 4.1 onwards.
- Scripted Install is available only with the Installer version of ESXi and is not available in the Embedded version of ESXi
- You cannot use a scripted install to install ESXi Installable to a USB device

9.2.7 Pre/post script tasks

These are two sections in the installation scripts;

- The %pre section runs immediately after the kickstart options have been parsed, but before the operating system installation begins.
- The %post commands run after the installation, but before the system reboots. There can be multiple %post sections and they execute in the order they appear.

Common uses for these scripts include;

- The %pre section is often used to copy files to the host so that data is persistent across the reboots during install. It also allows you to take user input and apply it to the post build tasks.
- The %pre section is also used with 'lookup' scripts to automatically detect and configure the networking based on hostname or user input (for example)
- %post can be used to configure services such as NTP, ESX licencing and virtual networking (create vSwitches, portgroups etc) so a host is ready to join a cluster for example

A good example of using both %pre and %post can be found [here](#).

9.2.8 Troubleshooting scripted installs

This isn't explicitly listed in the blueprint but it's certainly worth knowing.

- The installer (Weasel) creates a logfile which can be found at /var/log/weasel.log
- You can use a second console *during* installation (press ALT+ F1) and then view the above logfile

Some common errors are covered in [VMware KB 1022308](#).

9.2.9 Further Reading

[Auto-deploy from the VMware labs](#) (plus a [great post from Simon Long](#))

[UDA/EDA](#) – deployment appliances to save you time

[Midwife script](#) – an ingenious way of building a base host and then applying post configuration using PowerCLI

[Example installation scripts covering advanced configuration](#)

[Using scripted install feature of ESXi](#)

9.3 Configure vCentre server Linked Mode

Knowledge

- Identify Linked Mode Prerequisites
- Identify differences between Linked and non-linked vCenter Server Configurations
- Identify when a role requires reconciliation

Skills and Abilities

- Reconcile Roles in a Linked Mode Configuration
- Create and Join a Linked Mode Group
- Determine use cases for vCenter Server Linked Mode
- Troubleshoot Linked Mode Configurations

Tools & learning resources

- Product Documentation
 - [ESX and vCenter Server Installation Guide](#)
 - vSphere Datacenter Administration Guide
- vSphere Client
- [vCentre Linked Mode](#) (VMworld '09) – well worth a watch!
- [Best practices for vCenter Linked Mode](#) (VIOPS)

9.3.1 Determine use cases

Simplify administration – in large environments this prevents administrators having to open multiple VI client sessions to multiple vCenters as all administration can be done through a single session.

Geographical or organisational boundary – if the infrastructure is split across a large geographical area latency could be an issue if the vCenter server is remote to some hosts. In this case locating a vCenter onsite and using Linked Mode may improve performance.

NOTE: Linked Mode can't be used to move VMs or ESX hosts between vCenter instances – it's view and search only.

Scalability is another reason to use Linked Mode. (NOTE: it's not a multiple of a single vCenter's maximums);

- 10 vCenter servers
- 1000 ESX hosts
- 10000 powered on VMs (15000 VMs total)

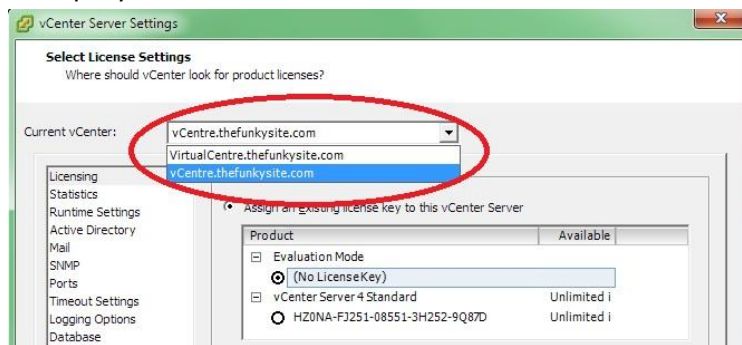
The above limits could be an issue for VDI deployments (lots of VMs per host).

NOT designed as a DR solution. VMware vCenter Server Heartbeat is the official DR product for vCenter, and SRM is the official DR product for protecting VMs.

Licences are shared between all linked vCenter instances so you can't stop someone at one site using any licence in the licencing portal.

9.3.2 Differences between Linked Mode and standalone vCenters

- VI client display – all vCenters shown in the tree hierarchy and at various other points. For example you can choose which vCenter to set Advanced Settings;



- Scalability limits
- Global role definitions vs per vCentre
- Global licencing vs per vCentre
- Ability to search across all vCenter instances
- ADAM service and replication

9.3.3 Linked Mode Prerequisites

Server compatibility is same as for vCenter.

AD considerations;

- Time synchronisation within 5 mins (Kerberos authentication).
- Working DNS
- User installing Linked Mode must be local Admin on both vCenters servers being linked.
- When vCenter servers are in multiple domains there must be two way trusts between domains.

Only included with Standard edition (not in vCenter Foundation edition).

People on the VMware communities site have had [no problems running vCenter 4.1 and linking to a vCenter running v4.0](#), though not sure if it's officially supported.

Linked Mode works with SRM and is compatible with vCSHB, but has limited compatibility with VMware Data Recovery. See the [Data Recovery FAQ](#) for details.

9.3.4 Create and Join a Linked Mode group

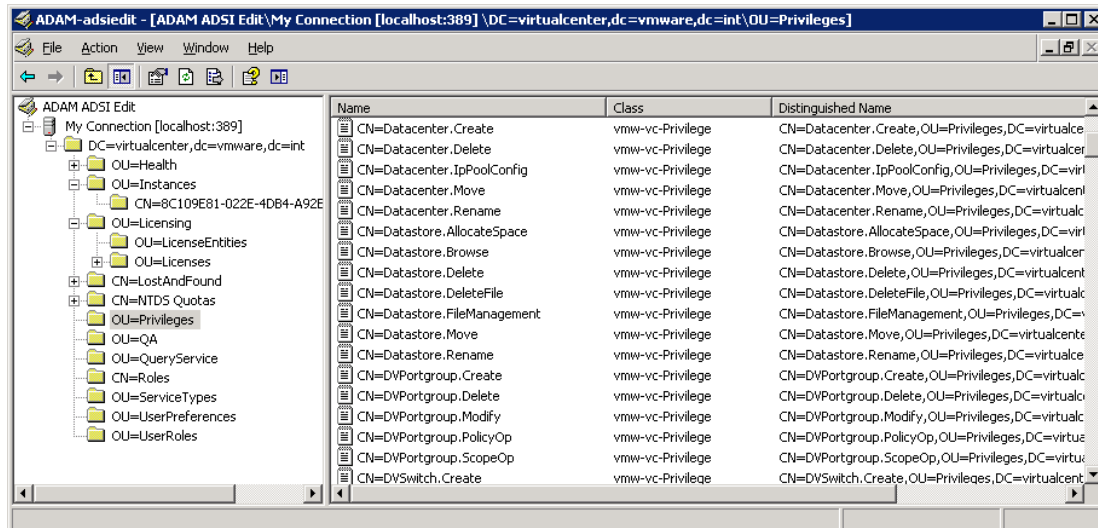
There are two ways to enable Linked Mode. You can choose it during vCenter installation (assuming you already have at least one vCenter server built) or you can configure it at a later date using Start -> Programs -> VMware -> vCenter Server Linked Mode Configuration.

NOTE: Setting up Linked Mode requires a restart of the vCenter services

When vCenter is installed, ADAM (Active Directory Application Mode, now renamed Lightweight Directory Service) is also installed regardless of whether Linked Mode is selected. This creates a lightweight LDAP server and an application specific 'partition' which stores configuration details for the vCenter instance. Details stored in the AD partition;

- Licence information
- Certificates (SSL etc)
- User roles and permissions

If Linked Mode is enabled the ADAM partition is replicated between all vCenter instances (see vCenter installation guide p.37). NOTE: Even if you're only using a single vCenter server NOT in linked mode the ADAM partition is still used to store licencing information (see [VMware KB1017480](#)). You can confirm this by starting 'ADAM ADSI Edit' on the vCenter server (Start -> Programs -> ADAM -> ADAM ADSI Edit);



9.3.5 Leaving a Linked Mode group (isolating a vCenter server)

The process is almost identical to joining a Linked Mode group. Go to Start -> Programs -> VMware -> vCenter Server Linked Mode Configuration but choose the 'Isolate this vCenter server...' option instead of the 'Join...' option. The vCenter Server restarts (the service, not the OS) and is no longer part of the Linked group.

9.3.6 Role Reconciliation

What is it? This is when roles defined at one vCenter server clash with the same role as defined on another vCenter and Linked Mode is used. For example if the Virtual Machine Administrator role is amended on one vCenter server (while in standalone mode) and it's then put into Linked Mode. As the same role has two conflicting definitions a conflict arises.

How to resolve it? The Linked Mode setup wizard will identify any conflicts and prompt the user. Choices are to automatically resolve the conflict or manually resolve it. In both cases the roles need to be renamed – if done automatically they're renamed with the name of the vCenter and role ie. 'vCentre01 VMAdmin'

9.3.7 Troubleshooting Linked Mode

Server name and DNS name for the vCenter server must match or connectivity errors will occur. See the [ESX and vCenter Server Installation Guide](#) (p106-107)

The following knowledgebase articles all use ADSI Edit to fix issues with vCenter Linked Mode;

- See [VMware KB1024036](#) for details of changing a host's name when in Linked Mode

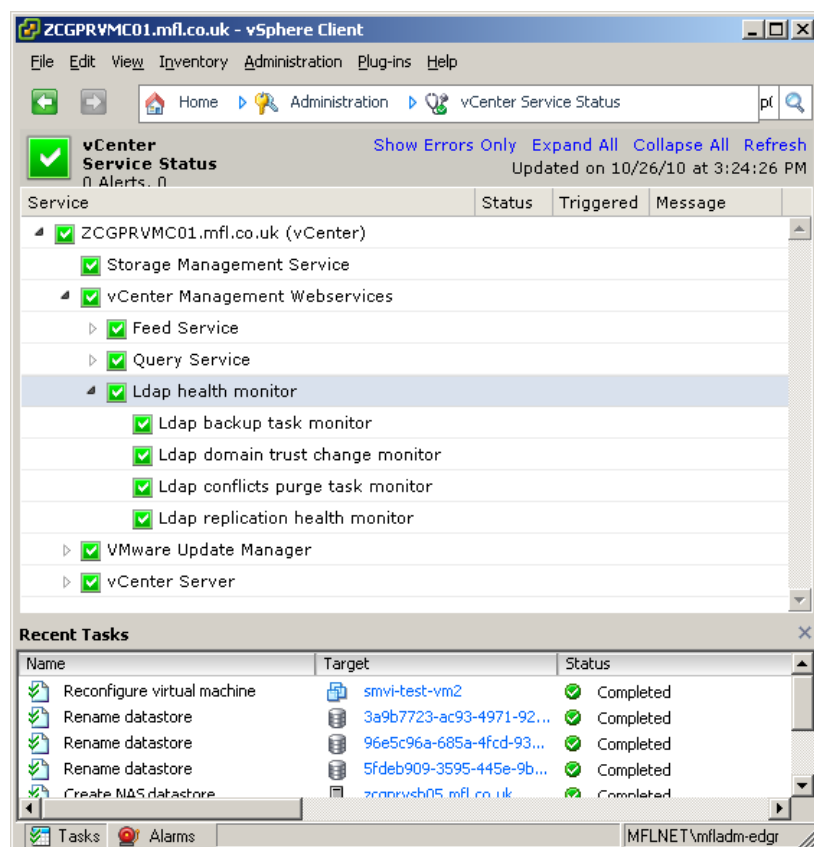
- See [VMware KB1017631](#) for details of how to force removal a vCenter server from Linked Mode
- See [VMware KB1024329](#) for details of how to rebuild the ADAM instance for a broken vCenter

As stated in the requirements section, time must be in sync across all vCenter servers (within 5 mins). If not (according to [VMware KB1009551](#)) there will be no obvious errors but replication will stop working.

If the user installing vCenter is not an administrator on both the source and destination vCenter servers the install may appear to complete OK but won't work as expected - see [VMware KB1016144](#) for details.

Check the ADAM service is started (named VMware vCMSDS) – this becomes a dependency for vCenter when Linked Mode is enabled. You can restart the VMware vCMSDS service without impacting any running operations (VM clones etc).

Replication is done via RPC so the relevant ports must be open on any firewalls. The default is for any changes to be replicated after 15 seconds although this is only for ADAM replicas in the same site. You can change replication schedules and monitor replication using the usual AD administration tools (more info can be found in this [Microsoft article](#)) or using vCenter (Home -> Administration -> vCenter Service Status, look at the LDAP Health Monitor). There is also a dedicated event log on the vCenter server named 'ADAM (VMwareVCMSDS)';



The actual files representing the ADAM partition are located here;

C:\Program Files\VMware\Infrastructure\VirtualCenter Server\VMwareVCMSDS

Logfiles

Logfile created when setting up Linked Mode;

C:\Program Files\VMware\Infrastructure\tomcat\temp\jointool.log

Logfiles for vCenter;

C:\Program Files\VMware\Infrastructure\logs

10 Appendix A - advanced parameters you might have to recall

.VMX settings

You may have to put these in a VMX file and looking them up in documentation could waste precious time;

- sched.mem.maxmemctl = max for balloon driver to reclaim (default 65%)
- sched.mem.pshare.enable = TRUE/FALSE to enable TPS per VM.
- monitor_control.disable_mmu_largepages = Disable large pages per VM
- monitor_control.restrict.backdoor = Enable for a virtual ESX host
- keyboard.typematicMinDelay = Used with WAN connected consoles (2000000 ms)

Advanced host settings

These can be configured through the VI client so you do get a checkbox but you have to remember where to look (as there are hundreds of options);

- boot.net.Netqueue = enable/disable Netqueue
- Mem.ShareScanGHZ = tune or disable TPS scan interval (0 to disable)
- Mem.allocGuestLargePages = enable/disable large pages (host level)
- nfs.max.volumes = Netapp advise set to 64. Default 8.

HA/DRS settings

- das.slotCPUinMHZ = set a custom slot size for CPU
- das.slotMeminMB = set a custom slot size for Memory
- das.failedetectioninterval = duration between HA heartbeats. In milliseconds.
- das.failedetectiontime = duration before isolation response (ms, default 30000)
- das.isolationaddress[x] = define another isolation network IP address
- das.usedefaultisolationaddress = TRUE/FALSE. Used with option above.
- das.maxvmrestartcount = number of retries during isolation response
- das.usevMotionnic = override not sharing vMotion and HA heartbeats
- das.ignoreRedundantNetWarning = suppress errors about mgmt network redundancy
- das.vmcputinmhz = change the default 256MHz CPU slot size
- das.AllowNetworks = specify a portgroup by name to use for HA heartbeats

vCenter advanced settings

- config.vpxd.filter.hostrescanFilter = enable/disable periodic host storage rescanning
- config.vpxd.filter.rdmFilter = enable/disable display of invalid/unavailable RDMS
- config.vpxd.filter.vmfsFilter = enable/disable VMFS integrity checking
- config.vpxd.filter.SameHostAndTransportsFilter = enable/disable