What's New in VMware® vSphere™ 4.1 — Availability and Resource Management

VMware vSphere 4.1

WHITE PAPER

vmware®

VMware® vSphere™ 4.1 ("vSphere") offers an extensive set of enhancements to increase the high availability, reliability and resource management of the vSphere platform. These include improvements to VMware High Availability (VMware HA), VMware Fault Tolerance (VMware FT), VMware Distributed Resource Scheduler (VMware DRS), VMware vMotion™, and VMware Enhanced vMotion Compatibility (EVC).

- Cluster scalability — the limits of cluster scalability have been improved.
- VMware FT and VMware HA enhancements — VMware FT no longer requires ESX hosts to have the same build numbers. VMware HA health check and operational status are visible to the user.
- VMware DRS VM–Host Affinity rules — sets constraints that restrict placement between a group of virtual machines and a group of hosts in a VMware DRS–enabled cluster.
- VMware Distributed Power Management (VMware DPM) improvements — schedules power management task for a cluster using the vSphere Client user interface.
- VMware DRS interoperability with VMware HA and VMware FT — VMware DRS is tightly integrated with VMware HA and VMware FT.
- Enhanced network logging performance — VMXNET3 is supported with VMware FT virtual machines.
- vMotion performance enhancements — more concurrent vMotion operations are supported.
- EVC improvements — usability of EVC has been improved.
- VMware Data Recovery (VDR) enhancements — concurrent VDR appliances are supported on vCenter Server.

Additional resource management enhancements in vSphere 4.1 include the following, which are fully described in other documentation.

- Networking: Network I/O Control (NetIOC) is a new traffic-management feature of the vNetwork Distributed Switch (vDS). NetIOC implements a software scheduler within the vDS to isolate and prioritize specific traffic types contending for bandwidth on the uplinks connecting ESX/ESXi 4.1 hosts with the physical network. Details on NetIOC can be found in the *What's New in VMware vSphere 4.1 — Networking*.
- Storage: Storage I/O Control (SIOC) is a new feature introduced in vSphere 4.1 to provide I/O prioritization of virtual machines running on a cluster of ESX hosts that access a shared storage pool. It extends the familiar constructs of shares and limits, which have existed for CPU and memory, to address storage utilization through a dynamic allocation of I/O queue slots across a cluster of ESX hosts. Details on SIOC can be found in the *What's New in VMware vSphere 4.1 — Storage*.

VMware Cluster Scalability Improvements

The scalability limits of VMware HA and VMware DRS clusters have been unified for vSphere 4.1. The maximum number of ESX hosts in a VMware HA or VMware DRS cluster is 32. The maximum number of virtual machines per ESX host in a VMware HA or VMware DRS cluster has been increased to 320, regardless of the number of ESX hosts in the cluster. The maximum number of virtual machines in a VMware HA or VMware DRS cluster has been increased to 3,000.

LIMITS	VSPHERE 4.0	VSPHERE 4.1	IMPROVEMENT
ESX hosts per cluster	32	32	1x
Virtual machines per ESX host in cluster	40[1]	320	8x
Virtual machines per cluster	1,280	3,000	>2x

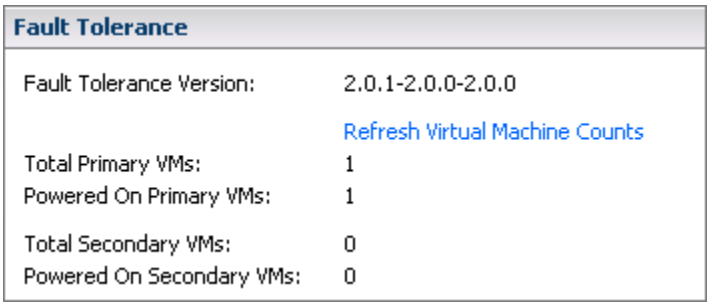
These are all soft limits.

NOTE: These VMware HA limits also apply to post-failover scenarios. Be sure that these limits will not be violated even after the maximum configured number of ESX host failovers. See the Configuration Maximums document for more details.

[1] For nine or more ESX hosts in HA cluster. 100 for eight or fewer ESX hosts in HA cluster, and 160 for eight or fewer ESX hosts in HA cluster with vSphere 4.0 Update 1.

VMware FT and VMware HA Enhancements

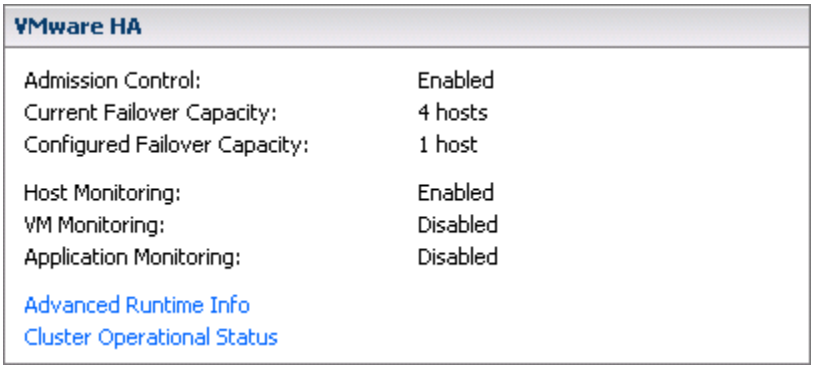
vSphere 4.1 introduces a VMware FT–specific versioning-control mechanism that allows the Primary and Secondary FT VMs to run on VMware FT–compatible hosts at different but compatible patch levels. No longer will the build numbers of the ESX hosts running the Primary and Secondary FT VMs have to be identical, as long as the Fault Tolerance Version numbers of the hosts match.



The Fault Tolerance Version appears on an ESX host's Summary tab in the vSphere Client. For ESX hosts prior to ESX/ESXi 4.1, this tab lists the ESX host build number instead.

vSphere 4.1 differentiates between events that are logged for a Primary FT VM and those that are logged for its Secondary FT VM by properly labeling them, and reports the reason why an ESX host might not support VMware FT. In addition, you can disable VMware HA when VMware FT–enabled virtual machines are deployed in a cluster, allowing for cluster maintenance operations without turning off VMware FT for every VMware FT–enabled virtual machine. See the *vSphere Availability Guide*.

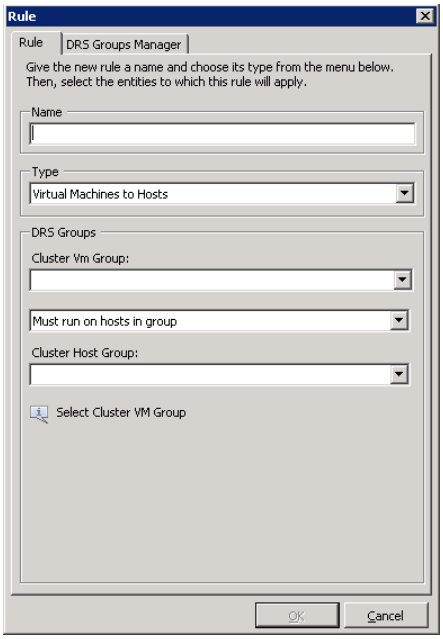
VMware HA provides an ongoing health-check facility to ensure that the required cluster configuration is met at all times. Deviations will result in an event or alarm on the cluster.



Configuration issues and other errors can occur for your VMware HA cluster or its ESX hosts, adversely affecting the proper operation of VMware HA. The VMware HA dashboard, which is accessible in the vSphere Client from the VMware HA section of the cluster's Summary tab, provides a new detailed window called Cluster Operational Status. Clicking on the Cluster Operational Status link (as shown in the figure) will open a window that displays more information about the current VMware HA operational status, including the specific status and errors for each host in the VMware HA cluster. See the *vSphere Availability Guide*.

VMware DRS VM–Host Affinity Rules

With vSphere 4.1, VMware DRS provides the ability to set constraints that restrict placement between a group of virtual machines and a group of hosts in a VMware DRS–enabled cluster. This feature is useful for enforcing host-based ISV licensing models, as well as for keeping sets of virtual machines on different racks or blade systems for availability reasons. See the *vSphere Resource Management Guide*.



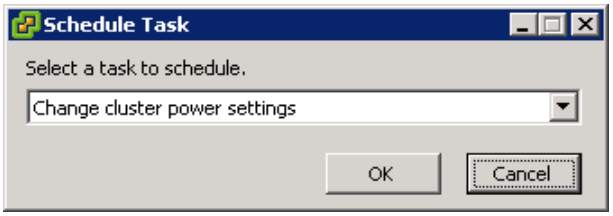
VMware advises customers that the VM–Host Affinity rules are not meant to be used often. The more constraints users put on virtual machine mobility, the harder it is for VMware DRS to balance load and to enforce resource allocation policies. There are two types of VM–Host Affinity rules: required rules (designated by "must" or "must not") and preferential rules (designated by "should" or "should not").

- Required VM–Host Affinity rules are only advised to be used for enforcing host-based licensing restrictions of ISV applications. VMware DRS, VMware HA, and VMware DPM never take any action that will violate required VM–Host Affinity rules. Manual violations are not permitted. Violations will only occur if DRS is prevented from migrating virtual machines to satisfy a required rule.
- Preferential VM–Host Affinity rules are meant for availability reasons, such as keeping two virtual machines on different racks or blade chassis's. Preferential rules can be violated to allow the proper functioning of VMware DRS, VMware HA and VMware DPM.

Finally, VM-VM anti-affinity rules can now incorporate more than two virtual machines. Here, VMware DRS tries to keep the specified virtual machines apart. Users could use such a rule to guarantee that certain virtual machines are always on different physical hosts.

VMware DPM Improvements

The vSphere Client now has a simple user interface allowing customers to create scheduled tasks enabling and disabling VMware DPM for a cluster by selecting "Change cluster power settings" in the Schedule Task wizard. For example, VMware DPM can be turned off prior to business hours in anticipation of higher resource demands, and turned on to conserve power after business hours.



Disabling VMware DPM will bring the ESX hosts in a VMware DRS cluster out of standby mode. This eliminates the risk of ESX hosts being stuck in standby mode while VMware DPM is disabled. Also, this ensures that when VMware DPM is disabled, all hosts are powered on and ready to accommodate load increases.

ESX 4.1 takes advantage of deep sleep states to further reduce power consumption during idle periods. In addition, users can view the history of host power consumption and power cap information on the vSphere Client Performance tab on newer platforms with integrated power meters. See the *vSphere Datacenter Administration Guide* and the *vSphere Resource Management Guide*.

VMware DRS Interoperability with VMware HA and VMware FT

VMware DRS will automatically place VMware FT–enabled virtual machines as they are powered on. In addition, setting DRS to Fully automated mode will allow VMware FT–enabled virtual machines to be automatically load balanced across the VMware DRS cluster.

NOTE: Enhanced vMotion Compatibility (EVC) must be enabled for better performance of VMware DRS to determine the location where VMware FT virtual machines will be automatically placed.

When a cluster has EVC enabled, VMware DRS makes the initial placement recommendations for VMware FT virtual machines, moves them during cluster load rebalancing, and allows users to assign a VMware DRS automation level to Primary FT VM (the Secondary FT VM always assumes the same setting as its associated Primary FT VM).

When VMware FT is used for virtual machines in a cluster that has EVC disabled, the VMware FT virtual machines are given VMware DRS automation levels of "disabled." In such a cluster, each Primary FT VM is powered on only on its registered host, its Secondary FT VM is automatically placed, and neither VMware FT virtual machine is moved for load-balancing purposes. For information on EVC, see the *VMware vSphere Datacenter Administration Guide*.

If you use affinity rules with a pair of VMware FT virtual machines, a VM-VM affinity rule applies to the Primary FT VM only, while a VM–Host Affinity rule applies to both the Primary FT VM and its Secondary FT VM.

In addition, VMware HA and VMware DRS are tightly integrated in vSphere 4.1, which allows VMware HA to restart virtual machines in more situations. Using VMware HA with VMware DRS combines automatic failover with load balancing. This combination can result in faster rebalancing of virtual machines after VMware HA has moved virtual machines to different hosts. However in some scenarios, VMware HA might not be able to failover virtual machines because of resource constraints for these reasons:

- VMware HA admission control is disabled and VMware DPM is enabled. This can result in VMware DPM consolidating virtual machines onto fewer hosts and placing the empty hosts in standby mode, leaving insufficient powered-on capacity to perform a failover.
- VM–Host Affinity (required) rules might limit the hosts on which certain virtual machines can be placed.
- There might be sufficient aggregate resources but these can be fragmented across multiple hosts so that they cannot be used by virtual machines for failover.

In such cases, VMware HA will use VMware DRS to try to adjust the cluster (for example, by bringing hosts out of standby mode or migrating virtual machines using vMotion to defragment the cluster resources) so that VMware HA can perform the failovers. See the *vSphere Availability Guide*.

Enhanced Network Logging Performance

VMware FT network logging performance allows improved throughput and reduced CPU usage. In addition, VMXNET3 vNICs are supported in VMware FT–enabled virtual machines. See the *vSphere Availability Guide*.

vMotion Performance Enhancements

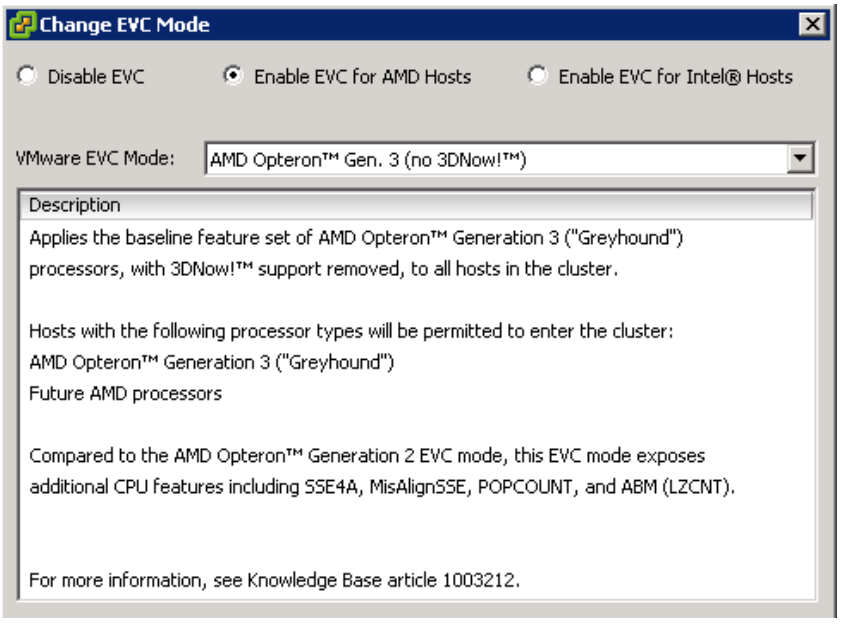
In vSphere 4.1, vMotion enhancements significantly reduce the overall time for host evacuations, with support for more simultaneous virtual machine migrations and faster individual virtual machine migrations. The result is a performance improvement of as much as 8x for an individual virtual machine migration, and support for four to eight simultaneous vMotion migrations per host, depending on the vMotion network adapter (1GbE or 10GbE, respectively). The number of simultaneous vMotion migrations touching a datastore is also increased to 128 for both VMFS and NFS datastores. Finally, when entering a host into Maintenance Mode, evacuation times are greatly decreased due to the above improvements. See the *vSphere Datacenter Administration Guide*.

OPERATION	VSPHERE 4.0	VSPHERE 4.1 (1GbE)	VSPHERE 4.1 (10GbE)
vMotions per ESX host	2	4	8
vMotions per datastore	8	128	128

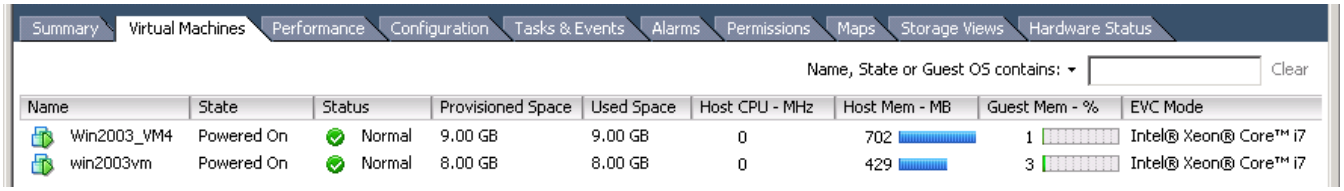
Improvements in Enhanced vMotion Compatibility

vSphere 4.1 includes usability improvements to EVC. These include the following:

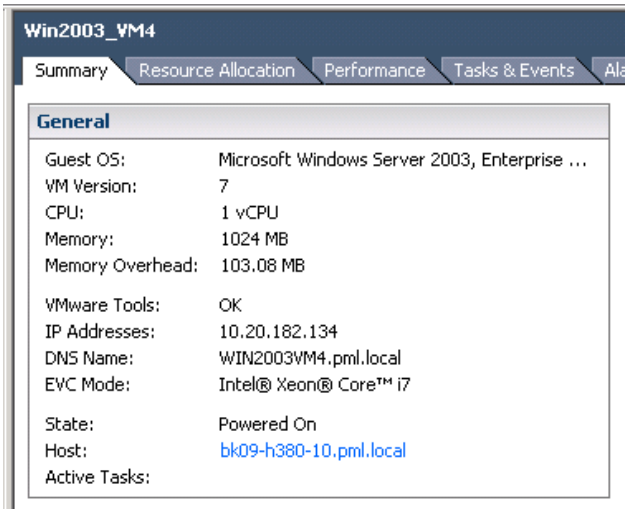
- Preparation for AMD next generation without 3DNow!™: Future generations of AMD processors may not include the 3DNow! set of CPU instructions. To prevent vMotion incompatibilities due to this change, vCenter Server 4.1 provides a new EVC mode called AMD Opteron™ Gen. 3 (no 3DNow!) that masks the 3DNow! instructions from virtual machines, allowing users to prepare clusters for the introduction of AMD hosts without 3DNow! instructions. Processors without the 3DNow! instructions will not be vMotion-compatible with older processors that have the 3DNow! instructions, unless an EVC mode or CPU compatibility mask is used to mask these instructions. Virtual machines in AMD EVC clusters should be transitioned to this new mode to preserve compatibility with future AMD processors.



- Better handling of powered-on virtual machines: vCenter Server now uses a running virtual machine's CPU feature set to determine if it can be migrated into an EVC cluster. Previously, it relied on the ESX host's CPU features. This will provide better granularity in error detection.
- Virtual machine's EVC capability: The Virtual Machines tab for hosts and clusters now displays the EVC mode corresponding to the features used by virtual machines. This makes it easier to determine which EVC modes the virtual machine can be migrated to. This also provides the opportunity to raise the EVC mode of a cluster and selectively power-cycle only those virtual machines that would benefit from it. You only need to power-cycle a virtual machine if you want it to pick up new features that have been exposed by an EVC mode change that happened while the virtual machine was active.



- Virtual machine Summary: The Summary tab for a virtual machine lists the EVC mode corresponding to the features used by the virtual machine. This information is available when the virtual machine is inside an EVC cluster.
- Earlier Add Host error detection: Host-specific incompatibilities are now displayed prior to the Add Host workflow when adding a host into an EVC cluster.



See the *vSphere Datacenter Administration Guide* for more information on improvements in Enhanced vMotion Compatibility.

VMware Data Recovery Enhancements

VMware Data Recovery (VDR) 1.2 is compatible with vSphere 4.1 and vSphere 4.0 Update 2. In addition to user-experience improvements, VDR 1.2 brings the ability to deploy up to 10 VDR appliances per vCenter Server instance. This allows multiple appliances to run concurrently to protect virtual machines in inventory and maximize the amount of data protected in the ever-shrinking backup window. The vSphere Client plug-in will allow quick switching among the deployed VDR appliances, and will automatically authenticate to each appliance. To complement the existing Windows File Level Restore (FLR) client, a Linux FLR client is also available with VDR 1.2. Finally, VDR 1.2 will leverage the additional Microsoft Volume Shadow Copy Service (VSS) quiescing support for Microsoft Windows Server 2008 and Windows 7 virtual machines being delivered in vSphere 4.1 to allow for application-level consistent snapshots. See the *VMware Data Recovery Administration Guide* for more details.

vStorage APIs for Data Protection (VADP) now offers VSS quiescing support for Windows Server 2008 and Windows Server 2008 R2 servers. This enables application-consistent backup and restore operations for Windows Server 2008 and Windows Server 2008 R2 applications.

Resources

Configuration Maximums
http://www.vmware.com/pdf/vsphere4/r41/vsp_41_config_max.pdf

vSphere Availability Guide
http://www.vmware.com/pdf/vsphere4/r41/vsp_41_availability.pdf

vSphere Resource Management Guide
http://www.vmware.com/pdf/vsphere4/r41/vsp_41_resource_mgmt.pdf

vSphere Datacenter Administration Guide
http://www.vmware.com/pdf/vsphere4/r41/vsp_41_dc_admin_guide.pdf

VMware Data Recovery Administration Guide
http://www.vmware.com/pdf/vdr_12_admin.pdf

What's New in VMware vSphere 4.1 — Networking
http://www.vmware.com/go/tp-new-in-networking-vsphere41

What's New in VMware vSphere 4.1 — Storage
http://www.vmware.com/go/tp-new-in-vsphere41-storage

VMware, Inc. 3401 Hillview Avenue Palo Alto CA 94304 USA Tel 877-486-9273 Fax 650-427-5001 www.vmware.com