



VMware Infrastructure

Resource Management with VMware DRS

VMware® Infrastructure 3 provides a set of distributed infrastructure services that make the entire IT environment more serviceable, available and efficient. Working with ESX Server 3, VirtualCenter 2, and VMotion™, VMware DRS (Distributed Resource Scheduler) dynamically allocates and balances computing capacity and virtual machine placement with resources pooled from multiple ESX Server hosts.

This white paper provides an architectural and conceptual overview of VMware DRS and describes how you can use DRS to simplify provisioning of applications, achieve higher levels of resource utilization, and better align use of IT resources with business values and priorities. It also describes how you can use new VMware Infrastructure capabilities such as resource pools and clusters, which allow you to allocate resources aggregated from multiple separate ESX Server hosts, but manage them with the simplicity of a single host.

This white paper covers the following topics:

- [Introduction to VMware DRS](#)
- [DRS Architecture and Conceptual Overview](#)
- [Using VMware DRS](#)
- [VMware DRS Requirements and Best Practices](#)

This white paper is intended for VMware partners, resellers, and VMware customers who want to implement virtual infrastructure solutions and want to know how to use new VMware Infrastructure features such as clusters, resource pools, and distributed infrastructure services such as VMware DRS.

Introduction to VMware DRS

With the introduction of ESX Server 3 and VirtualCenter 2, VMware extends the evolution of virtual infrastructure and virtual machines that began with the first VMware ESX Server release. ESX Server 3 and VirtualCenter 2 also introduce a revolutionary new set of infrastructure-wide services for resource optimization, high availability and data protection built on the VMware platform. These new services deliver capabilities which previously required complex or expensive solutions to implement using only physical machines. Use of these services also provides significantly higher hardware utilization and better alignment of IT resources with business goals and priorities. In the past, companies have had to assemble a patchwork of operating system or software application-specific solutions to obtain the same benefits.



VMware Infrastructure provides two new capabilities: **resource pools**, to simplify control over the resources of a host, and **clusters**, to aggregate and manage the combined resources of multiple hosts as a single collection. VMware Infrastructure virtualizes and aggregates industry-standard servers (processors, memory, their attached network and storage capacity) into logical resource pools (from a single ESX Server host or from a VMware cluster) that can be allocated to virtual machines on demand. Resource pools can be nested and organized hierarchically so that the IT environment matches company organization. In addition, individual business units can receive dedicated infrastructure while still profiting from the efficiency of resource pooling.

A set of virtualization-based distributed infrastructure services provide virtual machine monitoring and management to automate and simplify provisioning, optimize resource allocation, and provide high-availability service levels to applications at lower cost than static, physical infrastructure. One of the distributed services available, VMware DRS, dynamically allocates and balances computing capacity across the logical resource pools defined for VMware Infrastructure. VMware DRS continuously monitors utilization across the resource pools and intelligently allocates available resources among virtual machines based on resource allocation rules that reflect business needs and priorities.

Virtual machines operating within a resource pool are not tied to the particular physical server on which they are running at any given point in time. When a virtual machine experiences increased load, DRS first evaluates its priority against the established resource allocation rules and then, if justified, allocates additional resources by redistributing virtual machines among the physical servers. VMware VMotion executes the live migration of the virtual machine to a different server with complete transparency to end users. The dynamic resource allocation ensures that capacity is preferentially dedicated to the highest priority applications, while at the same time maximizing overall resource utilization.

DRS Architecture and Conceptual Overview

Before discussing the details of how VMware DRS works and how to use it to optimize resource allocation, it would be helpful to review a few basics about VMware infrastructure and some of the key elements with which VMware distributed services such as VMware DRS interact. The following sections provide some basic architectural information on VMware Infrastructure along with some of the newer features, such as clusters and resource pools, included with VMware Infrastructure 3 products.

VMware Infrastructure

At the core of VMware Infrastructure, VMware ESX Server is the foundation for delivering virtualization-based distributed services to IT environments. ESX Server provides a robust virtualization layer that abstracts processor, memory, storage and networking resources into multiple virtual machines that run side-by-side on the same physical server.

ESX Server installs directly on the server hardware, or "bare metal", and inserts a robust virtualization layer between the hardware and the operating system. ESX Server partitions a physical server into multiple secure and portable virtual machines that run on the same physical server. Each virtual machine represents a complete system—with processors, memory, networking, storage and BIOS—so that Windows, Linux, Solaris, and NetWare operating systems and software applications run in virtualized environment without any modification.

VirtualCenter, another key building block of VMware Infrastructure, manages all aspects of your virtual infrastructure—ESX Server hosts, virtual machines, provisioning, migration, resource allocations, and so on.



Figure 1 shows the architecture and typical configuration of VMware Infrastructure:

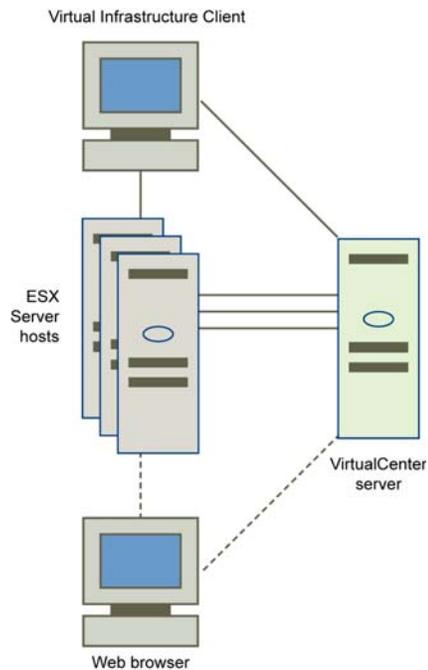


Figure 1. VMware Infrastructure Configuration

VMware Infrastructure simplifies management with a single client called the Virtual Infrastructure (VI) Client that you can use to perform all tasks. Every ESX Server configuration task from configuring storage and network connections, to managing the service console, can be accomplished centrally through the VI Client.

The VI Client connects to ESX Server hosts, even those not under VirtualCenter management, and also lets you remotely connect to any virtual machine for console access. There is a Windows version of the VI Client, and for access from any networked device, a web browser application provides virtual machine management and VMware Console access. The browser version of the client, Virtual Infrastructure Web Access, makes it as easy to give a user access to a virtual machine as sending a bookmark URL.

VirtualCenter user access controls provide customizable roles and permissions, so you create your own user roles by selecting from an extensive list of permissions to grant to each role. Responsibilities for specific VMware Infrastructure components such as resource pools can be delegated based on business organization, or ownership. VirtualCenter also provides full audit tracking to provide a detailed record of every action and operation performed on the virtual infrastructure.

Users can also access virtualization-based distributed services provided by VMotion, DRS, and HA directly through VirtualCenter and the VI client.



In addition, VirtualCenter exposes a rich programmatic Web Service interface for integration with third party system management products and extension of the core functionality:

- **VMware VMotion** enables the live migration of running virtual machines from one physical server to another. Live migration of virtual machines enables companies to perform hardware maintenance without scheduling downtime and disrupting business operations. VMotion allows the mapping of virtual machines to hosts to be continuously and automatically optimized within clusters for maximum hardware utilization, flexibility, and availability.
- **VMware DRS** works with VMotion to provide automated resource optimization and virtual machine placement and migration, to help align available resources with pre-defined business priorities while maximizing hardware utilization.
- **VMware HA** enables broad-based, cost-effective application availability, independent of hardware and operating systems.
- **VMware Consolidated Backup** provides an easy to use, centralized facility for LAN-free backup of virtual machines. Full and incremental file-based backup is supported for virtual machines running Microsoft Windows operating systems. Full image backup for disaster recovery scenarios is available for all virtual machines regardless of guest operating system.

VMware Clusters

Clusters are a new concept in virtual infrastructure management. They give you the power of multiple hosts with the simplicity of managing a single entity. A cluster is a group of loosely connected computers that work together, so that, from the point of view of aggregating resources such as CPU processing capability and memory, they can be viewed as though they are a single computer.

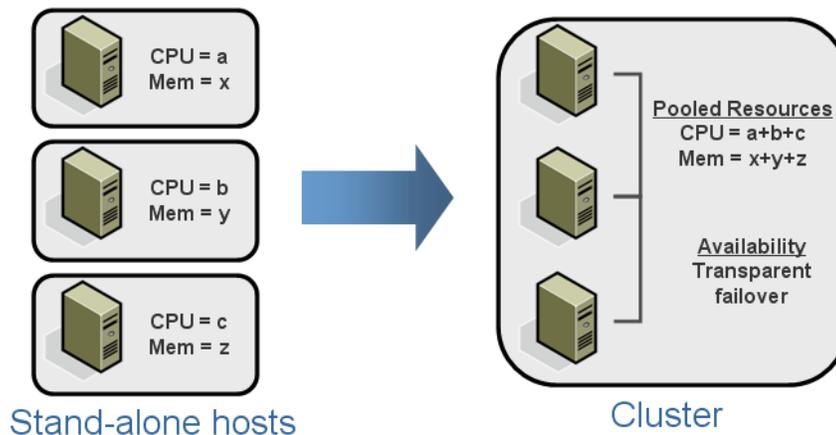


Figure 2. Resource Aggregation in VMware Clusters

VMware clusters let you aggregate the various hardware resources of individual ESX Server hosts but manage the resources as if they resided on a single host. When you power on a virtual machine, it can be given resources from anywhere in the cluster, rather than be tied to a specific ESX Server host.



VMware Infrastructure 3 provides two services to help with the management of VMware clusters: VMware HA (high availability) and VMware DRS. VMware HA allows virtual machines running on specific hosts to be switched over to use other host resources in the cluster in the case of host machine failures. VMware DRS provides automatic initial virtual machine placement and makes automatic resource relocation and optimization decisions as hosts are added or removed from the cluster or the load on individual virtual machines changes.

Resource pools can also be defined using the resources of a cluster. Defining resource pools through clusters complements efforts to align allocation of IT resources with company organization, business goals, and priorities.

Distributed Resource Scheduling

VMware DRS and VirtualCenter provide a view and management of all resources in the cluster. A global scheduler within VirtualCenter enables resource allocation and monitoring for all virtual machines running on ESX Servers that are part of the cluster.

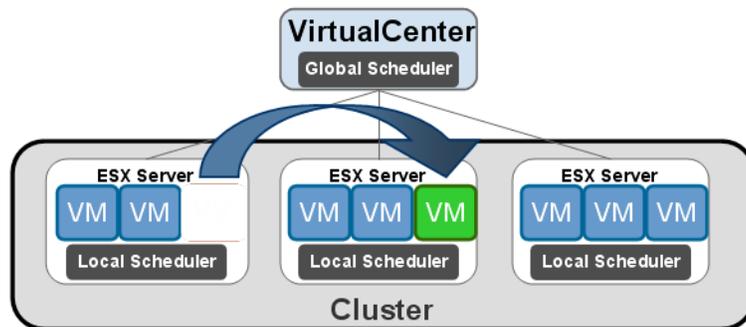


Figure 3. DRS Global Scheduler of Clusters in VirtualCenter

DRS provides automatic initial virtual machine placement on any of the hosts in the cluster, and also makes automatic resource relocation and optimization decisions as hosts or virtual machines are added or removed from the cluster. DRS can also be configured for manual control, in which case it only makes recommendations that you can review and carry out.

DRS provides several additional benefits to IT operations:

- Day-to-day IT operations are simplified as staff members are less affected by localized events and dynamic changes in their environment. Loads on individual virtual machines invariably change, but automatic resource optimization and relocation of virtual machines reduces the need for administrators to respond, allowing them to focus on the broader, higher-level tasks of managing their infrastructure.
- DRS simplifies the job of handling new applications and adding new virtual machines. Starting up new virtual machines to run new applications becomes more of a task of high-level resource planning and determining overall resource requirements, than needing to reconfigure and adjust virtual machines settings on individual ESX Server machines.
- DRS simplifies the task of extracting or removing hardware when it is no longer needed, or replacing older host machines with newer and larger capacity hardware. To remove hosts from a cluster, you can simply place them in maintenance mode, so that all virtual machines currently running on those hosts get reallocated to other resources of the cluster. After monitoring the performance of remaining systems to ensure that adequate resources remain for currently running virtual machines, you can remove the hosts from



the cluster to allocate them to a different cluster, or remove them from the network if the hardware resources are no longer needed. Adding new resources to the cluster is also straightforward, as you can simply drag and drop new ESX Server hosts into a cluster.

Before drilling down deeper into the details of how DRS works and how you can use it, it's helpful to have a better understanding of resource pools and how they use clusters.

What are Resource Pools?

Resource pools can be used to hierarchically partition available CPU and memory resources.

- Each standalone host and each DRS cluster has an (invisible) root resource pool that groups the resources of that host or cluster. The root resource pool is not displayed because the resources of the host (or cluster) and the root resource pool are always the same. If you don't create child resource pools, there are only the root resource pools.
- Users can create child resource pools of the root resource pool or of any user-created child resource pool. Each child resource pool owns some of the parent's resources and can, in turn, have a hierarchy of child resource pools to represent successively smaller units of computational capability.

A resource pool can contain child resource pools, virtual machines, or both. This creates a hierarchy of shared resources. The resource pools at a higher level are called parent resource pools, while the resource pools and virtual machines that are at the same level are called siblings. The following illustration shows an example:

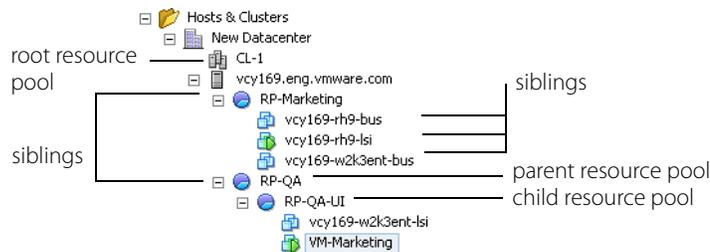


Figure 4. Resource Pool Nesting

In the example above, RP-QA is the parent resource pool for RP-QA-UI. RP-Marketing, and RP-QA are siblings. The three virtual machines immediately below RP-Marketing are also siblings.

Why Use Resource Pools?

Resource pools provide more flexibility, better hardware utilization, and more fine-grain delegation and control over resources of a host, but the benefits are especially evident when resource pools are used to compartmentalize all resources in a cluster. You can create multiple resource pools as direct children of the host or cluster and configure them, then delegate control over them to other business unit or operation managers. Using resource pools can provide the following benefits:

- Flexible hierarchical organization – You can add, remove, or reorganize resource pools or change resource allocations as needed.
- Isolation between pools, sharing within pools – Top-level administrators can make a pool of resources available to an administrator for a department. Allocation changes that are internal to one departmental resource pool do not unfairly affect other unrelated resource pools.



- Access control and delegation – When a top-level administrator makes a resource pool available to a department-level administrator, all virtual machine creation and management can then be performed by that administrator within the boundaries of the resources to which the resource pool is entitled by the current Shares, Reservation, and Limit settings.
- Separation of resources from hardware – if you are using clusters, the resources of all hosts are always assigned to the cluster. That means you can perform resource management independently of the actual host machines that contribute the resources. If you replace four 4-way hosts with four 8-way hosts, you don't need to make changes to your resource allocations. This allows administrators to think more about aggregate computing capacity and less about individual hosts.
- Management of sets of virtual machines running a distributed application – You don't need to set resources on each virtual machine. Instead, you can control the aggregate allocation of resources to the set of virtual machines by changing settings on their enclosing resource pool.

Resource Pools using Clusters

Resource pools, when used with clusters, provide easier management of large scale virtual system deployments, and enable allocation and optimization of resources aligned with an organization's business goals and priorities. The following illustration shows an example of resource pools drawing resources from a cluster:

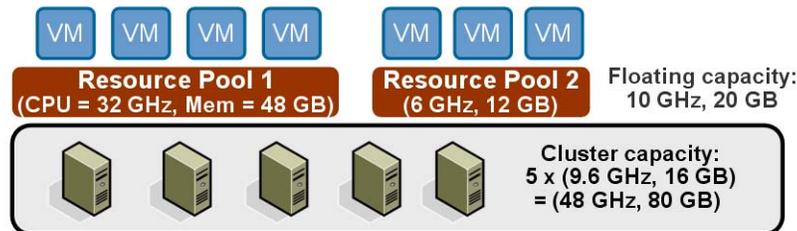


Figure 5. Resource Pools in a Cluster

In this case, five different hosts contribute resources to the cluster, with each host providing 9.6GHz processor bandwidth and 16GB memory. The total capacity of the cluster is 48GHz processor bandwidth and 80GB memory. The example shows two different resource pools. Resource Pool 1 has allocated resources of a 32GHz processor and 48GB memory. Resource Pool 2 has allocated resources of a 6GHz processor and 12GB memory.

Using DRS and VirtualCenter, virtual machines from various business units or applications can be assigned to the different resource pools, without regard for the physical hosts on which virtual machines will actually be run. In addition, the cluster has extra "floating" capacity, so additional resources are available and can be dynamically added to support virtual machines running in the resource pools, based on virtual machine load or other factors. So, a virtual machine's peak resource usage can exceed its reserved capacity by "borrowing" from extra floating cluster capacity or other "reserved" resources, if those resources are not currently being used.



Using VMware DRS

This section describes some of the setup and operation tasks you can perform using DRS and VirtualCenter: adding and removing hosts from clusters, setting up and allocating resource pools to virtual machines, and delegating resource administration to pool administrators. In addition, it describes configuration and operational controls in VirtualCenter to set VMotion automated migration levels, monitor CPU and memory usage, and perform runtime load balancing and adjustment.

Enabling DRS

VMware DRS is included as an integrated component in VMware Infrastructure 3 Enterprise. It is also available as add-on license options to VMware Infrastructure 3 Starter and VMware Infrastructure 3 Standard. To use DRS when you create VMware clusters, you need to set the Enable VMware DRS option, so that DRS can use the cluster load distribution information for initial virtual machine placement, to make load balancing recommendations, and to perform automatic runtime virtual machine migration.

For clusters enabled for DRS, the resources of all included hosts are assigned to the cluster. DRS uses the per-host resource pool hierarchies to implement the cluster-wide resource pool hierarchy. When you view the cluster using a VI Client connected to a VirtualCenter Server, you see the resource pool hierarchy implemented by DRS.

Note: When you view individual hosts using a VI Client connected to an ESX Server host, the underlying hierarchy of resource pools is presented. However, because DRS implements the most balanced resource pool hierarchy it can, you should not modify the hierarchy visible on the individual ESX Server host. If you do, DRS will undo your changes immediately for better balance if you are in automatic mode. DRS will make migration recommendations if you are in partially automatic or manual mode.

Creating a VMware Cluster

A cluster is a collection of ESX Server hosts and associated virtual machines with shared resources and a shared management interface. When you add a host to a cluster, the host's resources become part of the cluster's resources immediately, without any user intervention. When you create a cluster, you can choose to enable it for DRS, HA, or both.

Note: You can create a cluster without a special license, but you must have a license to enable a cluster for DRS or HA. In VirtualCenter 2.0.1, the maximum number of hosts that can be included in a cluster is 16.

The system must meet certain prerequisites to use VMware cluster features successfully:

- In general, DRS will work best if the virtual machines meet VMotion requirements. (See [VMware DRS Requirements and Best Practices](#), later in this white paper, for more information.)
- If you want to use DRS for load balancing, the hosts in your cluster must be part of a VMotion network. If the hosts are not in the VMotion network, DRS can still make initial placement recommendations.

VirtualCenter provides a New Cluster wizard to take you through the steps of creating a new cluster. When you first invoke the wizard, you are prompted to choose whether to create a cluster that supports VMware DRS, VMware HA, or both. That choice affects subsequent wizard pages. When you create a cluster, it initially does not include any hosts or virtual machines.



Selecting Automation Level

If you have chosen the Enable VMware DRS option, the wizard's DRS panel allows you to select the level of automation you want DRS to use regarding automatic or manual migration of virtual machines.

Setting	Initial VM Placement	Load Balancing
Manual	Recommended host is displayed.	Migration recommendation is displayed.
Partially Automatic	Automatic placement.	Migration recommendation is displayed.
Automatic	Automatic placement.	Automatic execution of migration recommendations.

Note: You can later change the level of automation for the whole cluster or for individual virtual machines. To do that, select the cluster in VirtualCenter and choose **Edit Settings** from the right-click mouse button menu. Choose the **VMware DRS** option to change settings for the entire cluster. To change settings for the individual virtual machine in the cluster, choose **VMware DRS -> Virtual Machine Options** and then right-click on a displayed virtual machine. Using the virtual machine options, you can, for example, put an individual virtual machine in manual or partially automatic mode, while the rest of the cluster is in automatic mode. It is also possible to disable DRS entirely for individual virtual machines.

After you've completed all selections for the cluster, the wizard presents a summary of the options selected. You can then complete the cluster creation.

Adding Hosts to a Cluster

The VirtualCenter inventory panel displays all clusters and hosts managed by that VirtualCenter Server. You can add a managed host to a cluster by selecting the host from either the inventory or list view and dragging the host to the target cluster object.

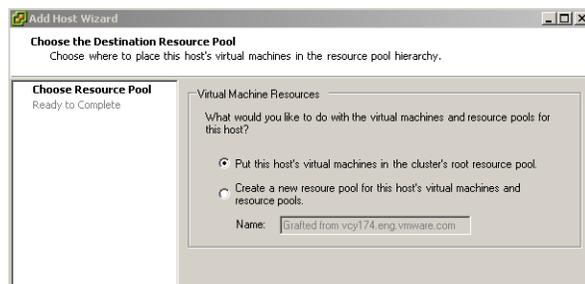


Figure 6. Adding Virtual Machines to Resource Pools

The wizard prompts to determine what you want to do with the host's virtual machines and resource pools:

- If you choose to put the host's virtual machines in the cluster's root resource, VirtualCenter makes any direct children of the host (virtual machines or resource pools) direct children of the cluster and discards the hierarchy. Any existing resource pools are removed.
- If you choose to create a new resource pool for the host's virtual machines and resource pools, VirtualCenter creates a top-level resource pool that becomes a direct child of the cluster and adds all children of the host to that new resource pool. You can supply a name for that new top-level resource pool. (The default is "grafted from *host_name*.") In this case,



the cluster keeps the grafted resource pool hierarchy if you later remove the host from the cluster. The host loses the existing resource pool hierarchy.

Note: If the host has no child resource pools or virtual machines, the host's resources are added to the cluster but no resource pool hierarchy with a top-level resource pool is created.

Adding Virtual Machines to Resource Pools

When you create a new virtual machine, the Virtual Machine Wizard in VirtualCenter allows you to select a resource pool in which to add the new virtual machine. You can also add an existing virtual machine to a resource pool by dragging or moving the virtual machine from any location in the inventory to the desired resource pool.

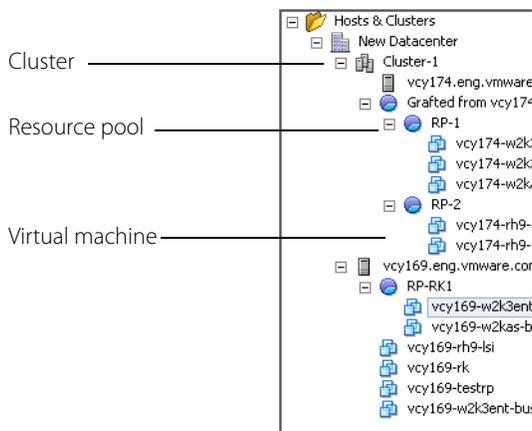


Figure 7. Cluster and Resource Pool Hierarchies

When you move a virtual machine to a new resource pool:

- The virtual machine's reservation and limit do not change.
- If the virtual machine's shares are high, medium, or low, %Shares adjusts to reflect the total number of shares in use in the new resource pool.
- If the virtual machine has custom shares assigned, the share value is maintained.

Note: If the virtual machine would receive a very large percentage of total shares, a warning is displayed.

- The information displayed in the Resource Allocation tab about the resource pool's reserved and unreserved CPU and memory resources changes to reflect the reservations associated with the virtual machine.

Note: CPU and memory, reserved and unreserved, change only if the virtual machine is powered on. If the virtual machine has been powered off or suspended, it can be moved, but overall available resources for the resource pool are not affected.

If a virtual machine is powered on, and the destination resource pool does not have enough CPU or memory to guarantee the virtual machine's reservation, the move fails because admission control does not allow it. An error dialog box compares available and requested resources, so you can consider whether an adjustment would resolve the issue.



Virtual Machine Attributes: Shares, Reservation, and Limit

For each virtual machine, you can specify Shares, Reservation (minimum) and Limit (maximum). There are other white papers posted on the VMware web site that provide more information and details of how to use these settings in different scenarios, but the following table provides a basic description of the effects these settings have on resource pools, particularly those defined using DRS-enabled clusters.

Setting	Description
Shares	The relative priority or importance of a virtual machine. If a virtual machine has twice as many shares for a resource as another virtual machine, it is entitled to consume twice as much of that resource. Shares are typically specified as high, normal, or low, using a 4:2:1 ratio. You can also choose Custom to assign a specific number of shares expressing a proportional weight to each virtual machine. Specifying shares makes sense only with regard to sibling virtual machines or resource pools, that is, virtual machines or resource pools with the same parent in the resource pool hierarchy. Siblings share resources according to their relative share values, bounded by the reservation and limit. When you assign shares to a virtual machine, you always specify the relative priority for that virtual machine.
Reservation	The guaranteed reservation for the virtual machine. When you attempt to power on a virtual machine, the virtual machine is only allowed to power on if the CPU and memory reservation is available and that amount of CPU and memory can be guaranteed even under heavy loads. Reservation defaults to 0. It is often a good idea to specify a reservation to guarantee that the necessary CPU or memory are always available for the virtual machine.
Limit	The upper limit for CPU or memory for a virtual machine. That is, more resources can be allocated to a virtual machine than the reservation amount, however, the allocated resources will never exceed the specified limit . The limit is expressed in concrete units (MHz or MB). CPU and memory limit default to unlimited. In that case, the amount of memory you assigned to the virtual machine when you created it becomes the virtual machine's memory. In most cases, it's not necessary to specify a limit. There are pros and cons for specifying a limit. On the plus side, assigning a limit can be useful if you start with a small number of virtual machines and want to manage user expectations; performance will deteriorate somewhat as you add more virtual machines, so you can simulate having fewer resources initially available by specifying a limit. The down side is that you might waste idle resources if you specify a limit, since virtual machines are not allowed to use more resources than the limit, even when the system is under-utilized and idle resources are available. You should therefore specify a limit only if there are good reasons for doing so.

Note: These settings are also used for resource pools.



Creating and Managing Resource Pools for DRS-Enabled Clusters

For clusters enabled for DRS, you can also create child resource pools of the cluster. When you create a child resource pool, you are prompted for resource pool attribute information. The system uses admission control to make sure you don't allocate resources that aren't available. After selecting the intended parent and choosing the New Resource Pool option, you can specify information listed in the following table for a new resource pool:

Field	Description
Name	Name of the new resource pool.
CPU Resources:	
Shares	Number of CPU shares the resource pool has with respect to the parent's total. Sibling resource pools share resources according to their relative share values bounded by the reservation and limit. You can choose Low, Normal, or High, or choose Custom to specify a number that assigns a share value.
Reservation	Guaranteed CPU allocation for this resource pool.
Expandable Reservation	Use this check box to indicate that, if virtual machines are powered on in this resource pool, and the reservations of the virtual machines combined are larger than the reservation of the resource pool, the resource pool can use a parent's or ancestor's resources. Default is checked.
Limit	Upper limit for the amount of CPU the host makes available to this resource pool. Default is Unlimited. To specify a limit, uncheck the Unlimited check box and type in the number.
Memory Resources:	
Shares	Number of memory shares the resource pool has with respect to the parent's total. Sibling resource pools share resources according to their relative share values bounded by the reservation and limit. You can choose Low, Normal, or High, or choose Custom to specify a number that assigns a share value.
Reservation	Guaranteed memory allocation for this resource pool.
Expandable Reservation	Use this check box to indicate that more than the specified reservation should be allocated if resources are available in a parent.
Limit	Upper limit for this resource pool's memory allocation. Default is Unlimited. To specify a different limit, uncheck the Unlimited check box.

When you power on a virtual machine or create a resource pool, the system checks whether the CPU and memory reservation is available for that action. If **Expandable Reservation** is not checked, the system considers only the resources available in the selected resource pool.

If **Expandable Reservation** is checked (the default), the system considers both the local resource pool and those of ancestors when performing admission control (deciding whether to allow a virtual machine to be powered on). Ancestors includes direct parents, parents of the parents, and so on. For ancestors, resources can only be used if the ancestor pool is set to **expandable** and no Limit is set that would stop it from borrowing more resources. Leaving the expandable option checked offers more flexibility, but, at the same time provides less protection from a child resource pool owner potentially reserving more resources than anticipated.

If any of the selected values are not legal values, because of limitations on total available CPU and memory, a yellow triangle is displayed to indicate so, and you cannot enter values larger than the legal values. For example, if you have a resource pool with a reservation of 10GB, and you created a child resource pool with a reservation of 6GB, you cannot create a second child resource pool with a reservation of 6MB and Type set to Fixed.



Admission Control

When you power on virtual machines inside a resource pool, or attempt to create a child resource pool, the system checks to ensure the resource pool's restriction and limit settings are not violated.

Before you power on a virtual machine or create a resource pool, you can check the CPU Unreserved and Memory Unreserved fields in the resource pool's Resource Allocation tab to see whether sufficient resources are available for your needs.

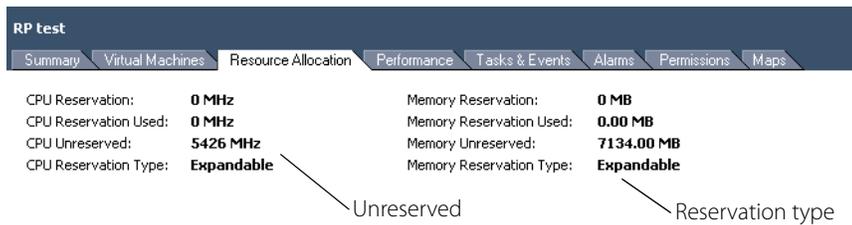


Figure 8. Resource Allocation Information

How unreserved CPU and memory are computed depends on the Reservation Type:

- If the Reservation Type is Fixed, the system checks whether the resource pool has sufficient unreserved resources. If it does, the action can be performed. If it does not, a message appears and the action cannot be performed.
- If the Reservation Type is Expandable, the system first checks whether the resource pool has sufficient resources to fulfill the requirements. If there are sufficient resources, the action is performed.
- If there are not sufficient resources, the managing server checks whether resources are available in a parent resource pool (direct parent or ancestor). If there are, the action is performed and the parent resource pool resources are reserved. If no resources are available, a message appears and the action is not performed.

When you power on a virtual machine on a host that is part of a DRS-enabled cluster, VirtualCenter does the following:

1. VirtualCenter performs admission control (that is, it checks whether the cluster or resource pool has enough resources for the virtual machine to be powered on).
2. VirtualCenter checks whether there's a host in the cluster that has enough resources for powering on the virtual machine. This must be a single host. It's not enough that two hosts jointly have sufficient resources. If there are not sufficient resources, a message appears. Otherwise, VirtualCenter proceeds as follows:
 - If DRS is in Manual mode, VirtualCenter displays a list of recommended hosts ordered from best to worst. You can choose one of the hosts.
 - If DRS is in Partially Automatic or Automatic mode, VirtualCenter places the virtual machine on the most suitable host.

Admission control is performed whenever you power on a virtual machine or create a resource pool. It is also performed whenever you modify the Reservation or Limit settings or when you change the structure of the resource pool hierarchy.



Viewing Resource Pool Information

When you select a resource pool in the VirtualCenter display, the Summary tab displays high-level statistical information about the resource pool.

Figure 9. Viewing Resource Pool Information

The following table describes items included in the summary:

Section	Description
General	The General panel displays statistical information for the resource pool. <ul style="list-style-type: none"> • Number of Virtual Machines — Number of virtual machines in this resource pool. Does not include the number of virtual machines in child resource pools. • Number of Running Virtual Machines — Number of running virtual machines in this resource pool. Does not include the number of virtual machines running in child resource pools. • Number of Child Resource Pools — Number of direct child resource pools. Does not include all resource pools in the hierarchy but only direct children.
CPU	Displays the CPU Shares, Reservation, Reservation Type, and Limit that were specified for this resource pool. Also displays the amount of CPU currently unreserved.
Commands	Allows you to invoke commonly used commands. <ul style="list-style-type: none"> • New Virtual Machine — Starts the New Virtual Machine wizard to create a new virtual machine in this resource pool. • Create Resource Pool — Displays the Create Resource Pool dialog, which allows you to create a child resource pool of the selected resource pool. • Edit Settings — Allows you to change the CPU and memory attributes for the selected resource pool.
Resources	Displays CPU Usage and Memory Usage for the virtual machines within the selected resource pool.
Memory	Displays the Memory Shares, Reservation, Reservation Type, and Limit that were specified for this resource pool. Also displays the amount of memory currently unreserved.



The resource pool's Resource Allocation tab gives detailed information about the resources currently reserved and available for the resource pool, and lists the user of the resources as described in the tables below:

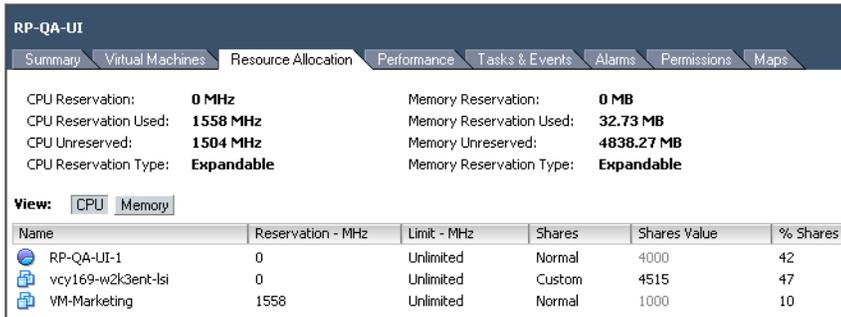


Figure 10. Resource Pool Allocation Information

The top portion of the display specifies information about the resource pool itself:

Field	Description
CPU Reservation/ Memory Reservation	Amount of CPU or memory specified in the reservation for this resource pool. The reservation can be specified during resource pool creation, or later by editing the resource pool.
CPU Reservation Used/ Memory Reservation Used	CPU or memory reservation used. Reservations are used by running virtual machines or by child resource pools with reservations.
CPU Unreserved / Memory Unreserved	CPU or memory currently unreserved and available to be reserved by virtual machines and resource pools. Look at this number when trying to determine whether you can create a child resource pool of a certain size, or whether you can power on a virtual machine with a certain reservation.
Reservation Type	Expandable or Fixed.

Below the resource pool specific information is a list of the resource pool's virtual machines and child resource pools of this resource pool.

Note: This table does not list virtual machines assigned to child resource pools of this resource pool.

You can click the CPU or Memory tab to display the following information:

Field	Description
Name	Name of the resource pool or virtual machine.
Reservation	Specified reservation for this virtual machine or resource pool. Default is 0, that is, the system reserves no resources for this resource pool.
Limit	Specified limit for this virtual machine or resource pool. Default is Unlimited, that is, the system allocates as many resources to this virtual machine as it can.
Shares	Specified shares for this virtual machine or resource pool. Choose High, Normal, or Low, if one of the default settings has been selected, or Custom, if a custom setting has been selected.
Shares Value	Number of shares allocated to this virtual machine or resource pool. This number depends on the Shares setting (High, Medium, Low, or Custom).



Field	Description
%Shares	Shares value for this resource pool or virtual machine divided by the total number of Shares in parent resource pool.
Type	Reservation type. Either Fixed or Expandable.

Monitoring Resource Pool Performance

Monitoring a resource pool's performance is useful if you want to understand the effectiveness of resource pool allocations. To monitor a resource pool's performance, select the resource pool in the inventory panel and click the Performance tab. VirtualCenter displays information about resource pool performance. You can click Change Chart Options to customize the performance chart.

Viewing Cluster Information

When you select a cluster from the VirtualCenter inventory panel, the Summary page displays high-level information about the selected cluster.

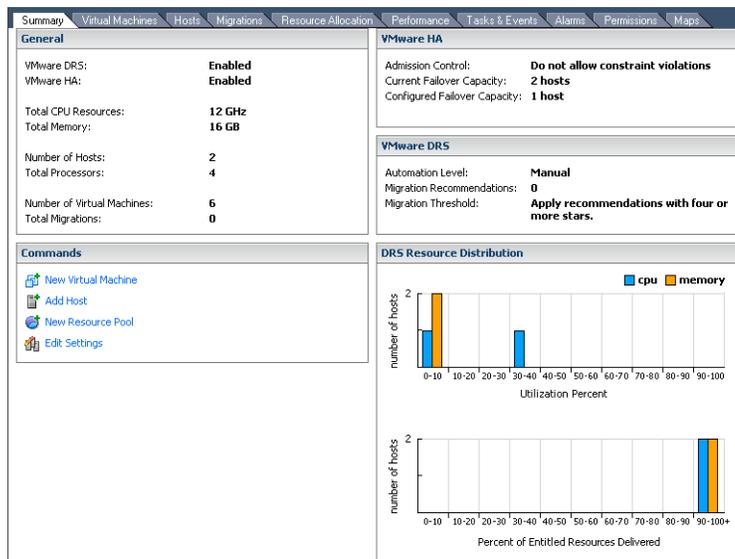


Figure 11. Viewing Cluster Information

The Summary page provides information about the cluster and the virtual machines assigned to the cluster. For clusters enabled for DRS, the Summary page displays automation level and migration threshold settings, outstanding migration recommendations and real-time histograms of "Utilization Percent" and "Percent of Entitled Resources Delivered" to show how balanced the cluster is. Migration recommendations are displayed if you select the Migrations tab. (Automation level and migration threshold are set during cluster creation.)



Cluster Status Information

As hosts and virtual machines are added or removed, clusters can become overcommitted or invalid because of HA or DRS violations. Messages displayed on the Summary page displays the status of the currently selected cluster.

The Virtual Infrastructure Client indicates whether a cluster is valid (green), overcommitted (yellow), or invalid (red).

Green (Valid) Cluster

A cluster is considered valid unless something happens that makes it overcommitted or it no longer satisfies failover capacity requirements. For example, a DRS cluster can become overcommitted if a host fails. A DRS cluster can also become invalid if VirtualCenter becomes unavailable and you power on virtual machines using a VI Client connected directly to an ESX Server host.

If a cluster is valid, this indicates there are enough resources to meet all reservations and to support all running virtual machines. In addition, there is at least one host with enough resources to run each virtual machine assigned to the cluster. If you use a particularly large virtual machine (for example, a virtual machine with an 8GB reservation), you must have at least one host with that much memory. It's not enough if two hosts together fulfill the requirement.

Yellow (Overcommitted) Cluster

A cluster becomes yellow when the tree of resource pools and virtual machines is internally consistent but the user has allowed violation of certain requirements. This happens when capacity is removed from the cluster because, for example, a host fails or is removed and there are no longer enough resources to support all requirements.

Red (Invalid) Cluster

A cluster becomes red when the tree is no longer internally consistent and does not have enough resources available. As a rule, this could happen if you use a VI Client connected to an ESX Server host to start a virtual machine while VirtualCenter is not available. As soon as VirtualCenter becomes available again, it detects the additional use of resources and marks that branch of the tree red. You can remedy the problem either by powering off one or more virtual machines or by adding resources. Until you have remedied the problem, you cannot allocate resources on the red branch of the tree, and DRS will not make any recommendation or migrate virtual machines automatically. As a result, the cluster is likely to become progressively more imbalanced.

Managing DRS

When you enable a cluster for DRS, VirtualCenter continuously monitors the distribution of CPU and memory resources for all hosts and virtual machines in the cluster. DRS compares these metrics to the ideal resource utilization, given the attributes of the resource pools and virtual machines in the cluster and the current load. DRS evaluates both the specified configuration settings and the current load.

Initial Placement

When you power on a virtual machine, VirtualCenter first checks that there are enough resources in the cluster to support that specific virtual machine. VirtualCenter then finds a host on which the virtual machine can run.

- If the cluster is manual, VirtualCenter displays a list of recommended hosts.
- If the cluster is partially automatic or automatic, VirtualCenter places the virtual machine on the appropriate host.



Virtual Machine Migration

When a cluster enabled for DRS becomes unbalanced, DRS makes recommendations or migrates virtual machines depending on the automation level:

- If the cluster is manual or partially automated, VirtualCenter does not take automatic actions to balance resources. Instead, the Summary page indicates that migration recommendations are available and the Migration page displays recommendations for the most efficient use of resources across the cluster.
- If the cluster is fully automated, VirtualCenter places virtual machines that join the cluster on appropriate hosts and migrates running virtual machines between hosts, as needed, to ensure the best possible use of cluster resources. VirtualCenter displays a history of migrations in the Migration tab.

Note: Even in an automatic migration setup, users can explicitly migrate individual virtual machines as they want, but VirtualCenter might move those virtual machines to other hosts to optimize cluster resources. By default, an automation level is specified for the whole cluster. You can also specify a custom automation level for individual virtual machines.

Setting the Migration Threshold

The migration threshold allows you to specify which recommendations are automatically applied when the cluster is in fully automated mode.

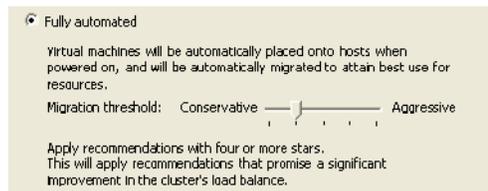


Figure 12. Migration Threshold Settings

You can move the slider to use one of five levels:

- **Level 1** – Apply only five-star recommendations. Includes recommendations that must be followed to satisfy cluster constraints, such as affinity rules and host maintenance. This level indicates a mandatory move, required to satisfy an affinity rule or evacuate a host that is entering maintenance mode.
- **Level 2** – Apply recommendations with four or more stars. Includes Level 1 plus recommendations that promise a significant improvement in the cluster's load balance.
- **Level 3** – Apply recommendations with three or more stars. Includes Level 1 and 2 plus recommendations that promise a good improvement in the cluster's load balance.
- **Level 4** – Apply recommendations with two or more stars. Includes Level 1-3 plus recommendations that promise a moderate improvement in the cluster's load balance.
- **Level 5** – Apply all recommendations. Includes Level 1-4 plus recommendations that promise a slight improvement in the cluster's load balance.

Note: See [VMware DRS Requirements and Best Practices](#), later in this white paper, for some more specific examples and guidelines on setting migration thresholds.



Following Migration Recommendations

If you create a cluster in manual or partially automated mode, VirtualCenter displays migration recommendations on the Migrations page. The system supplies as many recommendations as necessary to properly balance the resources of the cluster. Each recommendation includes a priority, the virtual machine to be moved, current (source) host and target host, CPU load and memory load, and a reason for the recommendation.

Priority	Virtual Machine	Reason	Source Host	Target Host	CPU Load	Memory Load
★★★★★	test06	Satisfy anti-affinity rule	vcuiqa010.en...	vcuiqa012.en...	source: 6019 MHz, target: 24 MHz	source: 2 GB, target: 511 MB
★★★★★	test01	Balance average CPU loads	vcuiqa010.en...	vcuiqa012.en...	source: 6019 MHz, target: 24 MHz	source: 2 GB, target: 511 MB
★★★★★	test02	Balance average CPU loads	vcuiqa010.en...	vcuiqa012.en...	source: 6019 MHz, target: 24 MHz	source: 2 GB, target: 511 MB
★★★	test08	Balance average CPU loads	vcuiqa010.en...	vcuiqa012.en...	source: 6019 MHz, target: 24 MHz	source: 2 GB, target: 511 MB

Apply Migration Recommendation

Figure 13. Migration Recommendations

Migrations that are all part of the same recommendation (for example because of DRS affinity) are indicated by a link icon.

At times, a virtual machine is part of a recommendation involving multiple virtual machines (for example, if the machine is part of an affinity or anti-affinity group). If you select such a virtual machine and the dependent virtual machines are not part of the selection, a warning dialog box asks whether you want to include the dependent virtual machines in this set of migrations.

Using DRS Affinity Rules

After you have created a DRS cluster, you can edit its properties to create rules that specify affinity. You can use these rules to specify that:

- DRS should try to keep certain virtual machines together on the same host, for example, to optimize performance.
- DRS should try to make sure that certain virtual machines are not together, for example, for high availability. You might want to guarantee certain virtual machines are always on different physical hosts. When there's a problem with one host, you don't lose both virtual machines.

Note: These rules are completely different from an individual host's CPU affinity rules. In general, don't specify affinity rules unless you have specific reasons to do so. This will give DRS the maximum flexibility to consider all options for resource management and hence enable the most efficient resource placement. See [VMware DRS Requirements and Best Practices](#), later in this white paper, for some specific examples of affinity rule application.



To create a DRS rule, select the cluster in VirtualCenter and choose **Edit Settings** from the right-click mouse button menu. In the Cluster Settings dialog box, choose the **Rules** option to define a new rule.

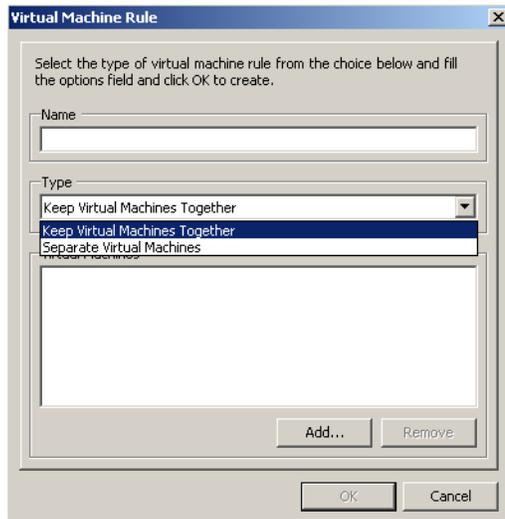


Figure 14. Adding DRS Affinity Rules

From the pop-up menu, choose one of the following options:

- Keep Virtual Machines Together
- Separate Virtual Machines

After you've added the rule, you can edit it, look for conflicting rules, or delete it.

When you add or edit a rule, and the cluster is immediately in violation of that rule, the system continues to operate and tries to correct the violation. For manual and partly automated DRS clusters, migration recommendations are based on both rule fulfillment and load balancing. You are not required to fulfill the rules, but the corresponding recommendations remain until the rules are fulfilled.

Host Removal and Virtual Machines

Both standalone hosts and hosts within a cluster support a Maintenance Mode, which restricts virtual machine operations on the host to allow the user to conveniently shut down running virtual machines in preparation for host shutdown. While in maintenance mode, the host does not allow you to deploy or power on new virtual machines. Virtual machines that are already running on the host continue to run normally. You can either migrate them to another host or shut them down. When there are no more running virtual machines on the host or cluster, its icon changes and its Summary indicates the new state. In addition, menu and command options involving virtual machine deployment are disabled when this host or cluster is selected.

Because a host must be in maintenance mode before you can remove it from a cluster, all virtual machines must be powered off first (unless DRS is also enabled, in which case virtual machines are automatically removed from the host). When you then remove the host from the cluster, the virtual machines that are currently associated with the host are also removed from the cluster.



Removing Hosts with Virtual Machines from a Cluster

If you remove a host from a cluster, the available resources for the cluster decrease. When you remove a host with virtual machines from a cluster, all its virtual machines are removed as well. You can remove a host only if it is in maintenance mode or disconnected.

If the cluster is enabled for DRS, host removal can have the following effect:

- If there are still enough resources to satisfy the reservations of all virtual machines and resource pools in the cluster, the cluster adjusts resource allocation to reflect the reduced amount of resources.
- If there are not enough resources to satisfy the reservations of all resource pools, but there are enough resources to satisfy the reservations for all virtual machines, an alarm is issued and the cluster is marked yellow.

Removing Virtual Machines from a Cluster

You can remove virtual machines by migrating them out of a cluster or by removing a host with virtual machines from the cluster.

You can migrate a virtual machine from a cluster to a standalone host, or from one cluster to another cluster, by using the standard drag-and-drop method or by selecting **Migrate** from the virtual machine's right-button menu or the VirtualCenter menu bar. If the cluster is DRS-enabled and the virtual machine is a member of a DRS cluster affinity rules group, VirtualCenter displays a warning before it allows the migration to proceed. The warning indicates that dependent virtual machines are not migrated automatically, so you have to acknowledge the warning before migration can proceed.

VMware DRS Requirements and Best Practices

Using VMware DRS helps simplify the setup and management of large virtual infrastructure deployments by allowing users to perform administration on a single entity, the cluster. However, for best results, your virtual infrastructure systems should also meet certain prerequisites so that VMware cluster features operate properly:

- In general, DRS will work best if the virtual machines meet VMotion requirements, as discussed in the next section.
- If you want to use DRS for load balancing, the hosts in your cluster must be part of a VMotion network. If the hosts are not in the VMotion network, DRS can still make initial placement recommendations.

VirtualCenter VMotion Requirements

To be configured for VMotion, each host in the cluster must meet the following requirements.

Shared Storage

Ensure that the managed hosts use shared storage. Shared storage is typically on a storage area network (SAN). See the *VMware SAN Configuration Guide* and the *VMware Infrastructure Server Configuration Guide* for additional information.

Shared VMFS Volume

Configure all of your managed hosts to use shared Virtual Machine File System (VMFS) volumes.

- Place the disks of all virtual machines on VMFS volumes that are accessible by both source and target hosts.



- Set access mode for the shared VMFS to “public”.
- Ensure the VMFS volume is sufficiently large to store all of the virtual disks for your virtual machines.
- Ensure all VMFS volumes on your source and destination hosts use volume names, and all virtual machines use those volume names for specifying the virtual disks.

Note: Virtual machine swap files also need to be on a VMFS accessible to both source and destination hosts.

Processor Compatibility

Ensure that the source and destination hosts have a compatible set of processors. VMotion transfers the running architectural state of a virtual machine between underlying VMware ESX Server systems. VMotion compatibility therefore means that the processors of the target host must be able to resume execution using the equivalent instructions, from the point where the processors of the source host were suspended.

Processor clock speeds and cache sizes might vary, but processors must be from the same vendor class (Intel vs. AMD) and same processor family to be compatible for migration with VMotion. Processor families, such as Xeon MP and Opteron, are defined by the processor vendors. Different versions within the same processor family can be distinguished by comparing a processor's model, stepping level, and extended features. In addition:

- In most cases, different processor versions within the same family are similar enough to maintain compatibility.
- In some cases, processor vendors have introduced significant architectural changes within the same processor family (such as 64-bit extensions and SSE3). VMware identifies these exceptions if it cannot guarantee successful migration with VMotion.

Note: VMware works on maintaining VMotion compatibility across the widest range of processors through partnerships with processor and hardware vendors. For current information, contact your VMware representative.

Miscellaneous Virtual Machine, VMFS, and VMotion Requirements

There are a few other specific VMware Infrastructure requirements to be aware of when using DRS:

- For ESX Server 3, the virtual machine configuration file for ESX Server hosts must reside on a VMFS.
- VMotion does not currently support raw or undoable virtual disks or migration of applications clustered using Microsoft Cluster Service (MSCS).
- VMotion requires a Gigabit Ethernet network between hosts.
- VMotion requires the setup of a private, Gigabit Ethernet migration network between all of the VMotion-enabled managed hosts. When VMotion is enabled on a managed host, configure a unique network identity object for the managed host and connect it to the private migration network.



Resource Management and DRS Best Practices

The following guidelines can help you achieve optimal performance for your virtual machines:

- If you expect frequent changes to the total available resources, use Shares, not Reservation, to allocate resources fairly across virtual machines. If you use Shares and you upgrade the host, each virtual machine stays at the same priority (keeps the same number of shares) even though each share represents a larger amount of memory or CPU.
- Use Reservation to specify the minimum acceptable amount of CPU or memory, not the amount you would like to have available. The host assigns additional resources as available based on the number of Shares and the Limit for your virtual machine.
- Use Reservations to specify the minimum reservation for each virtual machine. In contrast to Shares, the amount of concrete resources represented by a reservation does not change when you change the environment, for example, by adding or removing virtual machines.
- Don't set Reservation too high. A reservation that's too high can limit the number of virtual machines in a resource pool.
- When specifying the reservations for virtual machines, always leave some headroom; do not commit all resources. As you move closer to fully reserving all capacity in the system, it becomes increasingly difficult to make changes to reservations and to the resource pool hierarchy without violating admission control. In a DRS-enabled cluster, reservations that fully commit the capacity of the cluster or of individual hosts in the cluster can prevent DRS from migrating virtual machines between hosts.
- Use resource pools for delegated resource management. To fully isolate a resource pool, make the resource pool type Fixed and use Reservation and Limit.
- Group virtual machines for a multi-tier service in a resource pool. This allows resources to be assigned for the service as a whole.
- Affinity rules
 - In general, don't specify affinity rules unless you have specific reasons to do so. This will give DRS the maximum flexibility to consider all options for resource management and hence enable the most efficient resource placement.
 - Reason to select **Keep Virtual Machines Together**: The virtual machines are known to communicate a lot with each other over the network. If the virtual machines reside on the same host, then all network communication takes place over a virtual switch, which is much faster than a physical switch and doesn't use any physical network resources.
 - Reason to select **Separate Virtual Machines**: You want to maximize availability of the virtual machines. For example, if the virtual machines both provide a web server front end to the same application, you wouldn't want both virtual machines to go down at the same time, in the event that the physical host on which they reside goes down.
- When a list of recommendations is made for a cluster in manual mode, the recommendations should be applied in order. Otherwise, performance results could potentially be slightly worse than the initial state.



- The migration threshold for clusters should be set to more aggressive levels (towards 5) in the following cases:
 - if the hosts in the cluster are relatively homogeneous.
 - if the virtual machine's resource utilization does not vary too much over time.
 - if you have relatively few constraints on where a virtual machine can be placed, such as VSMP requirements, large memory requirements, specific virtual network configurations, access to specific datastores, etc.
- The migration threshold for clusters should be set to more conservative levels in the reverse situations to those listed above.
- When deciding which hosts to group into a DRS cluster, try to choose hosts that are as homogeneous as possible, in terms of CPU, memory, and storage, to allow the most flexible management of resources.

Summary

VMware DRS, along with the capabilities of clusters and resource pools, greatly simplifies virtual machine provisioning, resource allocation, load balancing, migration, and the overall management of large virtual infrastructure installations. Providing the capability to aggregate resources pooled from multiple ESX Server hosts allows companies to achieve higher levels of resource availability and utilization, while managing multiple ESX Servers in a cluster, with the simplicity of a single host. In addition, resource pools allow better alignment of IT resources with a company's own organizational structure, business priorities, and values. Access control granularity lets top-level administrators delegate selected permissions and control to business function and department-level administrators for resource pools and the associated virtual machines and applications for which they are responsible.