

# EMC AutoStart: High Availability in Virtual Environments

*Applied Technology*

---

## **Abstract**

EMC® AutoStart™ is an extremely flexible and powerful high availability product. VMware® ESX Server is a robust, production-proven virtualization layer that partitions a physical server into multiple virtual machines. This white paper outlines the support that AutoStart provides for applications in a VMware ESX Server environment and describes the supported data source configurations.

June 2009

---

---

Copyright © 2009 EMC Corporation. All rights reserved.

EMC believes the information in this publication is accurate as of its publication date. The information is subject to change without notice.

THE INFORMATION IN THIS PUBLICATION IS PROVIDED “AS IS.” EMC CORPORATION MAKES NO REPRESENTATIONS OR WARRANTIES OF ANY KIND WITH RESPECT TO THE INFORMATION IN THIS PUBLICATION, AND SPECIFICALLY DISCLAIMS IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Use, copying, and distribution of any EMC software described in this publication requires an applicable software license.

For the most up-to-date listing of EMC product names, see EMC Corporation Trademarks on [EMC.com](http://EMC.com)

All other trademarks used herein are the property of their respective owners.

Part number h6330

---

## Table of Contents

<b>Executive summary .....</b>	<b>4</b>
<b>Introduction .....</b>	<b>4</b>
Audience .....	4
Terminology .....	4
<b>AutoStart and VMware.....</b>	<b>5</b>
Supported ESX Server.....	5
VM clustering configurations.....	5
Cluster in a Box .....	5
Cluster Across Boxes .....	6
Standby Host.....	6
Virtual machine network configuration .....	8
Network redundancy .....	9
AutoStart modules .....	9
AutoStart data sources .....	9
EMC Mirroring .....	10
Shared Disk.....	10
Mount Point .....	11
MirrorView .....	11
SRDF.....	11
Windows Network Share.....	12
RepliStor.....	12
Veritas Volume Manager.....	12
Composite .....	13
VMware VMotion.....	13
VMware HA.....	13
<b>VMware caveats .....</b>	<b>13</b>
Network connection loss .....	14
VMotion .....	14
<b>Conclusion .....</b>	<b>14</b>
<b>References .....</b>	<b>14</b>

---

## Executive summary

IT organizations and data centers are increasingly adopting server virtualization for ease of operation, better hardware utilization, and higher ROI. VMware Infrastructure is the industry-leading and widely deployed infrastructure virtualization software that virtualizes servers, storage, and networking, allowing multiple operating systems and their applications to run independently in virtual machines (VMs) while sharing physical resources. This white paper provides an overview of how EMC® AutoStart™ can be used to provide high availability (HA) for applications in a VMware environment.

VMware itself provides an optional HA service at the virtual machine level. This is a cold standby solution, albeit an automated one. However, customers may need some kind of clustering solution for monitoring their mission-critical applications running in the virtual machines and providing a hot standby at the application level. The hot standby may need to be configured on the local network or on a remote site as a wide area disaster recovery solution. AutoStart can be deployed to provide these application HA clustering solutions in the VMware environment. AutoStart is also extremely flexible and supports most of the storage configurations that customers use in their VMware environments.

## Introduction

The purpose of the white paper is to outline the functionality provided by AutoStart for applications running in a VMware environment.

The paper first discusses the clustering configurations that are normally deployed in a VMware setup. It then describes the behavior of the supported AutoStart modules and data sources inside a VM. The paper does not cover the details of how to configure a particular AutoStart module or data source. For these details, refer to the *EMC AutoStart Administrator's Guide*.

## Audience

The primary target audiences for this white paper are AutoStart solution implementers working with VMware. It is assumed that readers are experienced with AutoStart implementations and have successfully installed an AutoStart domain, at least in a non-VMware environment. In addition, a basic understanding of VMware networking and storage concepts is highly desirable.

## Terminology

The following list is a glossary of important VMware terms.

Term	Description
Guest OS	An operating system that runs inside a virtual machine.
Host	The physical computer on which the virtual machines are installed.
RDM (Raw Device Mapping)	A mechanism that enables a virtual machine to have direct access to a LUN on the physical storage subsystem (Fibre Channel or iSCSI only). RDM can be configured in two modes — physical compatibility mode and virtual compatibility mode.
RDM – Physical Mode	RDM physical compatibility mode allows direct access of the SCSI device to the application. This is a true pass-through mode for all SCSI commands except REPORT LUNS.
RDM – Virtual Mode	RDM virtual compatibility mode provides full virtualization for the mapped RDM device. The guest operating system sees this exactly the same as a virtual disk file in a VMFS volume.
Virtual disk	A file or set of files that appears as a physical disk drive to a guest operating system. These files can be on the host machine or on a remote file system.
VMFS	VMware Virtual Machine File System. A cluster-aware file system that is optimized for storing virtual machines. One VMFS partition is supported per SCSI storage device or SAN.
VMware ESX Server	The VMware infrastructure virtualization layer runs on physical servers that abstracts

	processor, memory, storage, and networking resources into multiple virtual machines.
VMware High Availability (HA)	In the event of a server failure or a virtual machine failure, this feature enables automatic restarting of the affected virtual machines on other production servers (ESX Server) that have spare capacity.
VMware Infrastructure	This is the full virtualization suite from VMware that provides comprehensive server virtualization, management, resource optimization, application availability and operational automation capabilities in an integrated offering.
VMware Storage VMotion	Feature that enables the migration of virtual machine files from one datastore to another without service interruption. Feature that enables the live migration of running virtual machines from one physical server (ESX Server) to another with zero down time, continuous service availability, and complete transaction integrity.
VMware VMotion	Feature that enables the live migration of running virtual machines from one physical server (ESX Server) to another with zero down time, continuous service availability, and complete transaction integrity. Feature that enables the migration of virtual machine files from one datastore to another without service interruption.

## AutoStart and VMware

This section describes the AutoStart features, and limitations if any, for supporting applications in a VMware environment.

### Supported ESX Server

AutoStart has been qualified to work with VMware ESX Server 3.0.x, 3.5, 4.0, and 3i. The following guest operating systems are currently supported:

- 32-bit Windows Guest OS (Windows 2003 and Windows 2008) when running on x86 and x64 platforms.
- 64-bit Windows Guest OS (Windows 2003 and Windows 2008) when running on x64 platforms.

### VM clustering configurations

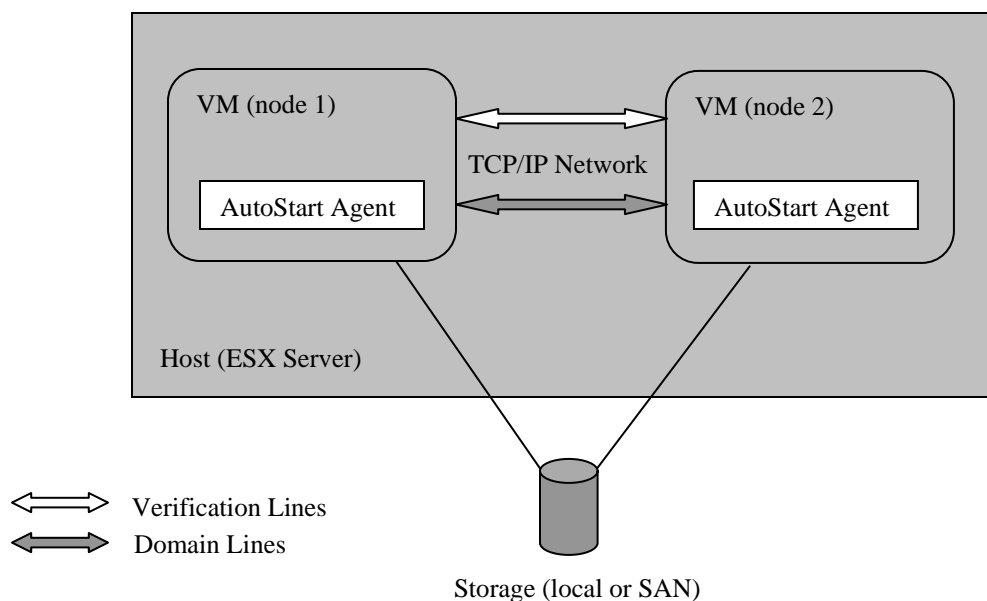
In general, AutoStart supports all the possible application failovers in a VMware setup: VM to VM on the same ESX Server, VM to another VM on a different ESX Server, VM to a physical host, and physical host to a VM. The actual failover scenarios supported depend on the specific application and the specific storage configurations used.

Typically three types of clustering solutions are used in a VMware environment. AutoStart can be used to realize all of these clustering configurations:

- Cluster in a Box, where virtual machines are clustered on a single host
- Cluster Across Boxes, where virtual machines are clustered across different physical hosts
- Standby Host, where physical machines are clustered with virtual machines

#### Cluster in a Box

This configuration consists of two or more clustered virtual machines running on the same ESX Server. The clustered virtual machines are part of the same AutoStart domain. This configuration protects against failures at the operating system or application level but provides no protection against hardware failures. Figure 1 illustrates a sample setup.



**Figure 1. Cluster in a Box**

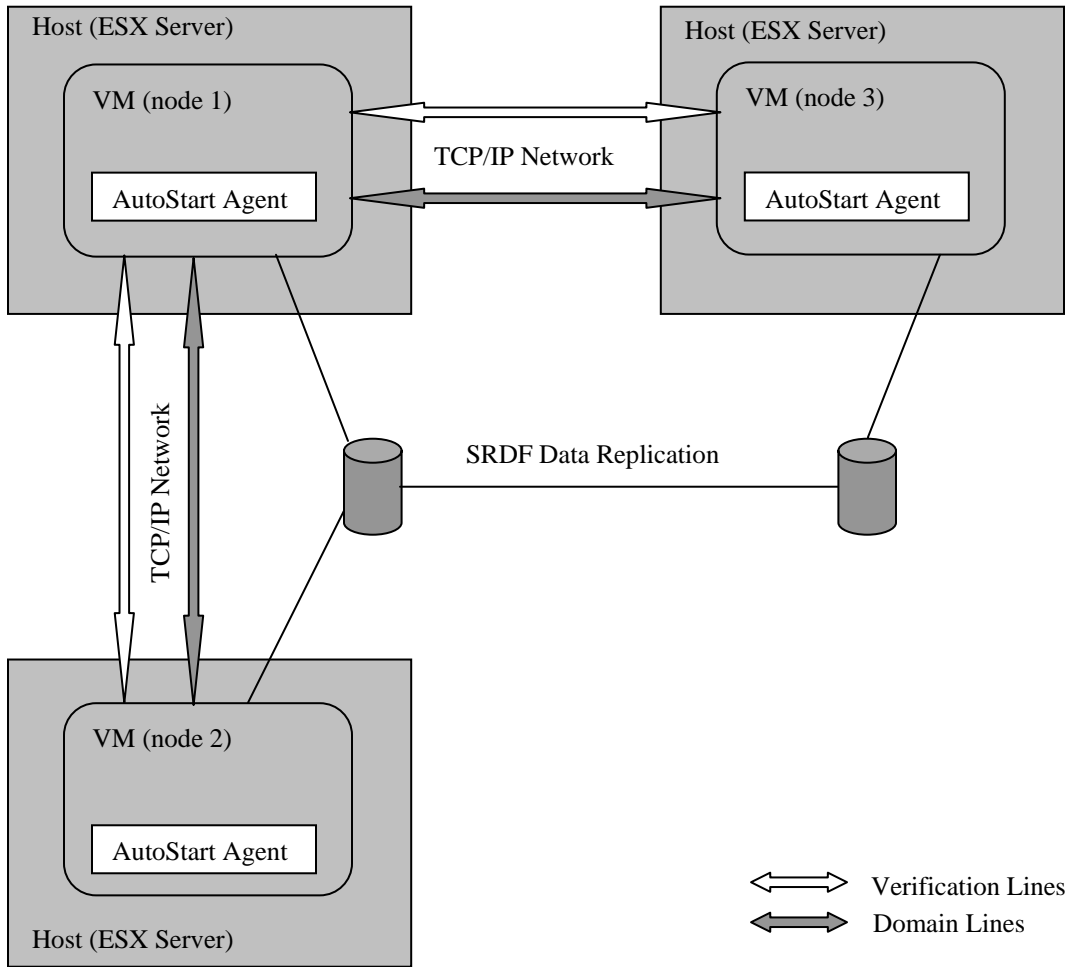
### Cluster Across Boxes

This configuration consists of two or more clustered virtual machines running on different physical hosts (different ESX Servers). The clustered virtual machines are part of the same AutoStart domain. This configuration protects against failures at the application, operating system, or hardware. Figure 2 illustrates a sample setup. Node 3 is at a remote site, while node 1 and node 2 are at the operational site. If node 1 fails, the application can fail over to node 2. If both node 1 and node 2 fail because of a site failure, the application can fail over to node 3 at the DR site.

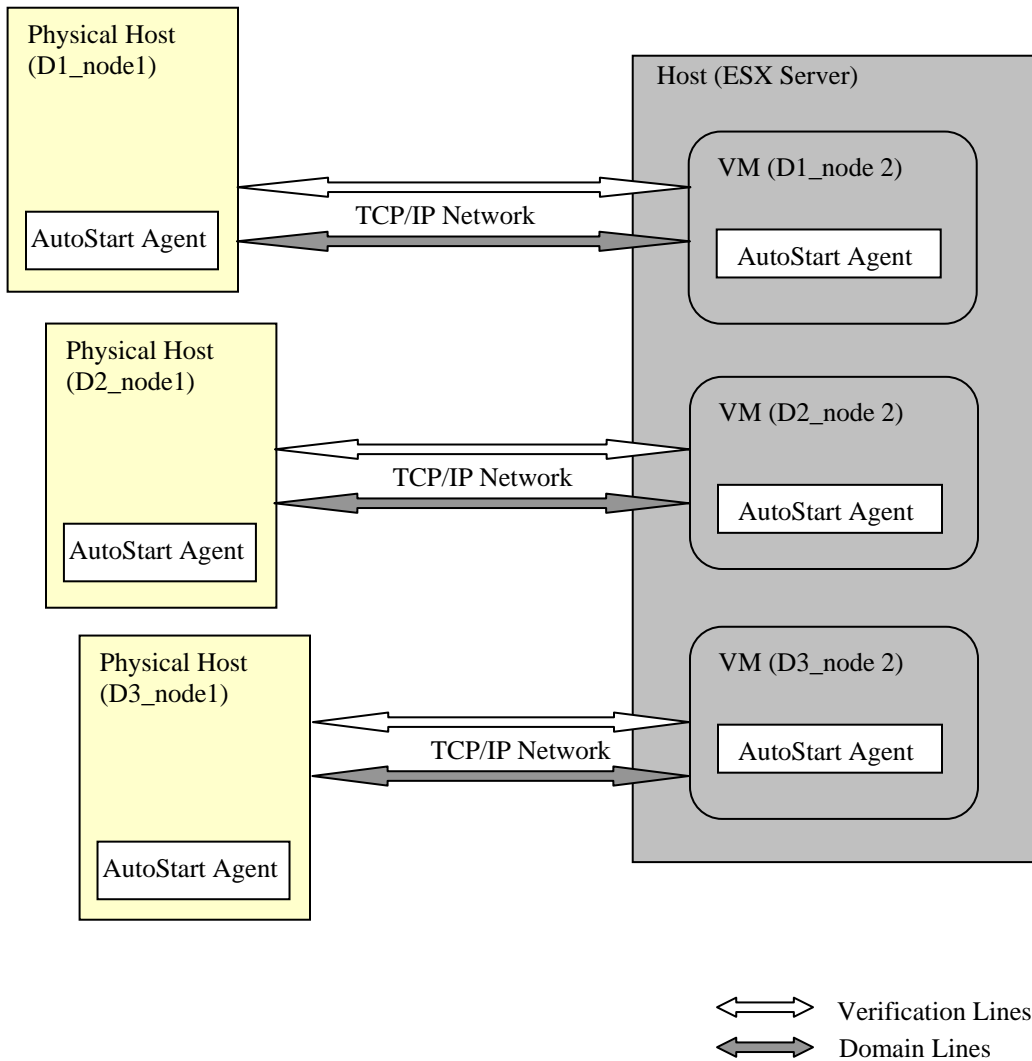
Note that each of the physical hosts (ESX servers) may be running other virtual machines that are not part of this AutoStart domain. Also note that explicit network connectivity has not been shown between node 3 and node 2; this is obvious from the fact that they are part of the same AutoStart domain.

### Standby Host

This configuration is also called the N+1 cluster configuration. The configuration involves N physical machines, with each of the physical machines backed up by a virtual machine running on the standby host. In the event of a failure in any of the physical hosts, the corresponding virtual machine takes over. This is a simple clustering solution that is useful where the hardware requirements are low. Figure 3 illustrates a sample setup. There are three physical machines involved and each has a corresponding virtual machine configured on the ESX Server. Each of the physical machines and its corresponding virtual machine constitute one AutoStart domain.



**Figure 2. Cluster Across Boxes**



**Figure 3. Standby Host**

### ***Virtual machine network configuration***

AutoStart recommends that every node in an AutoStart domain be configured with at least three different network lines, two domain lines, and one verification line. The nodes exchange heart beats over the domain lines. It would be advisable to enable multicast for the node heart beats. It is also recommended that external IP addresses be configured as the Isolation Addresses for the node. The virtual switch in ESX Server will make the nodes within the ESX Server box reachable even if external lines are down.

Each virtual machine has five PCI slots available by default. An AutoStart cluster would use three of these slots for the network adapters and the remaining two probably for the SCSI host bus adapters. So the recommendation is for the ESX Server host to be configured with at least three different network cards.

---

## Network redundancy

VMware supports a feature called NIC teaming. NIC teaming groups one or more NICs together for network load balancing and for supporting NIC-to-NIC failover. This NIC teaming is a feature at the ESX Server level and is transparent to the Guest OS, and hence the AutoStart. However NIC teaming at the operating system level is an unsupported configuration with AutoStart.

Users who want to deploy path level network redundancy for virtual machines can use the support AutoStart provides for managed IP addresses and Windows node aliases. AutoStart supports migration of the managed IP address or the Windows node alias to any other node in the domain along with the relevant application services. As noted in Table 1, all possible failover scenarios are supported.

**Table 1. Failover scenarios**

	VM <=> VM	VM<=>VM (Another ESX server)	PHY<=>VM
Managed IP address	yes	yes	yes
Windows node alias	yes	yes	yes

## ***AutoStart modules***

AutoStart provides the following modules for Windows operating systems:

- SQL Server 2005
- SQL Server 2008
- Exchange 2003 (32-bit Windows only)
- Exchange 2007 (64-bit Windows only)
- Oracle
- Print Services

All the modules have been qualified to work in the VMware environment. These modules provide HA service to the corresponding applications and users can deploy them to support all possible application failover modes: VM to VM on the same ESX Server, VM to VM on another local or remote ESX Server, and VM to a physical host, or vice versa.

## ***AutoStart data sources***

Data sources allow any AutoStart node or applications running on the AutoStart node to access data from a common storage device. AutoStart data sources manage access to the disk resources so that in the event of a failure, a node is not cut off from its data. AutoStart supports the following data sources in a VMware environment.

- EMC Mirroring
- Shared Disk
- Mount point
- EMC MirrorView™/S and MirrorView/A
- EMC SRDF®/S and SRDF/A
- Windows Network Share
- EMC RepliStor®
- Veritas Volume Manager
- Composite

The disk resources mapped to the Windows Guest OS could have been configured as part of a VMFS volume or may have been configured as RDM devices. The subsequent sections describe the support provided by AutoStart data sources for each of the possible device configurations in VMware.

---

## EMC Mirroring

The EMC Mirroring data source provides built-in volume level mirroring for Windows servers. Mirroring is synchronous and is supported only between two nodes. Mirroring is supported between two VM or between a VM and a physical machine. Table 2 shows the different configurations in VMware that support mirroring.

**Table 2. VMware configurations supporting mirroring**

	VM <=> VM	VM<=>VM (Another ESX server)	PHY<=>VM
VMFS	yes	yes	yes
RDM – physical mode	yes	yes	yes
RDM – virtual mode	yes	yes	yes

## Shared Disk

The Shared Disk data source is used when two or more nodes in the domain are connected to the same shared disk drive. The Shared Disk data source type allows AutoStart to manage which node has access to the disk to perform read and write operations. Table 3 lists the data source support for the different configurations in which a shared disk may be deployed.

**Table 3. VMware configurations supporting shared disk**

	VM <=> VM	VM<=>VM (Another ESX server)	PHY<=>VM
VMFS	yes	yes	no
RDM – physical mode	yes	yes	yes
RDM – virtual mode	yes	yes	yes

The table lists shared disk as supported between two (or more) virtual machines on a single ESX Server and between virtual machines across ESX Servers. However, it is important to note that a given device cannot be configured to be shared between two virtual machines on an ESX Server and between virtual machines across ESX Servers at the same time. The important configuration steps to be carried out in VMware for creating a shared disk are noted next.

To create a shared disk between VMs on a single ESX Server (VMFS volume):

- Create the shared disk from the command line using the “thick” option.
- Change the SCSI Bus Sharing to Virtual for the associated SCSI controller.
- Assign the disk to the desired virtual machines after powering them off.

To create a shared disk between VMs across ESX Servers (VMFS volume):

- Create the shared disk from the command line using the “thick” option.
- Change the SCSI Bus Sharing to Physical for the associated SCSI controller.
- Assign the disk to the desired virtual machines after powering them off.

To create a shared disk between VMs on a single ESX Server (RDM device):

- Create a new RDM disk device from the ESX Server.
- Change the SCSI Bus Sharing to Virtual for the associated SCSI controller.
- Assign the disk to the desired virtual machines after powering them off.

To create a shared disk between VMs across ESX Servers (RDM device):

- Create a new RDM disk device from the ESX Server.
- Change the SCSI Bus Sharing to Physical for the associated SCSI controller.
- Assign the disk to the desired virtual machines after powering them off.

## Mount Point

The prerequisite for creating a mount point is to have a shared disk configured. So the behavior of the mount point data source is exactly similar to that of the shared disk. Table 4 gives a list of the supported configurations and failover modes for the mount point data source.

**Table 4. VMware configurations supporting a mount point**

	VM <=> VM	VM<=>VM (Another ESX server)	PHY<=>VM
VMFS	yes	yes	no
RDM – physical mode	yes	yes	yes
RDM – virtual mode	yes	yes	yes

## MirrorView

The MirrorView data source controls access from the AutoStart node to the devices configured on the EMC CLARiiON® system. Both synchronous and asynchronous modes of MirrorView are supported in a VMware environment. Table 5 gives a list of the supported configurations and failover modes for the MirrorView data source.

**Table 5. VMware configurations supporting MirrorView**

	VM <=> VM	VM<=>VM (Another ESX server)	PHY<=>VM
VMFS	NA	yes	no
RDM – physical mode	NA	yes	yes
RDM – virtual mode	NA	no	no

The important configuration steps to be carried out in a VMware environment for configuring MirrorView are as follows:

- Create the remote mirror in the CLARiiON between the ESX Servers (VM<=>VM) or the ESX Server and the physical host (PHY<=>VM).
- Assign the LUNs as RDM or VMFS devices to the virtual machines.
- Create the MirrorView data source from the AutoStart console with the remote mirror created above.

In the RDM mode, the MirrorView data source is used in combination with the Shared Disk data source to support local and remote failovers. In VMFS mode, the MirrorView data source is used in combination with the VMFS data source to support local and remote application failovers. The VMFS data source is provided free with AutoStart and enables the MirrorView data source to support failovers with VMFS disks. The VMFS data source communicates with the VirtualCenter host and does an add or remove of the associated virtual disks in a virtual machine when the Resource Group is brought online or offline.

## SRDF

The SRDF data source controls access from the AutoStart node to the devices configured on the EMC Symmetrix® system. Both synchronous and asynchronous modes of SRDF are supported in a VMware environment.

Table 6 lists the supported configurations and failover modes for the SRDF data source.

---

**Table 6. VMware configurations supporting SRDF**

	VM <=> VM	VM<=>VM (Another ESX server)	PHY<=>VM
VMFS	NA	yes	no
RDM – physical mode	NA	yes	yes
RDM – virtual mode	NA	no	no

The important configuration steps to be carried out in a VMware environment for configuring SRDF are as follows:

- Create the device group or consistency group from the virtual machines using SYMCLI commands.
- Assign the LUNs as RDM or VMFS devices to the virtual machines.
- Create the SRDF data source from the AutoStart console.

In RDM mode, the SRDF data source is used in combination with the Shared Disk data source to support local and remote failovers. In the VMFS mode, the SRDF data source is used in combination with the VMFS data source to support local and remote application failovers. VMFS data source is provided free with AutoStart and enables the SRDF data source to support failovers with VMFS disks. The VMFS data source communicates with the VirtualCenter host and does an add or remove of the associated virtual disks in a virtual machine when the Resource Group is brought online or offline.

## Windows Network Share

The Windows Network Share data source allows AutoStart to manage access to network shares defined in a Windows domain from nodes within the AutoStart domain. The network shares operate at a much higher level using the TCP/IP protocol. So there is no dependency on the VMware environment. As Table 7 indicates, all possible node failover scenarios are supported.

**Table 7. VMware configurations supporting Windows Network Share**

	VM <=> VM	VM<=>VM (Another ESX server)	PHY<=>VM
Windows Network Share	yes	yes	yes

## RepliStor

The RepliStor data source works with the RepliStor product that provides asynchronous replication for Windows servers. Table 8 lists the supported configurations and failover modes for the RepliStor data source.

**Table 8. VMware configurations supporting RepliStor**

	VM <=> VM	VM<=>VM (Another ESX server)	PHY<=>VM
VMFS	yes	yes	yes
RDM – physical mode	yes	yes	yes
RDM – virtual mode	yes	yes	yes

## Veritas Volume Manager

AutoStart provides support for Veritas Volume Manager (VxVM) on VMware. VxVM is a logical volume manager. A disk group can be made up of multiple volumes, which in turn are made up of one or more disks. VxVM provides the underlying management of these disks. The AutoStart data source manages the access to the disk group. Table 9 lists the supported configurations and failover modes for the VxVM data source.

---

**Table 9. VMware configurations supporting Veritas Volume Manager**

	VM <=> VM	VM<=>VM (Another ESX server)	PHY<=>VM
VMFS	yes	yes	no
RDM – physical mode	yes	yes	yes
RDM – virtual mode	no	no	no

## Composite

The Composite data source is used to group two or more individual data sources into a single data source. The behavior of the Composite data source depends on the individual data sources; there are no specific VMware dependencies for this data source.

**Table 10. VMware configurations supporting the Composite data source**

	VM <=> VM	VM<=>VM (Another ESX server)	PHY<=>VM
VMFS	yes	yes	yes
RDM – physical mode	yes	yes	yes
RDM – virtual mode	yes	yes	yes

Composite data source can have four types of data sources:

- EMC Mirroring
- Shared disk
- RepliStor
- Veritas Volume Manager

The configurations supported here are defined based on the individual data source support for each configuration (for example, if the Composite data source contains only EMC Mirroring and RepliStor in it, all three configuration mentioned in the table are supported and if the Composite data source contains Veritas Volume Manager in it, then any of the configurations in RDM – virtual mode are *not* supported).

## VMware VMotion

VMware VMotion enables the live migration of running virtual machines from one host (ESX Server) to another host (ESX Server). VMotion is a hot migration and there is no virtual machine downtime during the migration. VMotion cannot be used to migrate virtual machines across data centers. Typical scenarios where VMotion is employed include taking a data backup, a hardware upgrade, or for load balancing in case of an excessive I/O on the host.

AutoStart is tolerant to a virtual machine migration using VMotion. AutoStart modules, data sources, node alias, and managed IP are tolerant to VMotion and continue to work after a virtual machine migration without any additional configurations. Existing network connections continue to be valid after a VMotion. Hence AutoStart console and AutoStart CLI sessions would be transparent to a virtual machine migration.

## VMware HA

AutoStart has not been qualified to work with VMware HA.

## VMware caveats

This section lists some of the known VMware issues and the current limitations with regard to AutoStart support for VMware.

---

## Network connection loss

Sometimes when a new disk device is being created in the virtual machine, the network connections are reset. The primary network interface will be deleted and a new network connection added. This is a known VMware issue. The following VMware KnowledgeBase article provides more information:

[http://kb.vmware.com/selfservice/microsites/search.do?language=en\\_US&cmd=displayKC&externalId=1513](http://kb.vmware.com/selfservice/microsites/search.do?language=en_US&cmd=displayKC&externalId=1513)

The problem happens because adding a new virtual SCSI controller pushes the virtual Ethernet controllers to new slots on the PCI bus. Typically, an AutoStart node would use two virtual SCSI controllers, SCSI 0 and SCSI 1. SCSI 0 is created by default when the ESX Server boot disks are created.

The suggested workaround is that the user also create SCSI controller 1 before AutoStart or networking is configured. After this any disks created in the future (on SCSI controller 0 or 1) would not exhibit this behavior. Note that virtual SCSI controller 1 would be created automatically when the user creates a disk on controller 1.

## VMotion

VMware does not allow VMotion to proceed if the associated virtual machine has storage shared with other virtual machines. This is a VMware limitation. One prerequisite for VMotion is that the SCSI bus sharing mode of the SCSI controller be set to none. The following VMware Communities page has more information:

<http://communities.vmware.com/message/1025000>

This means that the AutoStart Shared Disk data source cannot be used with VMotion. The Shared Disk data source requires that the associated storage be accessible from all the virtual machines involved.

## Conclusion

AutoStart provides a viable clustering solution for applications deployed on VMware virtual machines. AutoStart supports most of the storage configurations possible in a VMware setup, and in this regard is very flexible for deploying customized clustering solutions. AutoStart clustering employs a replicated database and, unlike some of the other competitor products, is not based on shared storage. This turns out to be more advantageous in the VMware world.

## References

The following can be found on Powerlink<sup>®</sup>, EMC's password-protected customer- and partner-only extranet.

- *EMC AutoStart Version 5.3 SP3 Installation Guide*
- *EMC AutoStart Version 5.3 SP3 Administrator's Guide*
- *EMC AutoStart Version 5.3 SP3 Release Notes*

The following can be found on the VMware website:

- *Introduction to VMware Infrastructure* (Item: EN-000019-00)
- *VMware ESX Server 3 Configuration Guide* (Item: EN-000031-00)
- *Setup for Microsoft Cluster Service* white paper
- *VMware Virtual Machine File System: Technical Overview and Best Practices* white paper