# What's New in VMware vSphere® 5.5 Platform

Version 1.4

vmware®

## Table of Contents

# Introduction

VMware vSphere® 5.5 introduces many new features and enhancements to further extend the core capabilities of the vSphere platform. This paper will discuss features and capabilities of the vSphere platform, including vSphere ESXi Hypervisor™, VMware vSphere High Availability (vSphere HA), virtual machines, VMware vCenter Server™, storage, networking and vSphere Big Data Extensions.

This paper is organized into the following five sections:

- **vSphere ESXi Hypervisor Enhancements**
  - Hot-Pluggable SSD PCI Express (PCIe) Devices
  - Support for Reliable Memory Technology
  - Enhancements for CPU C-States
- **Virtual Machine Enhancements**
  - Virtual Machine Compatibility with VMware ESXi™ 5.5
  - Expanded Virtual Graphics Support
  - Graphic Acceleration for Linux Guests
- **VMware vCenter Server Enhancements**
  - VMware® vCenter™ Single Sign-On
  - VMware vSphere Web Client
  - VMware vCenter Server Appliance™
  - vSphere App HA
  - vSphere HA and VMware vSphere Distributed Resource Scheduler™ (vSphere DRS) Virtual Machine–Virtual Machine Affinity Rules Enhancements
  - vSphere Big Data Extensions
- **vSphere Storage Enhancements**
  - Support for 62TB VMDK
  - MSCS Updates
  - vSphere 5.1 Feature Updates
  - 16GB E2E support
  - PDL AutoRemove
  - vSphere Replication Interoperability
  - vSphere Replication Multi-Point-in-Time Snapshot Retention
  - vSphere Flash Read Cache
- **vSphere Networking Enhancements**
  - Link Aggregation Control Protocol Enhancements
  - Traffic Filtering
  - Quality of Service Tagging
  - SR-IOV Enhancements
  - Enhanced Host-Level Packet Capture
  - 40GB NIC support

# vSphere ESXi Hypervisor Enhancements

## Hot-Pluggable PCIe SSD Devices

The ability to hot-swap traditional storage devices such as SATA and SAS hard disks on a running vSphere host has been a huge benefit to systems administrators in reducing the amount of downtime for virtual machine workloads. Solid-state disks (SSDs) are becoming more prevalent in the enterprise datacenter, and this same capability has been expanded to support SSD devices. Similarly as with SATA and SAS hard disks, users are now able to hot-add or hot-remove an SSD device while a vSphere host is running, and the underlying storage stack detects the operation.

## Support for Reliable Memory Technology

The most critical component to vSphere ESXi Hypervisor is the VMkernel, which is a purpose-built operating system (OS) to run virtual machines. Because vSphere ESXi Hypervisor runs directly in memory, an error in it can potentially crash it and the virtual machines running on the host. To provide greater resiliency and to protect against memory errors, vSphere ESXi Hypervisor can now take advantage of new hardware vendor–enabled Reliable Memory Technology, a CPU hardware feature through which a region of memory is reported from the hardware to vSphere ESXi Hypervisor as being more "reliable." This information is then used to optimize the placement of the VMkernel and other critical components such as the initial thread, hostd and the watchdog process and helps guard against memory errors.

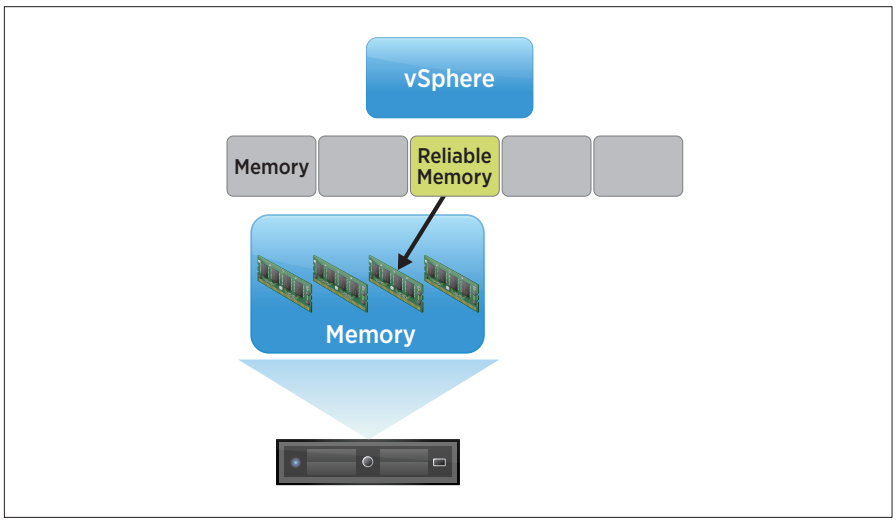


**Figure 1.**

## Enhancements to CPU C-States

In vSphere 5.1 and earlier, the balanced policy for host power management leveraged only the performance state (P-state), which kept the processor running at a lower frequency and voltage. In vSphere 5.5, the deep processor power state (C-state) also is used, providing additional power savings. Another potential benefit of reduced power consumption is with inherent increased performance, because turbo mode frequencies on Intel chipsets can be reached more quickly while other CPU cores in the physical package are in deep C-states.

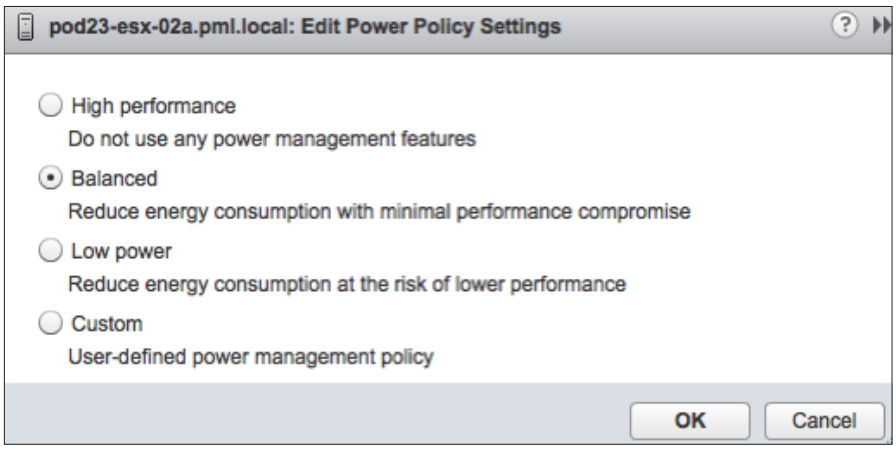


**Figure 2.**

# Virtual Machine Enhancements

## Virtual Machine Compatibility with VMware ESXi 5.5

vSphere 5.5 introduces a new virtual machine compatibility with several new features such as LSI SAS support for Oracle Solaris 11 OS, enablement for new CPU architectures, and a new advanced host controller interface (AHCI). This new virtual-SATA controller supports both virtual disks and CD-ROM devices that can connect up to 30 devices per controller, with a total of four controllers. This enables a virtual machine to have as many as 120 disk devices, compared to the previous limit of 60.

Table 1 summarizes the virtual machine compatibility levels supported in vSphere 5.5.

| vSphere Releases | Virtual Machine Hardware Version | vSphere 5.5 Compatibility |
|---|---|---|
| Virtual Infrastructure 3.5 | Version 4 | VMware ESX/ESXi 3.5 and later |
| vSphere 4.0 | Version 7 | VMware ESX/ESXi 4.0 and later |
| vSphere 4.1 | Version 7 | VMware ESX/ESXi 4.0 and later |
| vSphere 5.0 | Version 8 | VMware ESXi 5.0 and later |
| vSphere 5.1 | Version 9 | VMware ESXi 5.1 and later |
| vSphere 5.5 | Version 10 | VMware ESXi 5.5 and later |

**Table 1.**

## Expanded Virtual Graphics Support

vSphere 5.1 was the first vSphere release to provide support for hardware-accelerated 3D graphics—virtual shared graphics acceleration (vSGA)—inside of a virtual machine. That support was limited to only NVIDIA-based GPUs. With vSphere 5.5, vSGA support has been expanded to include both NVIDIA- and AMD-based GPUs. Virtual machines with graphic-intensive workloads or applications that typically have required hardware-based GPUs can now take advantage of additional GPU vendors, makes and models. See the *VMware Compatibility Guide* for details on supported GPU adapters.

There are three supported rendering modes for a virtual machine configured with a vSGA: automatic, hardware and software. Virtual machines still can leverage VMware vSphere vMotion® technology, even across a heterogeneous mix of GPU vendors, without any downtime or interruptions to the virtual machine. If automatic mode is enabled and a GPU is not available at the destination vSphere host, software rendering automatically is enabled. If hardware mode is configured and a GPU does not exist at the destination vSphere host, a vSphere vMotion instance is not attempted.

vSGA support can be enabled using both the vSphere Web Client and VMware Horizon View™ for Microsoft Windows 7 OS and Windows 8 OS. The following Linux OSs also are supported: Fedora 17 or later, Ubuntu 12 or later and Red Hat Enterprise Linux (RHEL) 7. Controlling vSGA use in Linux OSs is supported using the vSphere Web Client.

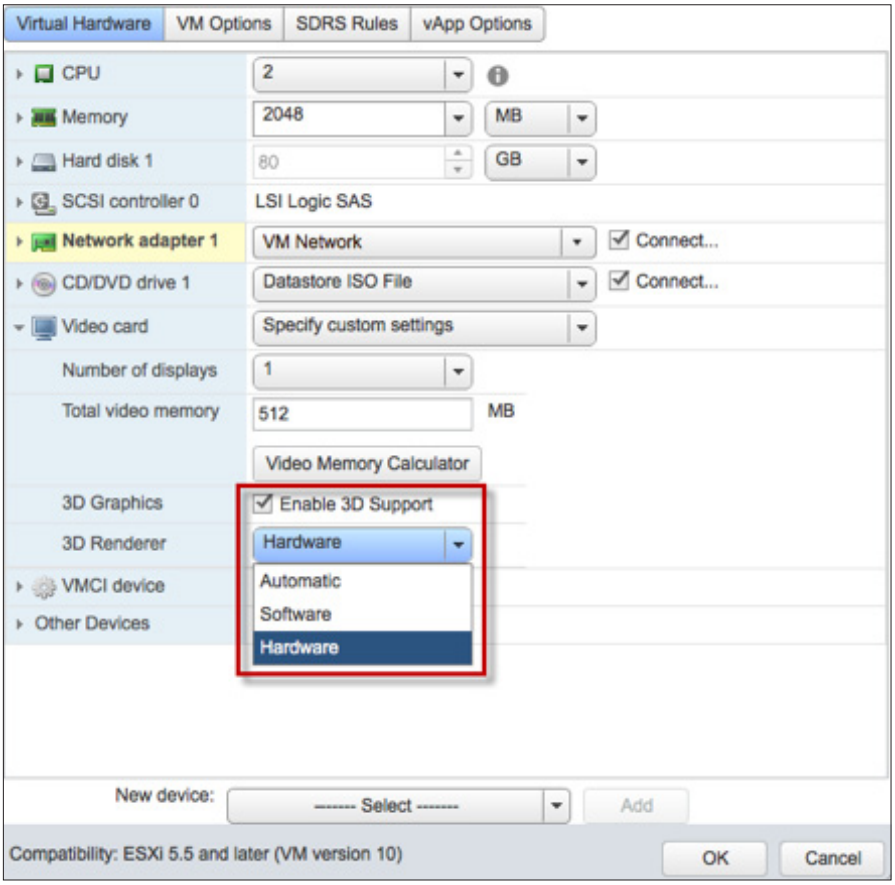


**Figure 3.**

## Graphic Acceleration for Linux Guests

With vSphere 5.5, graphic acceleration is now possible for Linux guest OSs. Leveraging a GPU on a vSphere host can help improve the performance and scalability of all graphics-related operations. In providing this support, VMware also is the first to develop a new guest driver that accelerates the *entire* Linux graphics stack for modern Linux distributions. VMware also is contributing 100 percent of the Linux guest driver code back to the open-source community. This means that any modern GNU/Linux distribution can package the VMware guest driver and provide out-of-the-box support for accelerated graphics without any additional tools or package installation.

The following Linux distributions are supported:

- Ubuntu: 12.04 and later
- Fedora: 17 and later
- RHEL 7

With the new guest driver, modern Linux distributions are enabled to support technologies such as the following:

- OpenGL 2.1
- DRM kernel mode setting
- Xrandr
- XRender
- Xv


**Figure 4.**

# VMware vCenter Server Enhancements

## vCenter Single Sign-On

vCenter Single Sign-On server 5.5, the authentication services of VMware vCloud® Suite, has been greatly enhanced to provide a richer experience that enables users to log in to vCloud Suite products in a true one-touch, single sign-on manner. This feature provided challenges for users in a previous release. As a result of extensive feedback, the following vCenter Single Sign-On enhancements have been made:

**Simplified deployment –** A single installation model for customers of all sizes is now offered.

**Enhanced Microsoft Active Directory integration –** The addition of native Active Directory support enables cross-domain authentication with one- and two-way trusts common in multidomain environments.

**Architecture –** Built from the ground up, this architecture removes the requirement of a database and now delivers a multimaster authentication solution with built-in replication and support for multiple tenants.

## vSphere Web Client

The platform-agnostic vSphere Web Client, which replaces the traditional vSphere Client™, continues to exclusively feature all-new vSphere 5.5 technologies and to lead the way in VMware virtualization and cloud management technologies.

Increased platform support – With vSphere 5.5, full client support for Mac OS X is now available in the vSphere Web Client. This includes native remote console for a virtual machine. Administrators and end users can now access and manage their vSphere environment using the desktop platform they are most comfortable with. Fully supported browsers include both Firefox and Chrome.

Improved usability experience – The vSphere Web Client includes the following key new features that improve overall usability and provide the administrator with a more native application feel:

- **Drag and drop –** Administrators now can drag and drop objects from the center panel onto the vSphere inventory, enabling them to quickly perform bulk actions. Default actions begin when the "drop" occurs, helping accelerate workflow actions. This enables administrators to perform "bulk" operations with ease. For example, to move multiple virtual machines, grab and drag them to the new host to start the migration workflow.
- **Filters –** Administrators can now select properties on a list of displayed objects and selected filters to meet specific search criteria. Displayed objects are dynamically updated to reflect the specific filters selected. Using filters, administrators can quickly narrow down to the most significant objects. For example, two checkbox filters can enable an administrator to see all virtual machines on a host that are powered on and running Windows Server 2008.
- **Recent items –** Administrators spend most of their day working on a handful of objects. The new recent-items navigation aid enables them to navigate with ease, typically by using one click between their most commonly used objects.

## vCenter Server Appliance

The popularity of vCenter Server Appliance has grown over the course of its previous releases. Although it offers matched API functionality to the installable vCenter Server version on Windows, administrators have found its widespread adoption prospects to be limited. One area of concern has been the embedded database that has previously been targeted for small datacenter environments. With the release of vSphere 5.5, the vCenter Server Appliance addresses this with a reengineered, embedded vPostgres database that can now support as many as 100 vSphere hosts or 3,000 virtual machines (with appropriate sizing). With new scalability maximums and simplified vCenter Server deployment and management, the vCenter Server Appliance offers an attractive alternative to the Windows version of vCenter Server when planning a new installation of vCenter Server 5.5.

## vSphere App HA

In versions earlier than vSphere 5.5, it was possible to enable virtual machine monitoring, which checks for the presence of "heartbeats" from VMware Tools™ as well as I/O activity from the virtual machine. If neither of these is detected in the specified amount of time, vSphere HA resets the virtual machine. In addition to virtual machine monitoring, users can leverage third-party application monitoring agents or create their own agents to work with vSphere HA using the VMware vSphere Guest SDK.

In vSphere 5.5, VMware has simplified application monitoring for vSphere HA with the introduction of vSphere App HA. This new feature works in conjunction with vSphere HA host monitoring and virtual machine monitoring to further improve application uptime. vSphere App HA can be configured to restart an application service when an issue is detected. It is possible to protect several commonly used, off-the-shelf applications. vSphere HA can also reset the virtual machine if the application fails to restart.

## Architecture Overview

vSphere App HA leverages VMware vFabric™ Hyperic® to monitor applications. Deploying vSphere App HA begins with provisioning two virtual appliances per vCenter Server: vSphere App HA and vFabric Hyperic. vSphere App HA virtual appliance stores and manages vSphere App HA policies. vFabric Hyperic monitors applications and enforces vSphere App HA policies, which are discussed in greater detail in the following section. It is possible to deploy these virtual appliances to a cluster other than the one running the protected applications; for example, a management cluster.

After the simple process of deploying the vFabric Hyperic and vSphere App HA virtual appliances, vFabric Hyperic agents are installed in the virtual machines containing applications that will be protected by vSphere App HA. These agents must be able to reliably communicate with the vFabric Hyperic virtual appliance.

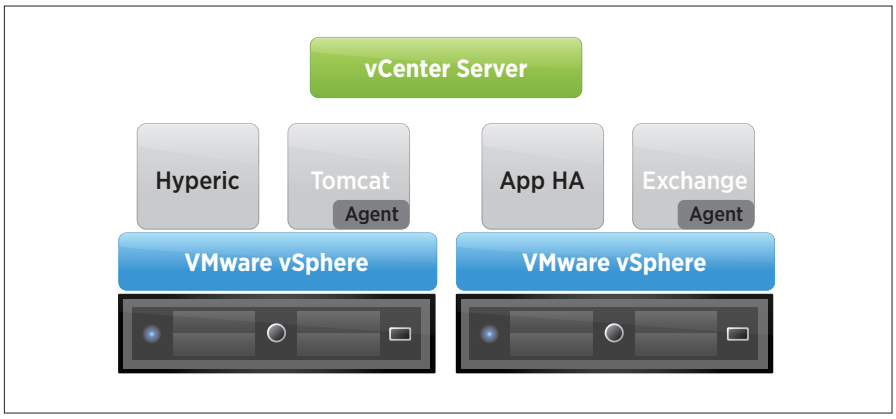


**Figure 5.** vSphere App HA Architecture Overview

## vSphere App HA Policies

Policies define items such as the number of minutes vSphere App HA will wait for the service to start, the option to reset the virtual machine if the service fails to start, and the option to reset the virtual machine when the service is unstable. Policies can be configured to trigger vCenter Server alarms when the service is down and the virtual machine is reset. Email notification is also available.

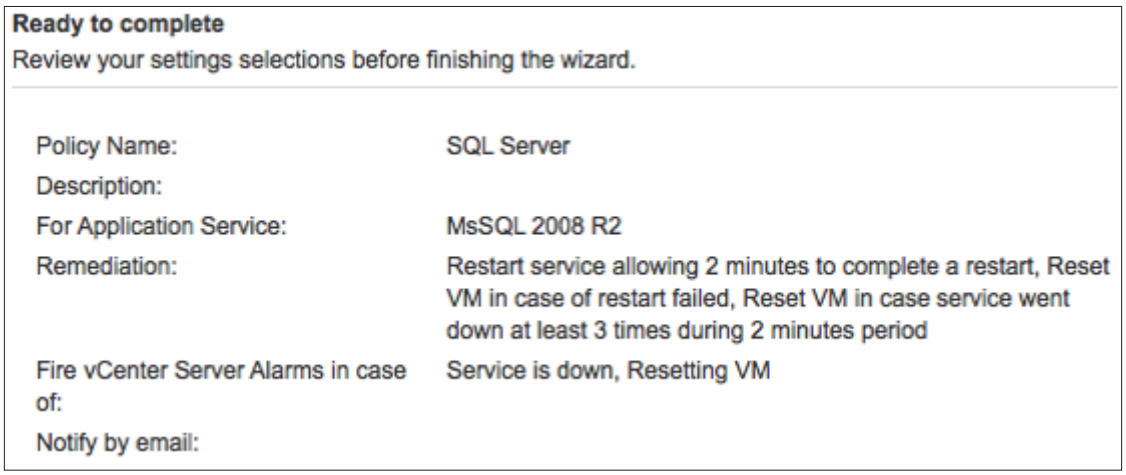


**Figure 6.** vSphere App HA Policy

## Enabling Protection for an Application Service

Application protection is enabled when a policy is assigned. Right-click the application service to assign a policy, as shown in Figure 7. vSphere HA virtual machine monitoring and application monitoring must be enabled.

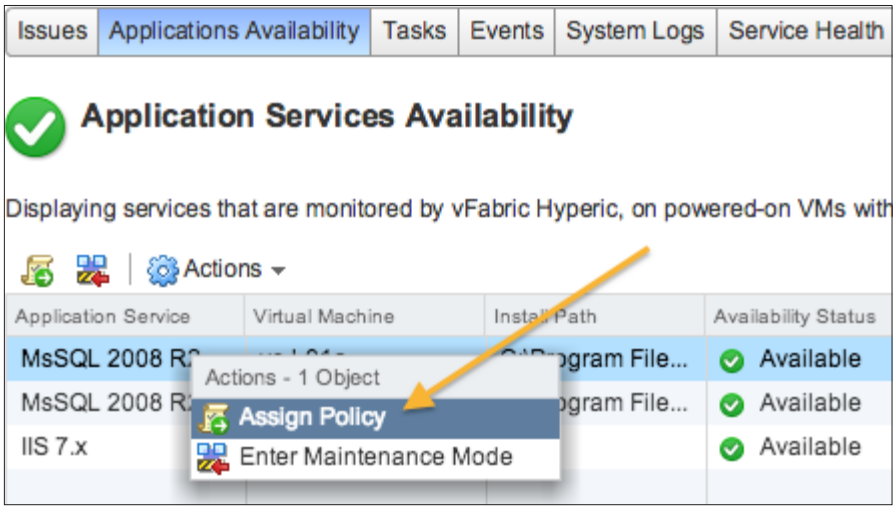


**Figure 7.** Assigning a vSphere App HA Policy

## vSphere HA and vSphere Distributed Resource Scheduler Virtual Machine–Virtual Machine Affinity Rules

vSphere DRS can configure DRS affinity rules, which help maintain the placement of virtual machines on hosts within a cluster. Various rules can be configured. One such rule, a virtual machine–virtual machine affinity rule, specifies whether selected virtual machines should be kept together on the same host or kept on separate hosts. A rule that keeps selected virtual machines on separate hosts is called a virtual machine–virtual machine antiaffinity rule and is typically used to manage the placement of virtual machines for availability purposes.

In versions earlier than vSphere 5.5, vSphere HA did not detect virtual machine–virtual machine antiaffinity rules, so it might have violated one during a vSphere HA failover event. vSphere DRS, if fully enabled, evaluates the environment, detects such violations and attempts a vSphere vMotion migration of one of the virtual machines to a separate host to satisfy the virtual machine–virtual machine antiaffinity rule. In a large majority of environments, this operation is acceptable and does not cause issues. However, some environments might have strict multitenancy or compliance restrictions that require consistent virtual machine separation. Another use case is an application with high sensitivity to latency; for example, a telephony application, where migration between hosts might cause adverse effects.

To address the need for maintaining placement of virtual machines on separate hosts—without vSphere vMotion migration—after a host failure, vSphere HA in vSphere 5.5 has been enhanced to conform with virtual machine–virtual machine antiaffinity rules. Application availability is maintained by controlling the placement of virtual machines recovered by vSphere HA without migration. This capability is configured as an advanced option in vSphere 5.5.

## VMware vSphere Data Protection Enhancements

VMware vSphere Data Protection™ is a backup and recovery solution for VMware virtual machines. It is fully integrated with vCenter Server and vSphere Web Client, providing easy, disk-based backup of virtual machines. vSphere Data Protection 5.5 is included with VMware vSphere 5.5 Essentials Plus Kit and higher. The following enhancements have been made to vSphere Data Protection 5.5:

- Direct-to-host emergency restore: vSphere Data Protection can be used to restore a virtual machine directly to a vSphere host without the need for vCenter Server and vSphere Web Client. This is especially helpful when using vSphere Data Protection to protect vCenter Server.
- Backup and restore of individual virtual machine hard disks (.vmdk files): Individual .vmdk files can be selected for backup and restore operations.
- Replication to EMC Avamar: vSphere Data Protection replicates backup data to EMC Avamar to provide offsite backup data storage for disaster recovery.
- Flexible storage placement: When deploying vSphere Data Protection, separate datastores can be selected for the OS partition and backup data partition of the virtual appliance.
- Mounting of existing backup data storage to new appliance: An existing vSphere Data Protection backup data partition can be mounted to a new vSphere Data Protection virtual appliance during deployment.
- Scheduling granularity: Backup and replication jobs can be scheduled at specific times; for example, Backup Job 1 at 8:45 p.m., Backup Job 2 at 11:30 p.m., and Replication Job 1 at 2:15 a.m.

## vSphere Big Data Extensions

vSphere Big Data Extensions (BDE) is a new addition in vSphere 5.5 for VMware vSphere Enterprise Edition™ and VMware vSphere Enterprise Plus Edition™. BDE is a tool that enables administrators to deploy and manage Hadoop clusters on vSphere from a familiar vSphere Web Client interface. It simplifies the provisioning of the infrastructure and software services required for multinode Hadoop clusters. BDE is based on technology from Project Serengeti, the VMware open-source virtual Hadoop management tool.

BDE is available as a plug-in for the vSphere Web Client. Administrators can deploy virtual Hadoop clusters through BDE, customizing variables such as number of Hadoop nodes in the cluster, size of Hadoop virtual machines, and choice of local or shared storage. BDE supports the deployment of all major Hadoop distributions, as well as ecosystem components such as Apache Pig, Apache Hive and Apache HBase.

BDE performs the following functions on the virtual Hadoop clusters it manages:

- Creates, deletes, starts, stops and resizes clusters
- Controls resource usage of Hadoop clusters
- Specifies physical server topology information
- Manages the Hadoop distributions available to BDE users
- Automatically scales clusters based on available resources and in response to other workloads on the vSphere cluster

Using BDE, administrators can provide multiple tenants with elastic, virtual Hadoop clusters that scale as needed to share resources efficiently. Another benefit of Hadoop on vSphere is that critical services in these Hadoop clusters can be protected easily using vSphere HA and VMware vSphere Fault Tolerance (vSphere FT). BDE offers ease of management and operational simplicity by automating many of these tasks for virtual Hadoop clusters.

# vSphere Storage Enhancements

## Support for 62TB VMDK

VMware is increasing the maximum size of a virtual machine disk file (VMDK) in vSphere 5.5. The previous limit was 2TB—512 bytes. The new limit is 62TB. The maximum size of a virtual Raw Device Mapping (RDM) is also increasing, from 2TB—512 bytes to 62TB. Virtual machine snapshots also support this new size for delta disks that are created when a snapshot is taken of the virtual machine.

This new size meets the scalability requirements of all application types running in virtual machines.

## MSCS Updates

Microsoft Cluster Service (MSCS) continues to be deployed in virtual machines for application availability purposes. VMware is introducing a number of additional features to continue supporting customers that implement this application in their vSphere environments. In vSphere 5.5, VMware supports the following features related to MSCS:

- Microsoft Windows 2012
- Round-robin path policy for shared storage
- iSCSI protocol for shared storage
- Fibre Channel over Ethernet (FCoE) protocol for shared storage

Historically, shared storage was supported in MSCS environments only if the protocol used was Fibre Channel (FC). With the vSphere 5.5 release, this restriction has been relaxed to include support for FCoE and iSCSI.

With regard to the introduction of round-robin support, a number of changes were made concerning the SCSI locking mechanism used by MSCS when a failover of services occurs. To facilitate this new path policy, changes have been implemented that make it irrelevant which path is used to place the SCSI reservation; any path can free the reservation.

## 16GB E2E Support

In vSphere 5.0, VMware introduced support for 16Gb FC HBAs. However these HBAs were throttled down to work at 8Gb. In vSphere 5.1, VMware introduced support to run these HBAs at 16Gb. However, there is no support for full, end-to-end 16Gb connectivity from host to array. To get full bandwidth, a number of 8Gb connections must be created from the switch to the storage array.

In vSphere 5.5, VMware introduces 16Gb end-to-end FC support. Both the HBAs and array controllers can run at 16Gb as long as the FC switch between the initiator and target supports it.

## PDL AutoRemove

Permanent device loss (PDL) is a situation that can occur when a disk device either fails or is removed from the vSphere host in an uncontrolled fashion. PDL detects if a disk device has been permanently removed—that is, the device will not return—based on SCSI sense codes. When the device enters this PDL state, the vSphere host can take action to prevent directing any further, unnecessary I/O to this device. This alleviates other conditions that might arise on the host as a result of this unnecessary I/O. With vSphere 5.5, a new feature called PDL AutoRemove is introduced. This feature automatically removes a device from a host when it enters a PDL state. Because vSphere hosts have a limit of 255 disk devices per host, a device that is in a PDL state can no longer accept I/O but can still occupy one of the available disk device spaces. Therefore, it is better to remove the device from the host.

PDL AutoRemove occurs only if there are no open handles left on the device. The auto-remove takes place when the last handle on the device closes. If the device recovers, or if it is readded after having been inadvertently removed, it will be treated as a new device.

## vSphere Replication Interoperability

In vSphere 5.0, there were interoperability concerns with VMware vSphere Replication and VMware vSphere Storage vMotion®, as well as with VMware vSphere Storage DRS™. There were considerations to be made at both the primary site and the replica site.

At the primary site, because of how vSphere Replication works, there are two separate cases of support for vSphere Storage vMotion and vSphere Storage DRS to be considered:

- Moving a subset of the virtual machine's disks
- Moving the virtual machine's home directory

This works fine in the first case—moving a subset of the virtual machine's disks with vSphere Storage vMotion or vSphere Storage DRS. From the vSphere Replication perspective, the vSphere Storage vMotion migration is a "fast suspend/resume" operation, which vSphere Replication handles well.

The second case—a vSphere Storage vMotion migration of a virtual machine's home directory—creates the issue with primary site migrations. In this case, the vSphere Replication persistent state files (.psf) are deleted rather than migrated. vSphere Replication detects this as a power-off operation, followed by a power-on of the virtual machine without the ".psf" files. This triggers a vSphere Replication "full sync," wherein the disk contents are read and checksummed on each side, a fairly expensive and time-consuming task. vSphere 5.5 addresses this scenario.

At the primary site, migrations now move the persistent state files that contain pointers to the changed blocks along with the VMDKs in the virtual machine's home directory, thereby removing the need for a full synchronization. This means that replicated virtual machines can now be moved between datastores, by vSphere Storage vMotion or vSphere Storage DRS, without incurring a penalty on the replication. The retention of the .psf means that the virtual machine can be brought to the new datastore or directory while retaining its current replication data and can continue with the procedure and with the "fast suspend/resume" operation of moving an individual VMDK.

At the replica site, the interaction is less complicated because vSphere Storage vMotion is not supported for the replicated disks. vSphere Storage DRS cannot detect the replica disks: They are simply "disks"—there is no "virtual machine." While the .vmx file describing the virtual machine is there, the replicated disks are not actually attached until test or failover occurs. Therefore, vSphere Storage DRS cannot move these disks because it only detects registered virtual machines. This means that there are no low-level interoperability problems, but there is a high-level one because it is preferable that vSphere Storage DRS detect the replica disks and be able to move them out of the way if a datastore is filling up at the replica site. This scenario remains the same in the vSphere 5.5 release. With vSphere Replication, moving the target virtual machines is accomplished by manually pausing—not "stopping," which deletes the replica VMDK—replication; cloning the VMDK, using VMware vSphere Command-Line Interface, into another directory; manually reconfiguring vSphere Replication to point to the new target; waiting for it to complete a full sync; and then deleting the old replica files.

## vSphere Replication Multi-Point-in-Time (MPIT) Snapshot Retention

vSphere Replication through vSphere 5.1 worked by creating a redo log on the disk at the target location. When a replication was taking place, the vSphere Replication appliance received the changed blocks from the source host and immediately wrote them to the redo log on the target disk.

Because any given replication has a fixed size according to the number of changed blocks, vSphere Replication could determine when the complete replication bundle (the "lightweight delta") had been received. Only then did it commit the redo log to the target VMDK file.

vSphere Replication then retained the most recent redo log as a snapshot, which would be automatically committed during a failover. This snapshot was retained in case of error during the commit; this would ensure that during crash or corruption, there was always a "last-known good snapshot" ready to be committed or recommitted. This prevents finding only corrupted data when recovering a virtual machine.

Historically, the snapshot was retained but the redo log was discarded. Each new replication overwrote the previous redo log, and each commit of the redo log overwrote the active snapshot. The recoverable point in time was always the most recent complete replication.

A new feature is introduced in vSphere 5.5 that enables retention of historical points in time. The old redo logs are not discarded; instead, they are retained and cleaned up on a schedule according to the MPIT retention policy.

For example, if the MPIT retention policy dictates that 24 snapshots must be kept over a one-day period, vSphere Replication retains 24 snapshots. If there is a 1-hour recovery-point objective (RPO) set for replication, vSphere Replication likely retains every replication during the day, because roughly 24 replicas will be made during that day.

If, however, a 15-minute RPO is set, approximately 96 replications will take place over a 24-hour period, thereby creating many more snapshots than are required for retention. On the basis of the retention policy cycle (for example, hourly—24 retained per day), vSphere Replication scans through the retained snapshots and discards those deemed unnecessary. If it finds four snapshots per hour (on a 15-minute RPO) but is retaining only one per hour (24-per-day retention policy), it retains the earliest replica snapshot in the retention cycle and discards the rest.

The most recent complete snapshot is always retained, to provide the most up-to-date data available for failover. This most recent complete point in time is always used for failover; there is no way to select an earlier point in time for failover. At the time of failover, the replicated VMDK is attached to the virtual machine within the replicated vmx, and the virtual machine is powered on. After failover, an administrator opens the snapshot manager for that virtual machine and selects from the retained historical points in time, as with any other snapshot.

## Additional vSphere 5.5 Storage Feature Enhancements

### VAAI UNMAP Improvements

vSphere 5.5 introduces a new and simpler VAAI UNMAP/Reclaim command:

```
# esxcli storage vmfs unmap
```

As before, this command creates temporary files and uses UNMAP primitive to inform the array that these blocks in this temporary file can be reclaimed. This enables a correlation between what the array reports as free space on a thin-provisioned datastore and what vSphere reports as free space. Previously, there was a mismatch between the host and the storage regarding the reporting of free space on thin-provisioned datastores.

There are two major enhancements in vSphere 5.5: the ability to specify the reclaim size in blocks rather than as a percentage value; dead space can now be reclaimed in increments rather than all at once.

### VMFS Heap Improvements

In previous versions of vSphere, there was an issue with VMware vSphere VMFS heap: There were concerns when accessing open files of more than 30TB from a single vSphere host. vSphere 5.0 p5 and vSphere 5.1 Update 1 introduced a larger heap size to confront this. In vSphere 5.5, VMware introduces a much improved heap eviction process, so there is no need for the larger heap size, which consumes memory. vSphere 5.5, with a maximum of 256MB of heap, enables vSphere hosts to access all address space of a 64TB VMFS.

### vSphere Flash Read Cache

vSphere 5.5 introduces a new storage solution called vSphere Flash Read Cache, a new Flash-based storage solution that is fully integrated with vSphere. Its design is based on a framework that enables the virtualization and management of local Flash-based devices in vSphere.

vSphere Flash Read Cache framework design is based on two major components:

- vSphere Flash Read Cache infrastructure
- vSphere Flash Read Cache software

vSphere Flash Read Cache enables the pooling of multiple Flash-based devices into a single consumable vSphere construct called vSphere Flash Resource, which is consumed and managed in the same way as CPU and memory are done today in vSphere.

The vSphere Flash Read Cache infrastructure is responsible for integrating the vSphere hosts' locally attached Flash-based devices into the vSphere storage stack. This integration delivers a Flash management platform that enables the pooling of Flash-based devices into a vSphere Flash Resource.

vSphere hosts consume the vSphere Flash Resource as vSphere Flash Swap Cache, which replaces the Swap to SSD feature previously introduced with vSphere 5.0.

The vSphere Flash Read Cache software is natively built into the core vSphere ESXi Hypervisor. vSphere Flash Read Cache provides a write-through cache mode that enhances virtual machines' performance without the modification of applications and OSs. Virtual machines cannot detect the described performance and the allocation of vSphere Flash Read Cache.

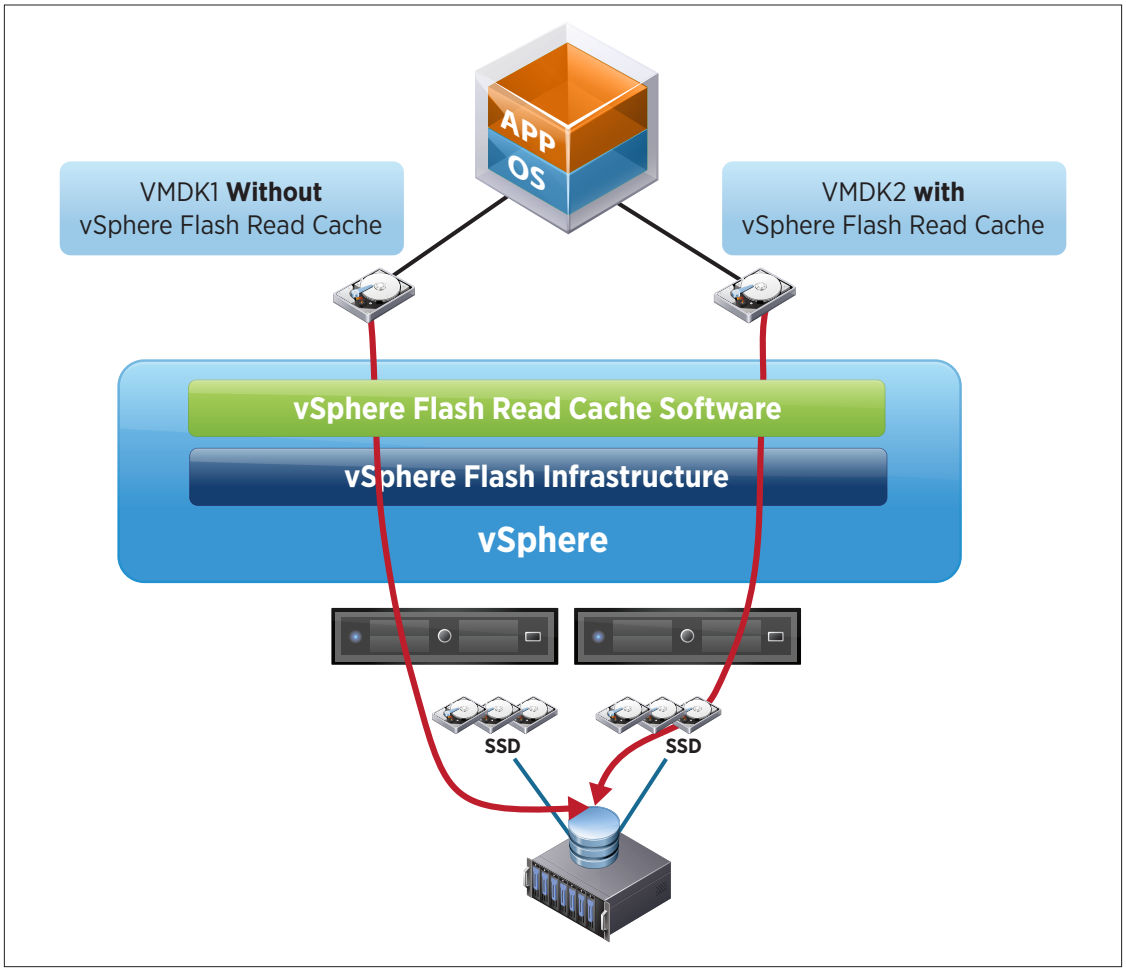


**Figure 8.**

The performance enhancements are introduced to virtual machines based on the placement of the vSphere Flash Read Cache, which is situated directly in the virtual machine's virtual disk data path. vSphere Flash Read Cache enhances virtual machine performance by accelerating read-intensive workloads in vSphere environments.

The tight integration of vSphere Flash Read Cache with vSphere 5.5 also delivers support and compatibility with vSphere Enterprise Edition features such as vSphere vMotion, vSphere HA and vSphere DRS.

# vSphere Networking Enhancements

vSphere 5.5 introduces some key networking enhancements and capabilities to further simplify operations, improve performance and provide security in virtual networks. VMware vSphere Distributed Switch™ is a centrally managed, datacenter-wide switch that provides advanced networking features on the vSphere platform. Having one virtual switch across the entire vSphere environment greatly simplifies management. The following are some of the key benefits of the features in this release:

- The enhanced link aggregation feature provides choice in hashing algorithms and also increases the limit on number of link aggregation groups.
- Additional port security is enabled through traffic filtering support.
- Prioritizing traffic at layer 3 increases quality of service support.
- A packet-capture tool provides monitoring at the various layers of the virtual switching stack.
- Other enhancements include improved single-root I/O virtualization (SR-IOV) support and 40GB NIC support.

## Link Aggregation Control Protocol (LACP) Enhancements

In vSphere 5.1, LACP is supported. LACP is a standards-based method to control the bundling of several physical network links together to form a logical channel for increased bandwidth and redundancy purposes. It dynamically negotiates link aggregation parameters such as hashing algorithms, number of uplinks, and so on, across vSphere Distributed Switch and physical access layer switches. In case of any link failures or cabling mistakes, LACP automatically renegotiates parameters across the two switches. This reduces the manual intervention required to debug cabling issues.

The following key enhancements are available on vSphere Distributed Switch with vSphere 5.5:

- Comprehensive load-balancing algorithm support – 22 new hashing algorithm options are available. For example, source and destination IP address and VLAN field can be used as the input for the hashing algorithm.
- Support for multiple link aggregation groups (LAGs) – 64 LAGs per host and 64 LAGs per VMware vSphere VDS.
- Because LACP configuration is applied per host, this can be very time consuming for large deployments. In this release, new workflows to configure LACP across a large number of hosts are made available through templates.

Figure 9 shows a deployment in which a vSphere host has four uplinks, and those uplinks are connected to the two physical switches. By combining two uplinks on the physical and virtual switch, LAGs are created. The LACP configuration on the vSphere host is performed on the VDS and the port groups.

First, the LAGs and the associated uplinks are configured on the VDS. Then, the port groups are configured to use those LAGs. In this example, the green port group is configured with LAG1; the yellow port group is configured with LAG2. All the traffic from virtual machines connected to the green port group follow the LAG1 path.

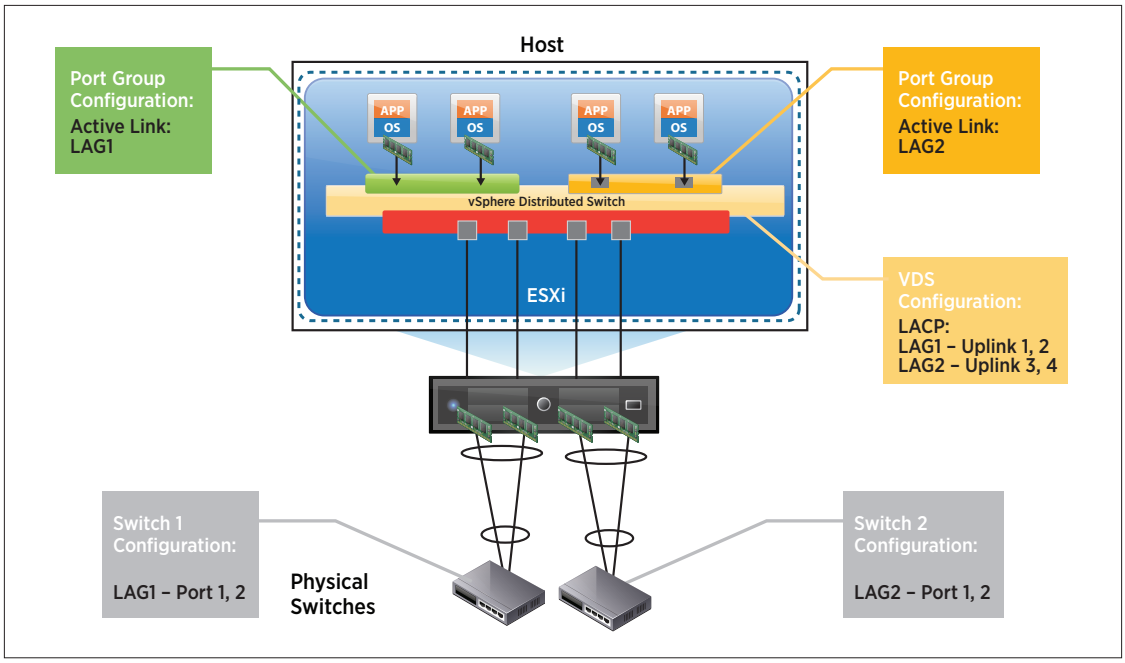


**Figure 9.** LACP Example – Two LAGs

## Traffic Filtering

Traffic filtering is the ability to filter packets based on the various parameters of the packet header. This capability is also referred to as access control lists (ACLs), and it is used to provide port-level security.

The VDS supports packet classification, based on the following three different types of qualifiers:

- MAC SA and DA qualifiers
- System traffic qualifiers – vSphere vMotion, vSphere management, vSphere FT, and so on
- IP qualifiers – Protocol type, IP SA, IP DA, and port number

After the qualifier has been selected and packets have been classified, users have the option to either filter or tag those packets.

When the classified packets have been selected for filtering, users have the option to filter ingress, egress, or traffic in both directions.

As shown in Figure 10, the traffic-filtering configuration is at the port group level.

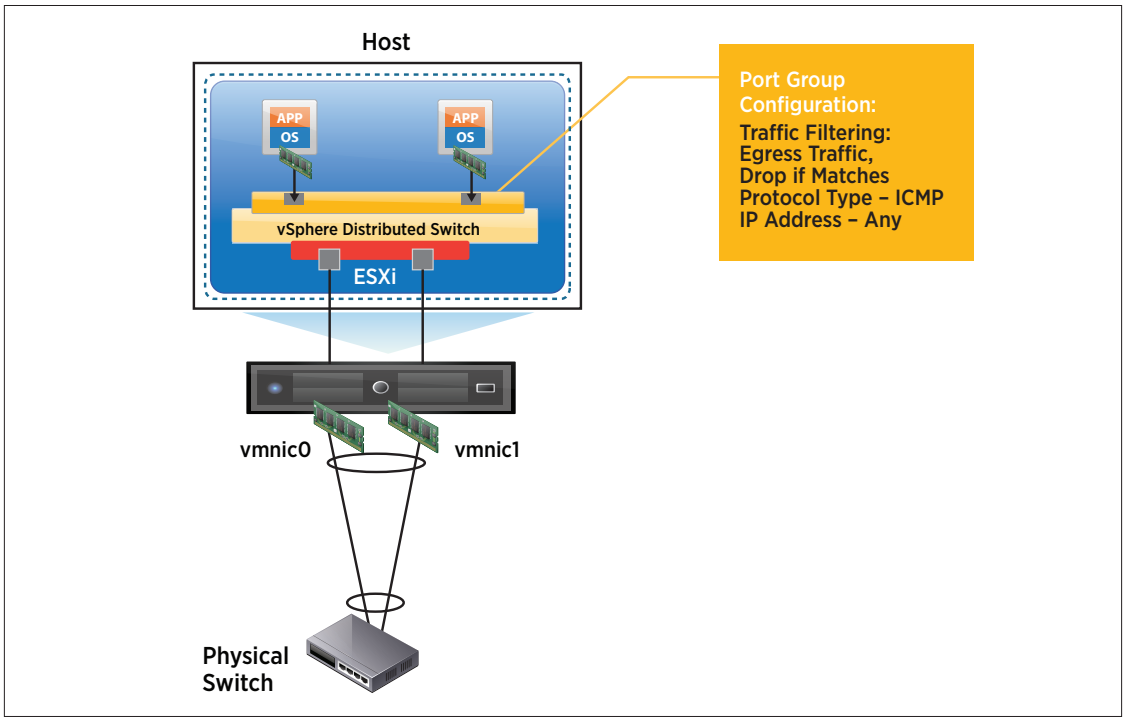


**Figure 10.** Traffic Filtering

## Quality of Service Tagging

Two types of Quality of Service (QoS) marking/tagging common in networking are 802.1p Class of Service (CoS), applied on Ethernet/layer 2 packets, and Differentiated Service Code Point (DSCP), applied on IP packets. The physical network devices use these tags to identify important traffic types and provide QoS based on the value of the tag. Because business-critical and latency-sensitive applications are virtualized and are run in parallel with other applications on an ESXi host, it is important to enable the traffic management and tagging features on VDS.

The traffic management feature on VDS helps reserve bandwidth for important traffic types, and the tagging feature enables the external physical network to detect the level of importance of each traffic type. It is a best practice to tag the traffic near the source and help achieve end-to-end QoS. During network congestion scenarios, the highly tagged traffic doesn't get dropped, providing the traffic type with higher QoS.

VMware has supported 802.1p tagging on VDS since vSphere 5.1. The 802.1p tag is inserted in the Ethernet header before the packet is sent out on the physical network. In vSphere 5.5, the DSCP marking support enables users to insert tags in the IP header. IP header–level tagging helps in layer 3 environments, where physical routers function better with an IP header tag than with an Ethernet header tag.

After the packets are classified based on the qualifiers described in the "Traffic Filtering" section, users can choose to perform Ethernet (layer 2) or IP (layer 3) header–level marking. The markings can be configured at the port group level.

## SR-IOV Enhancements

Single-root I/O virtualization (SR-IOV) is a standard that enables one PCI Express (PCIe) adapter to be presented as multiple, separate logical devices to virtual machines. In this release, the workflow of configuring the SR-IOV–enabled physical NICs is simplified. Also, a new capability is introduced that enables users to communicate the port group properties defined on the vSphere standard switch (VSS) or VDS to the virtual functions.

The new control path through VSS and VDS communicates the port group–specific properties to the virtual functions. For example, if promiscuous mode is enabled in a port group, that configuration is then passed to virtual functions, and the virtual machines connected to the port group will receive traffic from other virtual machines.

## Enhanced Host-Level Packet Capture

Troubleshooting any network issue requires various sets of tools. In the vSphere environment, the VDS provides standard monitoring and troubleshooting tools, including NetFlow, Switched Port Analyzer (SPAN), Remote Switched Port Analyzer (RSPAN) and Encapsulated Remote Switched Port Analyzer (ERSPAN). In this release, an enhanced host-level packet capture tool is introduced. The packet capture tool is equivalent to the command-line tcpdump tool available on the Linux platform.

The following are some of the key capabilities of the packet capture tool:

- Available as part of the vSphere platform and can be accessed through the vSphere host command prompt
- Can capture traffic on VSS and VDS
- Captures packets at the following levels
  - Uplink
  - Virtual switch port
  - vNIC
- Can capture dropped packets
- Can trace the path of a packet with time stamp details

### 40GB NIC Support

Support for 40GB NICs on the vSphere platform enables users to take advantage of higher bandwidth pipes to the servers. In this release, the functionality is delivered via Mellanox ConnextX-3 VPI adapters configured in Ethernet mode.

# Conclusion

VMware vSphere 5.5 introduces many new features and enhancements that further extend the core capabilities of the vSphere platform. The core vSphere ESXi Hypervisor enhancements in vSphere 5.5 include the following:

- Hot-pluggable SSD PCIe devices
- Support for Reliable Memory Technology
- Enhancements to CPU C-states

Along with the core vSphere ESXi Hypervisor improvements, vSphere 5.5 provides the following virtual machine–related enhancements:

- Virtual machine compatibility with VMware ESXi 5.5
- Expanded virtual graphics support to include added support for an additional hardware-accelerated graphics vendor
- Graphic acceleration support for Linux guest operating systems

In addition, the following vCenter Server enhancements include:

- vCenter Single Sign-On Server security enhancements
- vSphere Web Client platform support and UI improvements
- vCenter Server Appliance configuration maximum increases
- Simplified vSphere App HA application monitoring
- vSphere DRS virtual machine–virtual machine affinity rule enhancements
- vSphere Big Data Extensions, a new feature that deploys and manages Hadoop clusters on vSphere from within vCenter

vSphere 5.5 also includes the following storage-related enhancements:

- Support for 62TB VMDK
- MSCS updates
- vSphere 5.1 enhancements
- 16GB E2E support
- PDL AutoRemove
- vSphere Replication interoperability and multi-point-in-time snapshot retention

vSphere 5.5 also introduces the following networking-related enhancements:

- Improved LACP capabilities
- Traffic filtering
- Quality of Service tagging

# About the Authors

Vyenkatesh (Venky) Deshpande is a senior technical marketing manager at VMware. His focus is on networking aspects of the vSphere platform and VMware vCloud® Networking and Security™ product. Venky blogs on the VMware vSphere Blog at http://blogs.vmware.com/vsphere/networking. Follow Venky on Twitter @VMWNetworking.

Cormac Hogan is a senior technical marketing architect within the Cloud Infrastructure Product Marketing group at VMware. He is responsible for storage in general, with a focus on core VMware vSphere storage technologies and virtual storage, including the VMware vSphere Storage Appliance. He has been with VMware since 2005 and in technical marketing since 2011. Cormac blogs on the VMware vSphere Blog at http://blogs.vmware.com/vsphere/storage. Follow Cormac on Twitter @VMwareStorage.

Jeff Hunter is a senior technical marketing manager at VMware, focusing on IT business continuity and disaster recovery. Jeff has been with VMware since 2007. Prior to VMware, Jeff spent several years in a systems engineer role, expanding the virtual infrastructures at a regional bank and a Fortune 500 insurance company. Jeff blogs on the VMware vSphere Blog at http://blogs.vmware.com/vsphere/uptime. Follow Jeff on Twitter @jhuntervmware.

Justin King has been involved in the IT industry for more than 15 years. He has had various roles and responsibilities, ranging from administration to architecting solutions. Since joining VMware in 2009, Justin has supported sales teams as a sales engineer and evangelized BCDR technologies. Currently, he is part of the Technical Marketing team, focusing on vCenter Server. Justin blogs on the VMware vSphere Blog at http://blogs.vmware.com/vsphere/vcenter-server. Follow Justin on Twitter @vCenterGuy.

William Lam is a senior technical marketing engineer in the Cloud Infrastructure Product Marketing group at VMware. William currently focuses on automation for both the vSphere and VMware vCloud Director® platform APIs and CLIs. Previous to VMware, he was a systems engineer, managing a large vSphere installation and UNIX/Linux systems. William blogs on the VMware vSphere Blog at http://blogs.vmware.com/vsphere/automation. Follow William on Twitter @lamw.

Ken Werneburg is senior technical marketing manager at VMware for business continuity and disaster recovery solutions. Ken blogs on the VMware vSphere Blog at http://blogs.vmware.com/vsphere/uptime. Follow Ken on Twitter @vmKen.

Rawlinson Rivera is a senior technical marketing manager within the Cloud Infrastructure Product Marketing group at VMware. He is responsible for storage in general, with a focus on VMware storage virtualization technologies, including the VMware vSphere Flash Read Cache, VMware virtual SAN, VMware Virsto, and vSphere Storage Appliance. Rawlinson blogs on the VMware vSphere Blog at http://blogs.vmware.com/vsphere/storage and http://www.punchingclouds.com. Follow Rawlinson on Twitter @PunchingClouds.

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

 Item No: VMW-WP-vSPHR-5.5-PLTFRM-A4-101 Docsource: OIC - 13VM004.05