

VMware Best Practice and Integration Guide



INTRODUCTION

Today's Data Centers are embracing Server Virtualization as a means to optimize hardware resources, energy resources, and real estate. In fact, most Data Centers, large and small, have already moved the majority of non-critical applications to virtualized environments. The next step in this evolution towards Virtualization is to migrate mission critical, tier one applications to the Virtual Machine. However, in order to gain the trust of IT departments, the virtualization environment must offer the features of a 5 9's High Availability infrastructure and provide guarantees of scalability, performance and reliability. In order to accomplish this, all of the components of the infrastructure must meet these new demands.

VMware's flagship vSphere virtualization product is ready to manage mission critical applications. It provides world class high availability features and resource management tools that allow customers to give priority to tier one applications. To complement these distributed virtualized computing environments, Dot Hill Systems offers networked storage solutions that provide equivalent high availability and reliability features for the underlying storage subsystem.

Dot Hill Systems is an active partner with VMware in bringing comprehensive virtualized solutions to customers. The Dot Hill AssuredSAN 3000 Series Storage Arrays are designed to be integral components in demanding High Availability computing environments, delivering data to servers from a robust, reliable and high performance platform.

This whitepaper will outline the advantages of using Dot Hill AssuredSAN 3000 Series storage products in VMware environments. To begin with, this guide will highlight elements of the strategic partnership between Dot Hill and VMware, including the advantages gained through product integration and certification. Implementation and tuning of disparate components in a system can sometimes prove daunting. This guide includes a section that provides a targeted list of best practices and configuration tips specific to VMware configurations. Finally, this guide will demonstrate the performance characteristics of AssuredSAN storage in VMware environments, based on actual test results from application specific test tools.

DOT HILL AND VMWARE'S STRATEGIC PARTNERSHIP

Dot Hill Systems maintains a strategic partnership with VMware through the VMware Technology Alliance Partnership (TAP) program. As an active VMware partner, Dot Hill actively integrates VMware specific features into its AssuredSAN storage products, and ensures that AssuredSAN systems integrate seamlessly into VMware environments.

The primary means used to ensure seamless integration is the participation in the Hardware Certification Program at VMware. All AssuredSAN products are subjected to compatibility tests defined by VMware for each VMware vSphere version, and are added to VMware's Hardware Compatibility List when the storage system passes this battery of tests. These processes and programs help to ensure the integrity and quality of the overall solution. Dot Hill ensures that all AssuredSAN products are certified with the latest releases of VMware vSphere.

The Dot Hill AssuredSAN family of products are designed for high availability computing environments, with redundant, active/active processing components, RAID protected storage, multiple data paths, and built-in snapshot and replication features. In addition, AssuredSAN systems offer an assortment of drive and interface options, providing custom solutions to a wide variety of applications. Finally, the AssuredSAN storage system is architected for speed. Every system is optimized for high performance delivery of data regardless of the workload. Figure 1 shows a few examples of base configurations of AssuredSAN 3000 Series Storage Systems.



FIGURE 1
Dot Hill AssuredSAN 3000 Series Storage Products

In addition to the quality features built into the AssuredSAN storage array, the VMware vSphere software suite includes many features to enhance the overall storage management solution. These features dovetail and integrate with the AssuredSAN, providing a complete portfolio for customers requiring a robust, scalable, high performance computing and storage environment. Table I outlines many of the common storage related features that complement or are integrated with the AssuredSAN storage systems.

TABLE I
AssuredSAN Storage Integration with VMware vSphere

Connectivity	✓	All Dot Hill AssuredSAN 3000 Series products are Certified with the latest revisions of VMware vSphere. Dot Hill offers connectivity options with FC, iSCSI and SAS interfaces.
High Availability	✓	VMware supports a clustered server environment with automatic failover. The AssuredSAN 3000 Series Storage Systems offer fully redundant components and storage, providing an ideal storage complement to a VMware cluster. In addition, storage volumes from the AssuredSAN can be presented to multiple hosts, enabling VMware vSphere HA features such as vSphere vMotion.
Multipathing	✓	The AssuredSAN is ALUA compliant and compatible with standard VMware multipathing drivers. This allows all data to be transferred along optimum paths for maximum speed and efficiency.

Thin Provisioning	✓	The AssuredSAN is fully compatible with the Thin Provisioning features of VMware. This feature can be leveraged to help simplify the management of storage resources.
vSphere vMotion	✓	The AssuredSAN is an ideal storage solution for VMware vSphere vMotion, as it allows Virtual Machines to migrate seamlessly between physical machines, while underlying storage remains online.
vSphere Storage vMotion	✓	The AssuredSAN is completely compatible with VMware vSphere Storage vMotion as a storage source or a storage target. Storage is efficiently and quickly migrated to or from the AssuredSAN as the need arises.
Storage I/O Control (SIOC)	✓	The AssuredSAN is a high performance solution for many VMware workloads. In cases where I/O needs to be managed to assure bandwidth for particular applications, AssuredSAN is fully compatible with VMware vSphere Storage I/O Control.
vStorage API for Array Integration (VAAI)	✓	The AssuredSAN supports all three of the VAAI SCSI primitives and implements an end-to-end solution via the Dot Hill plug-in for VAAI with vSphere 4.1 and vSphere 5. The VAAI solution with AssuredSAN has been certified with VMware and is listed in the VMware Hardware Compatibility List.
vSphere Site Recovery Manager (SRM)	✓	The AssuredSAN fully supports vSphere Site Recovery Manager via the AssuredRemote replication technology and the Dot Hill Storage Replication Agent (SRA) for vSphere 5.

Finally, a key aspect of the relationship between Dot Hill and VMware is the service and support available to customers to ensure seamless integration of the software and the hardware. Dot Hill is committed to providing world class support and services to resolve issues, assist in implementing best practices and provide proactive maintenance and upgrades to their infrastructure. For more information, visit www.dothill.com.

SAN STORAGE BEST PRACTICES

The Dot Hill Systems AssuredSAN 3000 Series Storage Arrays are ideal data repositories for any VMware deployment. They are especially suited for clustered deployments, where the customer can take advantage of the High Availability features of AssuredSAN, such as redundant components. In addition, distributed, SAN based storage is a key enabler for clustered servers, and in the case of VMware, it enables important features such as vMotion. Configuring a SAN environment for use with VMware is not much different than configuring a SAN for any other type of server. This section covers some general purpose SAN configuration Best Practices, along with some VMware specific guidelines.

In brief, Best Practice guidelines aim to accomplish two specific goals:

- **Redundancy.** Having no single point of failure is critical in the design of High Availability systems.

- Performance. In many cases, this translates to distributing I/O loads among available resources to realize the highest overall system performance.

To get the most from AssuredSAN storage in a VMware environment, the following Best Practice guidelines have been established. These guidelines should help in deploying a solution that provides optimum performance, reliability and availability.

In addition to these guidelines, VMware has published a Fibre Channel connectivity guide, which expands on some of the topics covered here in a more general fashion. You can refer to the VMware guide at this link:

http://www.vmware.com/pdf/vsphere4/r41/vsp_41_san_cfg.pdf

1) Ensure proper cabling for performance and redundancy

One of the most important, yet challenging aspects of storage network design and implementation is in the cabling. Proper cabling provides redundancy, so that the failure of any one component does not compromise access to any of the storage. At the same time, this redundancy should also provide a mechanism for aggregating throughput, delivering higher performance to the host systems. Following are a few rules and guidelines to follow when connecting components:

- a) Deploy redundant switches. A switch in a network represents a single point of failure. A Best Practice in High Availability systems is to utilize two switches between a storage system and a host.
- b) On each storage controller in the storage system, connect at least one of the host ports to each switch. If a controller has two host port connections, then connect one port to Switch A, and the other port to Switch B.
- c) Each Host server should have at least 2 ports dedicated for connectivity into the SAN. Once again, this eliminates single points of failure in the network. Connect each port to a separate switch.

These principles are demonstrated in the following reference Fibre Channel configuration for AssuredSAN storage connectivity to VMware servers:

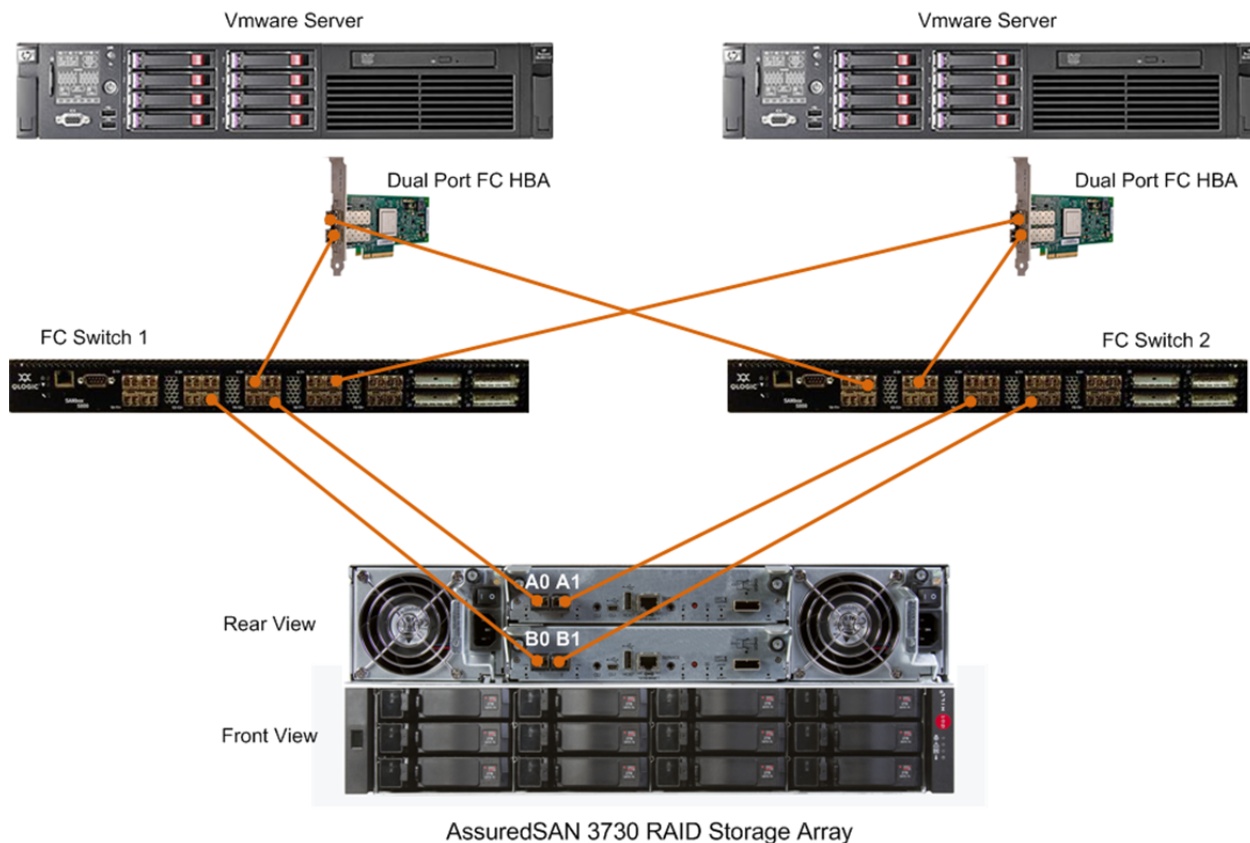


FIGURE 2

SAN Connectivity Best Practice

2) Distribute the Load among Controllers

The DH AssuredSAN 3000 Series RAID Array typically comes with dual Active-Active controllers. With two controllers, it is important to optimize the processing power of both controllers in order to realize the fastest throughput and I/O to the system. To do this, the RAID groups that are defined on the array should be evenly distributed among the two controllers (each RAID group is assigned to an “owning” controller). For example, if there are 12 drives in the array, then it makes sense to create two RAID groups of 6 drives each; one RAID group assigned to controller A, and the other RAID group assigned to controller B.

3) Distribute the load among Virtual Machines

As is demonstrated in the test results presented later in this whitepaper, virtualized environments are not optimized for hosting just one Virtual Machine. While this may seem obvious and intuitive, it bears stating that performance is optimized when it is distributed among several Virtual Machines in a system.

4) Understand your Multipathing (MPIO) setup

If the host system is connected to the storage system using the Best Practices listed above (redundant switches, multiple host connections...), then VMware will be presented with multiple paths to each Volume.

In addition, some of the paths will be considered to be Optimized or Preferred, and some of the paths will be considered Non-preferred. The storage system will process I/O on any of the paths, but it will process I/O faster on the preferred paths. The reasons for having different types of paths stems from the fact that RAID groups are assigned an “owning” controller, and all I/O to the RAID group passes through the “owning” controller. Thus, I/O routed directly to the “owning” controller will be processed very quickly. On the other hand, I/O routed on paths that are connected to the alternate controller will ultimately be passed through the “owning” controller for processing, which slows things down. Storage devices that distinguish path types as being preferred or non-preferred are known as ALUA (Asymmetric Logical Unit Access) compliant devices, as defined in the SCSI 3 (SPC-3) protocol.

All Dot Hill AssuredSAN 3000 Series storage products are ALUA compliant devices. They should be recognized by VMware as such and assigned the default multipathing module for ALUA storage devices. This module is known as “VMW_SATP_ALUA”.

The multipathing module selected for the AssuredSAN array can be verified by examining the “manage paths” dialog for the associated devices in the “Configuration->Storage Adapters” section of an ESX server in vCenter. Alternatively, the following command can be issued in the ESX console:

```
esxcli nmp device list
```

The output from this command will list each device path in the system along with associated details. Pay attention to the following entries:

```
Device Display Name: Dot Hill Fibre Channel Disk...
Storage Array Type: VMW_SATP_ALUA
Path Selection Policy: VMW_PSP_MRU
```

The `Device Display Name` will identify the device. The name “Dot Hill” should be in the output string. The `Storage Array Type` identifies the array as ALUA, Active-Active, or Active-Passive. If the policy is not `VMW_SATP_ALUA`, then the policy will need to be specifically set (see below). Finally, the `Path Selection Policy` will identify the load balancing policy. By default, `VMW_PSP_MRU` (Most Recently Used) is selected. This policy may not provide the best load balancing across the available ports. The policy of `VMW_PSP_RR` (Round Robin) may be selected instead to improve load balancing. This policy can be easily modified in vCenter for each path in the system. Alternatively, the default policy for the array can be modified to `VMW_PSP_RR` (see below) so that all paths are automatically configured to the new default of Round Robin, thus eliminating the need for manual adjustments in vCenter.

To setup specific default Storage Array Type settings and Path Selection Policy settings for the AssuredSAN 3000 Series array, the following command must be entered in the ESX console:

```
esxcli nmp satp addrule --vendor="DotHill"
                        --model="DH3000"
                        --claim-option="tpgs_on"
                        --satp="VMW_SATP_ALUA"
                        --psp="VMW_PSP_RR"
                        --description="Dot Hill AssuredSAN 3000 support"
```

NOTE: *The command must be entered on a single line. The command is shown on multiple lines here for readability.*

Here are some points to consider when using this command:

- 1) The server must be re-booted for the change to take effect.
- 2) Each server in a cluster must be modified independently.
- 3) The new entry cannot be modified. To change the entry, issue the “deleterule” command to remove the existing entry first.
- 4) To verify that the new entry is in the list of multipathing rules, issue the following command:

```
esxcli nmp satp listrules
```

Please note that adding this rule is usually NOT necessary. It is only needed if the AssuredSAN 3000 series array is NOT identified as ALUA, or if a *default* policy of Round Robin is desired for all paths.

5) Select an appropriate storage configuration

An important element of planning storage usage involves selection of RAID groups and RAID levels. While each customer scenario is unique and trade-offs are necessary to balance performance with capacity, it is understood that RAID 10 will deliver superior I/O performance in random workloads (such as Databases), whereas RAID 5 will deliver superior performance in sequential workloads. Given the distributed, virtualized nature of VMware environments, most I/O workloads will be random – regardless of the application.

In addition to RAID selection, a few tuning parameters may be applied to optimize performance in specific environments:

- a) **Chunk Size.** A chunk size of 512K has been shown to provide the best performance in standard random and sequential workloads. However, some mixed workloads generate superior results when the chunk size is lowered. Where possible, the customer is encouraged to experiment with a few different chunk sizes to determine which works best in their environment.
- b) **Read-Ahead Cache.** The read-ahead cache provides significant performance gains in purely sequential workloads. However in some mixed workloads, this mechanism may not provide

optimum results. In largely random workloads, the read-ahead cache should be disabled to ensure the best possible performance.

6) Take advantage of Storage I/O Control in vSphere

The Storage I/O Control (SIOC) feature of vSphere provides for policy based distribution of I/O resources. The benefit of this feature is granular control of bandwidth consumption by applications running on Virtual Machines. As an example, one Virtual Machine may be hosting an Oracle database while a second Virtual Machine hosts a file server. Both Virtual Machines utilize the same underlying storage system. Therefore, it is possible for the file server to heavily utilize the storage resources at the same time that the Oracle database is attempting to perform storage operations. This type of contention can compromise the overall performance of critical applications in a virtualized environment. SIOC can alleviate contention by guaranteeing bandwidth to critical applications.

SIOC is a crucial element in virtual server environments where tier one applications need precedence over lower priority systems. Even in more modest deployments, SIOC can provide an element of balance among applications, so that no application is starved for access to storage.

PERFORMANCE CHARACTERIZATION

Customers using AssuredSAN storage in virtual server environments want to have confidence that the solution does not compromise on the performance of data delivery between the storage system and the virtual machine host. In addition to VMware certification processes, Dot Hill has engaged in performance characterization programs to establish the relative benefit or detriment of using AssuredSAN storage with VMware host systems. The following sections highlight the results and analysis of this testing.

METHODOLOGY

The objective of this program was to establish the overall performance of a storage system with a single x64 based host server, both with and without the VMware virtualization software. To accomplish this goal, a host operating system was selected that could readily be installed as a stand-alone OS on the server, as well as function within the Virtual Machine environment of VMware. Along with the selection of the OS, multiple performance test tools were selected to measure data transfer rates on the OS, both in the stand-alone configuration, as well as the Virtual Machine configuration.

The following hardware and software elements were selected for testing:

Hosts (quantity: 2)

HP ProLiant DL380 G6 Server

Two (2) quad core Xeon processors (E5530)

32 GB memory

QLogic QLE2562 dual port 8Gbit Fibre Channel HBA

Storage

Dot Hill AssuredSAN 3730 RAID Array with Dual Active-Active Controllers

TS230 Firmware

48 Disk drives: 450 GB, 15K rpm, 6Gbit, 3.5" SAS

Operating System

Microsoft Windows Server 2008 R2 SP1 Enterprise Edition

Test Tools

Iometer 2008.06.18

Microsoft JetStress 2010

Hypervisor

VMware vSphere 4.1

The following diagram shows the connectivity of these elements for testing:

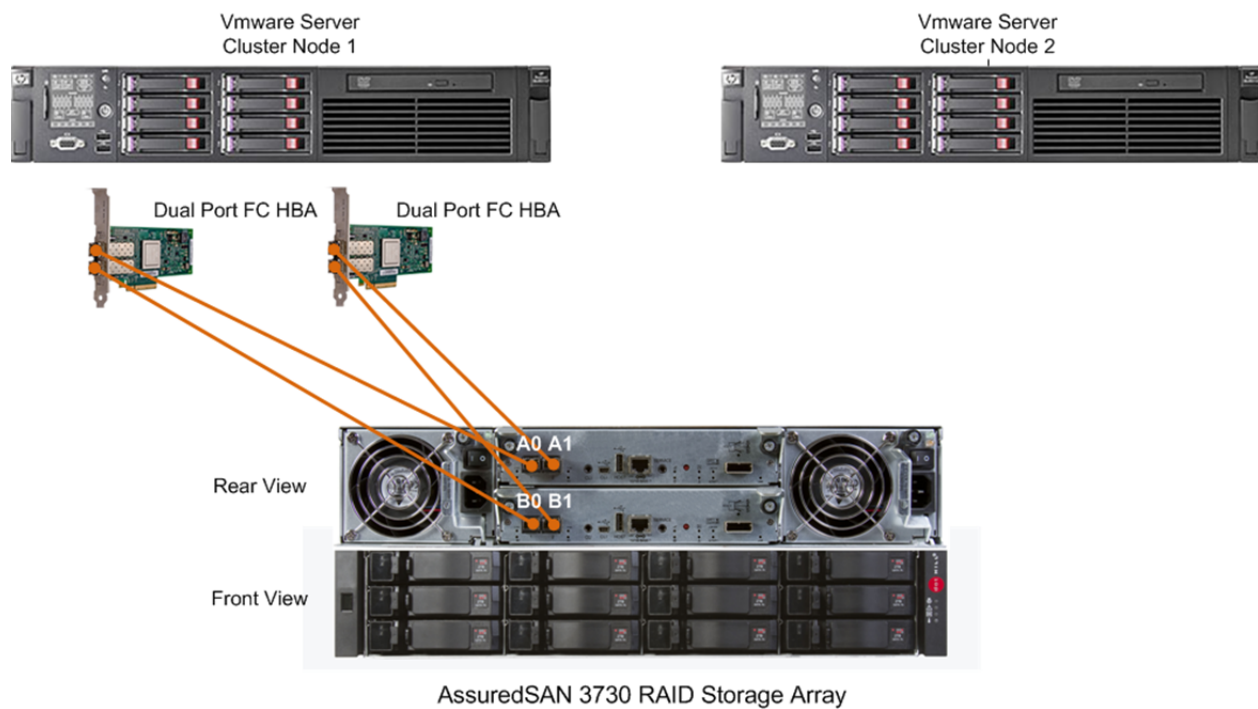


FIGURE 3

Performance Characterization Testing: Components and Connectivity

NOTE: In multi-host tests, one of the FC HBAs is moved from Cluster Node 1 to Cluster Node 2.

Using these hardware and software elements, the following configurations were tested for best data transfer performance:

Stand-Alone	Windows Server loaded onto single bare metal DL380 Server. All storage presented to server via two QLE2562 HBAs.
1 VM, 1 Server	VMware vSphere loaded onto single bare metal DL380 Server. One virtual machine running with Windows Server 2008 loaded on the VM. All storage presented to server via two QLE2562 HBAs.
2 VMs, 1 Server	VMware vSphere loaded onto single bare metal DL380 Server. Two virtual machines running with Windows Server 2008 loaded on each VM. All storage presented to server via two QLE2562 HBAs.
4 VMs, 1 Server	VMware vSphere loaded onto single bare metal DL380 Server. Four virtual machines running with Windows Server 2008 loaded on each VM. All storage presented to server via two QLE2562 HBAs.
2 VMs, 2 Servers	VMware vSphere loaded onto two bare metal DL380 Servers. The two servers are clustered together with vSphere High Availability Services. Two virtual machines running on each server (4 VMs total) with Windows Server 2008 loaded on each VM. Half of the volumes from the storage system are presented to one server; the other half is presented to the other server. Each server has a single QLE2562 HBA for storage connectivity.

In addition to these configurations, various options were used in presenting storage up to the host for testing:

RAID levels: RAID 0, RAID 5, RAID 6, RAID 10

VMware storage presentation: VMFS, RDM

In all cases, the following rules were observed:

- 1) The entire amount of available storage is exercised in each test.
- 2) Four VDisks of 12 drives each are created for each test. Chunk size of 512K.
- 3) Either 8 or 16 Volumes are created for each test, depending upon the amount of available space. The primary rule is to consume all the available space, but keep each Volume under 2 TB in size (a VMFS limitation in vSphere 4.1).
- 4) Storage Volumes are distributed evenly among the hosts under test. For example, if 4 VMs are being tested, and 16 Volumes have been created, then each host will be presented with 4 of the Volumes.

Iometer Specifics

For each configuration tested, the following test profiles were exercised with Iometer using various I/O sizes and queue depths:

- Sequential Read
- Sequential Write
- Random Read
- Random Write

JETSTRESS SPECIFICS

For each configuration test, a “disk subsystem throughput” test was executed for 2 hours. Each volume hosted one database which consumed 100% of the volume.

All jetstress testing was performed using RAID 10 disk groups. This RAID level is commonly used in database configurations due to its superior performance in random transactions.

IOMETER TEST RESULTS

The following charts show representative results from Iometer testing in RAID 5 configurations. While other RAID levels are not presented here, it is important to note that all tests at all RAID levels delivered results consistent with the representative results shown here.

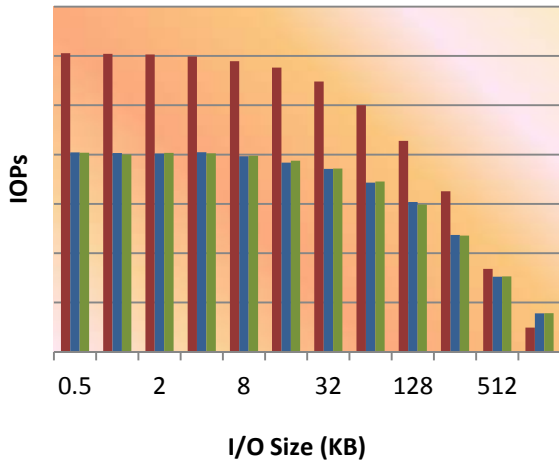
SINGLE MACHINE IOMETER PERFORMANCE

Iometer load is concentrated on a single machine – virtual or stand alone

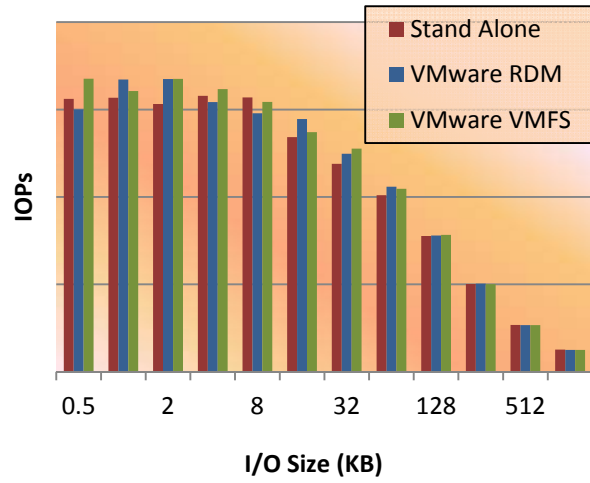
48 Disk Drives: 450 GB, 15K rpm SAS

4 RAID 5 Groups; 16 Volumes

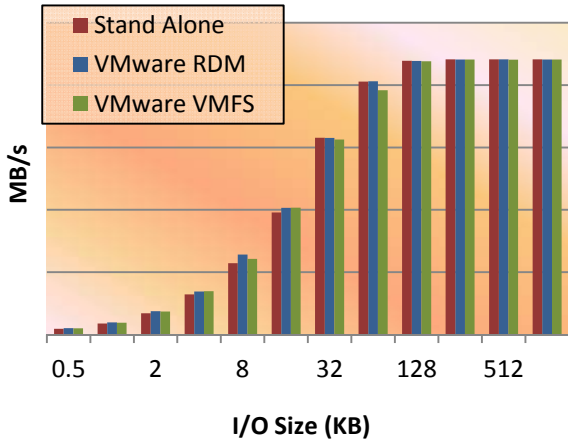
Random Read IOPs



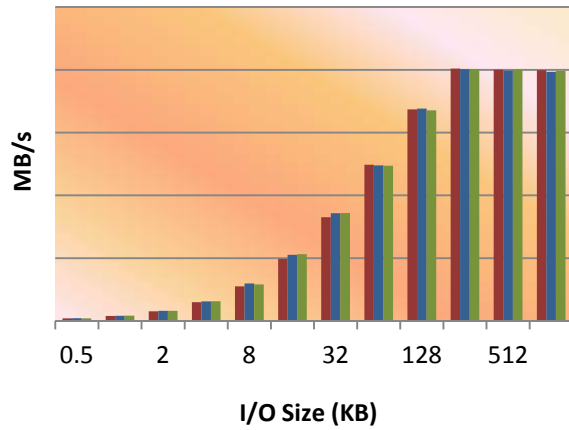
Random Write IOPs



Sequential Read Throughput



Sequential Write Throughput



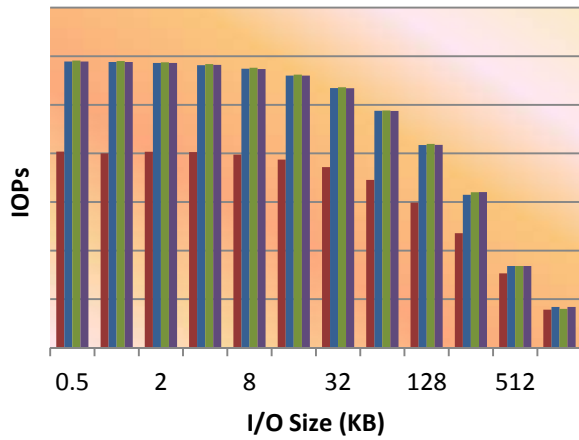
VMFS STORAGE – MULTIPLE VIRTUAL MACHINES

Iometer load is distributed among multiple virtual machines

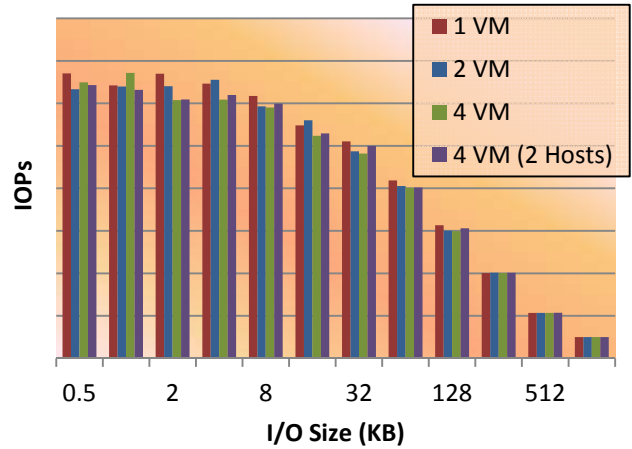
48 Disk Drives: 450 GB, 15K rpm SAS

4 RAID 5 Groups; 16 Volumes

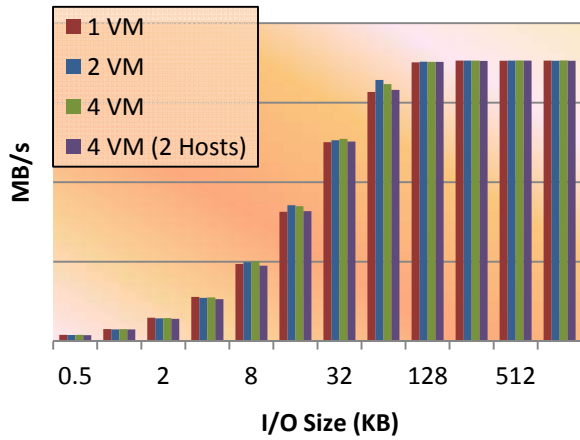
Random Read IOPs



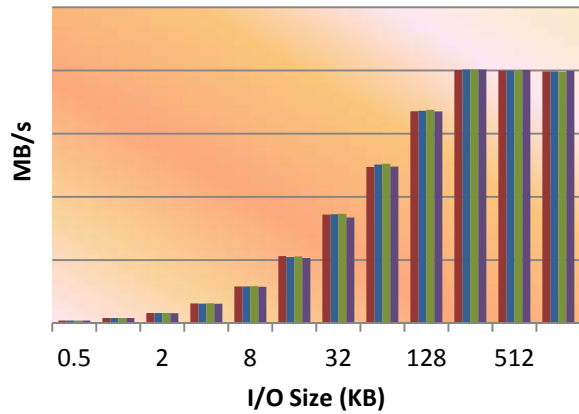
Random Write IOPs



Sequential Read Throughput

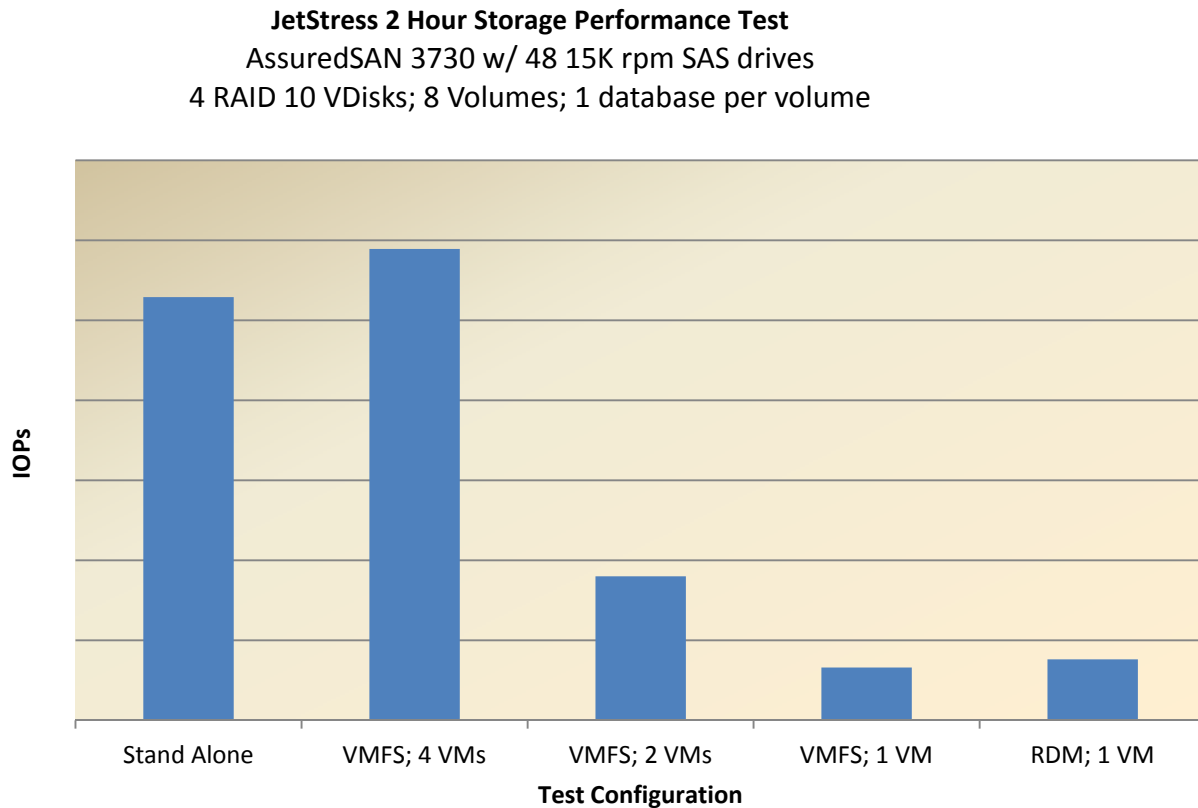


Sequential Write Throughput



JETSTRESS TEST RESULTS

The following chart shows the overall results of JetStress testing. All results were measured using the greatest load on the system that would pass the test (all average latency measurements are below 20 ms).



ANALYSIS

In most workloads and configurations, there is almost no measurable difference between an Operating System running on dedicated physical resources, and an Operating System running within a vSphere 4.1 virtualized environment. However, Iometer testing clearly demonstrates a deficiency in Random Read performance when all of the storage is concentrated on a single VMware virtual machine. In these configurations, Random Read performance can only achieve two thirds of the Random Read I/O of a stand-alone system.

However, when the same tests are run in a VMware vSphere environment where the load is distributed among multiple Virtual Machines, then this discrepancy is negated, and the system is able to deliver nearly the same Random Read performance as a stand-alone system.

It is important to note that the results were consistent across all RAID levels. In addition, no difference was observed in the performance of storage that was presented as Raw Device Map (RDM) or via a Datastore (VMFS).

JetStress test results reinforced the conclusions drawn from Iometer testing: Performance of a single Virtual Machine lags the performance of a stand-alone system. JetStress tests performed on a single Virtual Machine generated only one-tenth of the I/O as compared to stand-alone systems, regardless of whether the storage was presented as RDM or VMFS. However, once the load was spread across multiple Virtual Machines, then the performance increased. In fact, with 4 Virtual Machines running on one physical machine, the JetStress result outperformed a stand-alone server by 10%.

These results reinforce and validate what a customer would expect from a virtualized environment. The I/O system is not designed for a single Virtual Machine. It is designed for many Virtual Machines. When the system is deployed as designed, then I/O performance is optimized and is demonstrated to be equivalent to a stand-alone server environment.

In summary, the following conclusions have been established from this testing:

- 1) Random Read performance of a single vSphere 4.1 Virtual Machine will lag the performance of a stand-alone system.
- 2) In all other workloads, a single vSphere 4.1 Virtual Machine can generate approximately the same I/O performance as a stand-alone system.
- 3) In configurations where storage is distributed to four Virtual Machines, then there are no measurable performance differences to a stand-alone system.

CONCLUSION

The Dot Hill AssuredSAN storage array is an ideal storage solution in VMware computing environments. The AssuredSAN has been demonstrated and certified to work seamlessly with VMware vSphere Servers and associated Virtual Machines. In addition to the inherent benefits of SAN based storage, the AssuredSAN storage array delivers:

- High Performance – 8Gbit FC speeds
- Reliability – redundant components and paths
- High Availability – Clustered controllers with failover technology
- Scalability & Flexibility – a variety of drive options and configurations

The High Availability features offered by VMware and Dot Hill unleash the potential of Virtualized computing for mission critical applications. These solutions provide the means by which applications can execute and migrate among various hardware platforms to ensure availability and meet performance requirements.

This paper has demonstrated how the strategic partnership between Dot Hill and VMware has resulted in a combined solution that meets the needs of today's dynamic Data Centers with High Availability features and scalable architecture. In addition, real world performance characterization has shown that customers can realize the benefits of virtualized applications without sacrificing speed.

© 2012 Copyright Dot Hill Systems Corporation. All rights reserved. Dot Hill Systems Corp., Dot Hill, the DotHill logo, and AssuredSAN are trademarks or registered trademarks of Dot Hill Systems. All other trademarks are the property of their respective companies in the United States and/or other countries.
1.11.12

