Using IBM® FlashSystem V840™ with the VMware® vStorage APIs for Array Integration

Summary: IBM FlashSystem V840 provides integration and performance benefits for VMware vSphere environments through support of the VMware vStorage APIs for Array Integration. The VAAI API functions: hardware-accelerated Block Zero, hardware-assisted Locking, and hardware accelerated Full Copy enable certain vSphere functions to be offloaded from the vSphere host to the VAAI-enabled IBM FlashSystem V840.

This document explores the usage and performance benefits of VAAI with IBM FlashSystem V840.

Introduction to VMware vStorage APIs for Array Integration

The vStorage APIs for Array Integration (VAAI) are a set of APIs available to VMware storage partners which when leveraged allow certain VMware functions to be delegated to the storage array, enhancing performance and reducing load on servers and storage area networks (SAN). The current implementation of VAAI in vSphere 5.5 includes the following functions supported by IBM FlashSystem V840: hardware-accelerated Block Zero, hardware-assisted Locking, and hardware-accelerated Full Copy. These functions each relate to and benefit common VMware operational tasks, as shown in Figure 1.

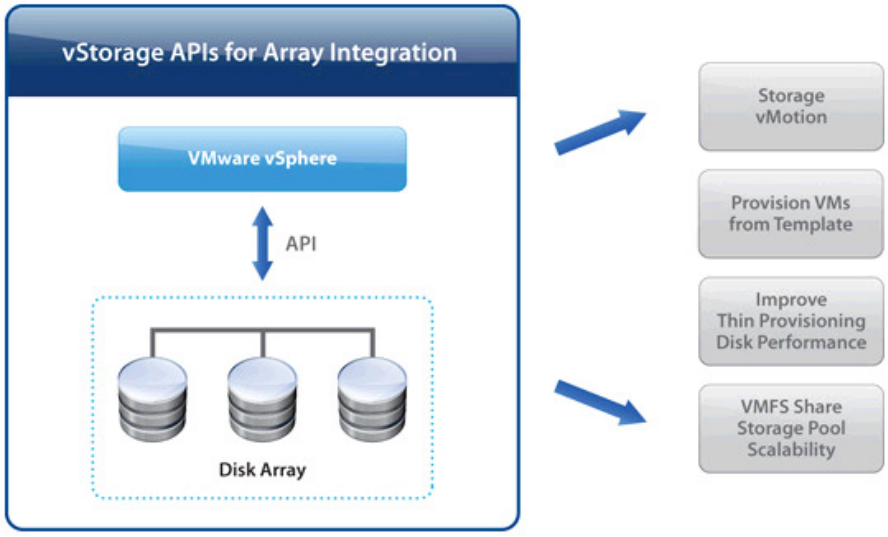


Figure 1 vStorage APIs for Array Integration relationship to VMware functions

Hardware-accelerated Block Zero

Certain operations within VMware vSphere, such as deploying a VMware Fault Tolerance (FT) compliant virtual machine or creating a new fully-allocated virtual disk, require that the .VMDK file be provisioned as eagerzeroedthick. By default, the VMware .VMDK files are provisioned as zeroedthick. In the zeroedthick format, the space for the .VMDK file is fully allocated, but blocks are initialized when they are first accessed. In the eagerzeroedthick format, the .VMDK file is fully allocated and all blocks are initialized immediately. Eagerzeroedthick .VMDK files should be deployed to ensure maximum throughput can be achieved by the virtual machine.

When provisioning an eagerzeroedthick .VMDK file, SCSI commands are issued from the host to write zeros and initialize disk blocks. These requests consume host CPU cycles and HBA queue slots, and consume bandwidth on the SAN. ESXi hosts utilizing IBM FlashSystem V840 storage can be configured to use hardware-accelerated Block Zero, which offloads the initialization of disk blocks to the IBM FlashSystem V840 array. This offloaded task will initialize large numbers of disk blocks very quickly

without transferring data over the SAN or consuming ESXi host resources. The following VMware provisioning tasks benefit from Block Zero:

- Cloning operations which include eagerzerodthick .VMDK files
- Allocating new file blocks for thin-provisioned .VMDK files
- Initializing previous unwritten file blocks for zeroedthick .VMDK files

Hardware-accelerated Full Copy

When a new virtual machine is created from cloning a template or virtual machine, data is read from the source virtual machine or template up through the running ESXi host and back down to the destination. Hardware-accelerated Full Copy allows the operations to be run on IBM FlashSystem V840 and not on the ESXi host. The ESXi host initiates and tracks the progress of the task, while IBM FlashSystem V840 performs the work.

In addition to cloning, Full Copy is also engaged during a Storage VMotion operation. When a task is created to migrate a virtual machine, the operation runs completely on IBM FlashSystem V840 as long as the source and destination VMFS data stores both reside there.

Full Copy greatly reduces the SAN traffic required for performing cloning and migration operations, while also saving CPU cycles and HBA queue slots on the ESXi hosts. In addition to these efficiencies, cloning and migration tasks run with Full Copy are significantly faster than standard tasks.

Hardware-assisted Locking

Many vSphere operations such as virtual machine power on, snapshot creation and deletion, VMotion, and many others require that a SCSI reservation be placed on the Virtual Machine File System (VMFS) volume from the initiating host. This lock prevents other host servers from accessing the logical unit number (LUN). This is a preventative measure used to prevent virtual machines from being modified by more than one ESX host at a time and to protect the VMFS metadata. However, it can also result in reduced performance when VMFS volumes with a high number of virtual machines are used. The hardware-assisted Locking function eliminates this problem by providing a more granular way to control access to the VMFS metadata. Rather than a lock being placed by a host on the entire LUN, hardware-assisted Locking enables the storage subsystem to place a lock on a single block of data. This enhances performance in many ways and significantly increases the scalability of VMFS volumes and vSphere clusters.

Enabling vStorage APIs for Array Integration

Enabling VAAI requires no administrative overhead as IBM FlashSystem V840 volumes are enabled by default. However, administrators can validate that a volume is supported through the vSphere Web

Client or the **esxcli.** For example, to see which functions are supported for a particular storage device, the following command can be used:

 # esxcli storage core device vaai status get

naa.x

VAAI Plugin Name:

ATS Status: supported

Clone Status: supported

Zero Status: supported

Delete Status: unsupported

The overall VAAI status for each ESXi host accessing a storage device is displayed in the general tab of a VMFS data store within the vSphere Web Client, as displayed in **Figure 2**~~Figure 2~~.

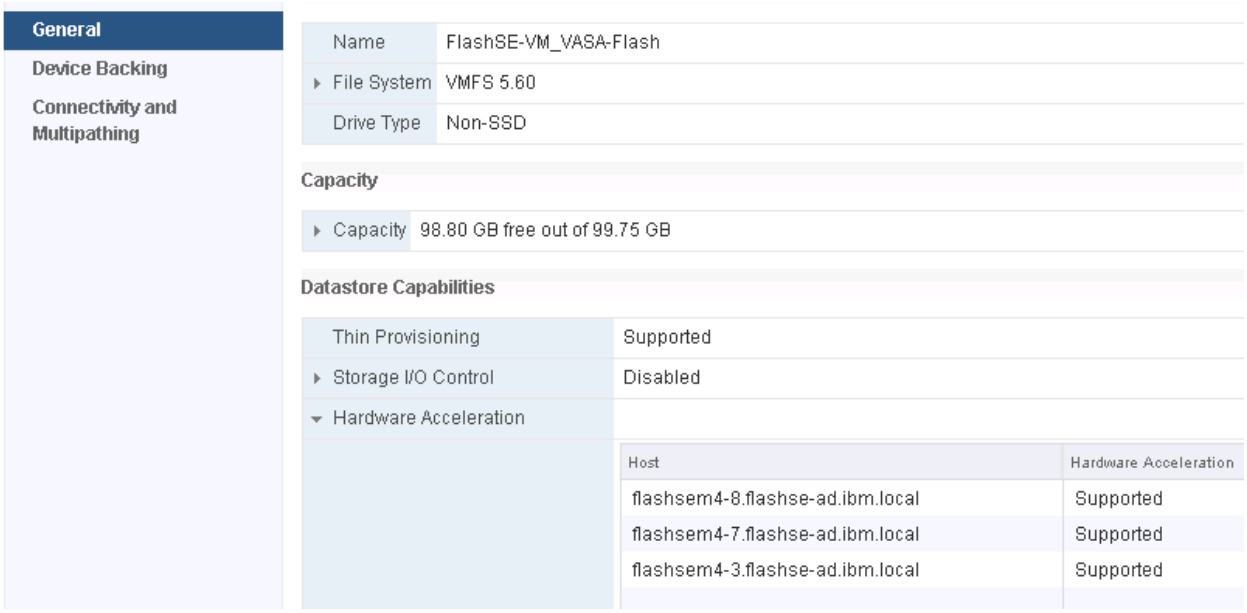


Figure 2 VAAI status per ESXi host

vStorage APIs for Array Integration is a feature that should remain enabled. However, should the need arise to disable VAAI (such as testing the performance benefits as in this white paper) the three functions can be disabled through both the vSphere client and the command line.

Controlling vStorage APIs for Array Integration through the vSphere Web Client

Either the vSphere Web Client or the traditional vSphere Client can be used to disable the vStorage APIs for Array Integration features. Each setting can be modified independently of the others. The procedure is as follows:

1. Browse to the host in the vSphere Web Client navigator.
2. Click the **Manage** tab, and click **Settings.**
3. Under System, click **Advanced System Settings.**
4. Change the value of any of the options to 0 (disabled):
    - VMFS3.HardwareAcceleratedLocking (hardware-assisted Locking)
    - DataMover.HardwareAcceleratedMove (hardware-accelerated Full Copy)
    - DataMover.HardwareAcceleratedInit (hardware-accelerated Block Zero)

Controlling vStorage APIs for Array Integration through the command line

There are multiple command line interfaces which can be used to modify the VAAI features, including PowerCLI, vSphere CLI, or ESXi console connection. This example utilizes commands on the ESXi console connection. The procedure is as follows:

 # esxcli system settings advanced set –int-value 0 –option /DataMover/HardwareAcceleratedMove

 # esxcli system settings advanced set –int-value 0 –option /DataMover/HardwareAcceleratedInit

 # esxcli system settings advanced set –int-value 0 –option /DataMover/HardwareAcceleratedLocking

VAAI performance and benefits with IBM FlashSystem V840

The VMware operations impacted by VAAI receive significant performance and efficiency benefits that are easily identified in production environments. For this technical paper, several test cases were constructed in a lab environment to demonstrate the benefits of the VAAI support provided by IBM FlashSystem V840.

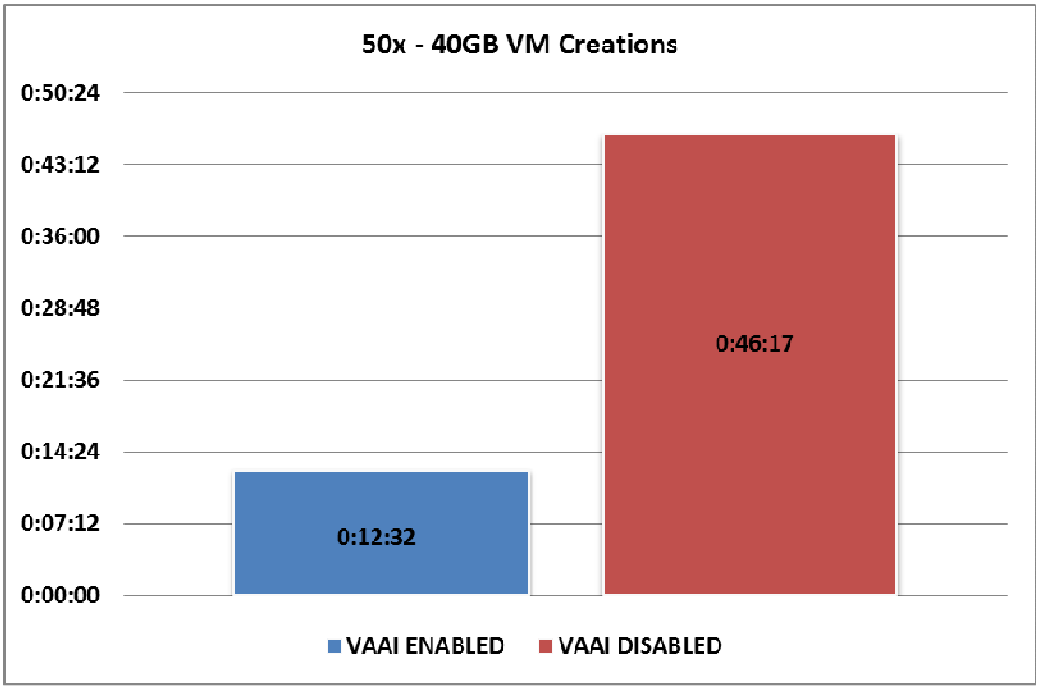
Hardware-accelerated Block Zero test validation

Hardware-accelerated Block Zero increases performance in several scenarios. The test case in this technical paper focuses on the scenario of creating VMware virtual machines with fully-allocated eagerzeroedthick .VMDK files. The test composed utilized a custom script to deploy virtual machines to VMFS data stores. Virtual machine configurations and quantities for each test were as follows:

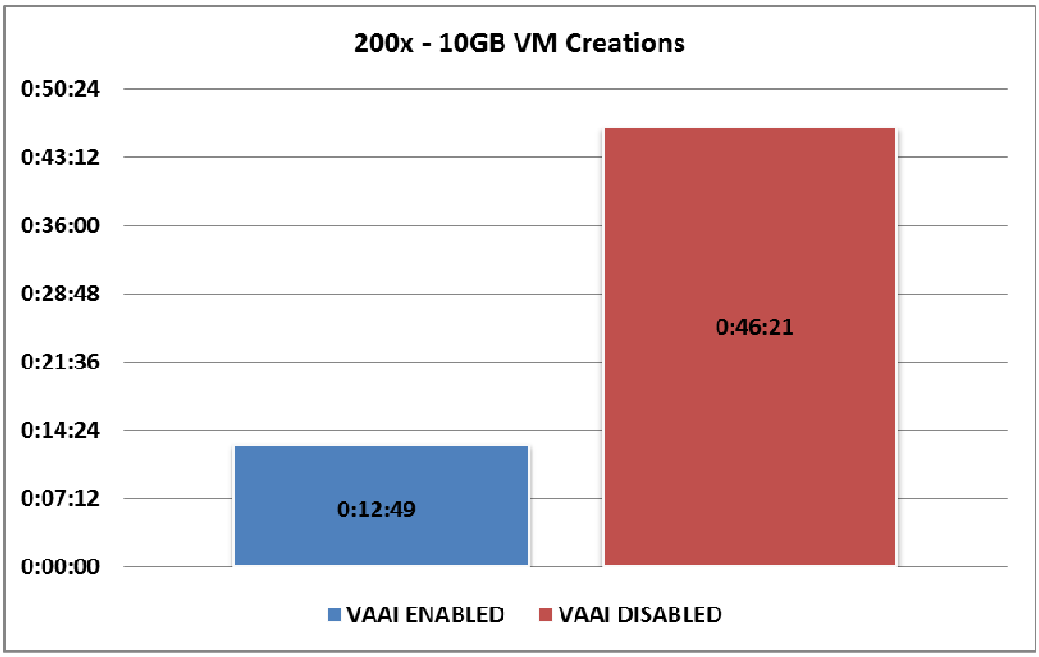
- **Large scenario** - 50 virtual machines – 40GB eagerzeroedthick .VMDK file configured on each virtual machine
- **Small scenario** - 200 virtual machines – 10GB eagerzeroedthick .VMDK file configured on each virtual machine

The metric recorded for this test case was aggregate time to deploy all virtual machines. Each configuration was tested three times to establish an average time to deploy 50 and 200 virtual machines respectively. Configurations were tested with and without hardware-accelerated Block Zero.

Block Zero validation – Large virtual machine scenario



Block Zero validation – Small virtual machine scenario



Each test exhibited an improvement of over 260% with hardware-accelerated Block Zero enabled. The improvement in time was also accompanied with an elimination of ESXi host and SAN resources used for the provisioning tasks.

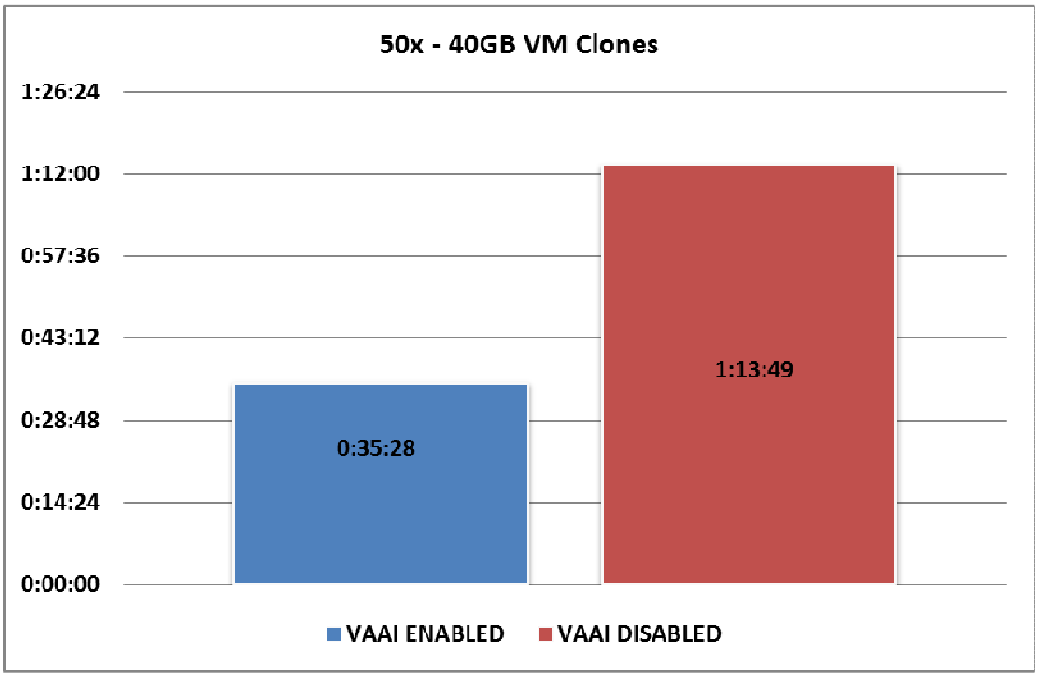
Hardware-accelerated Full Copy test validation

Hardware-accelerated Full Copy is leveraged when creating new virtual machines by cloning a template or existing virtual machine, or by using Storage vMotion to migrate a virtual machine between VMFS data stores. Like the previous Block Zero validation, the Full Copy validation tests used a custom script. This script was used to clone a number of virtual machines from base images. Virtual machine configurations and quantities for each test were as follows:

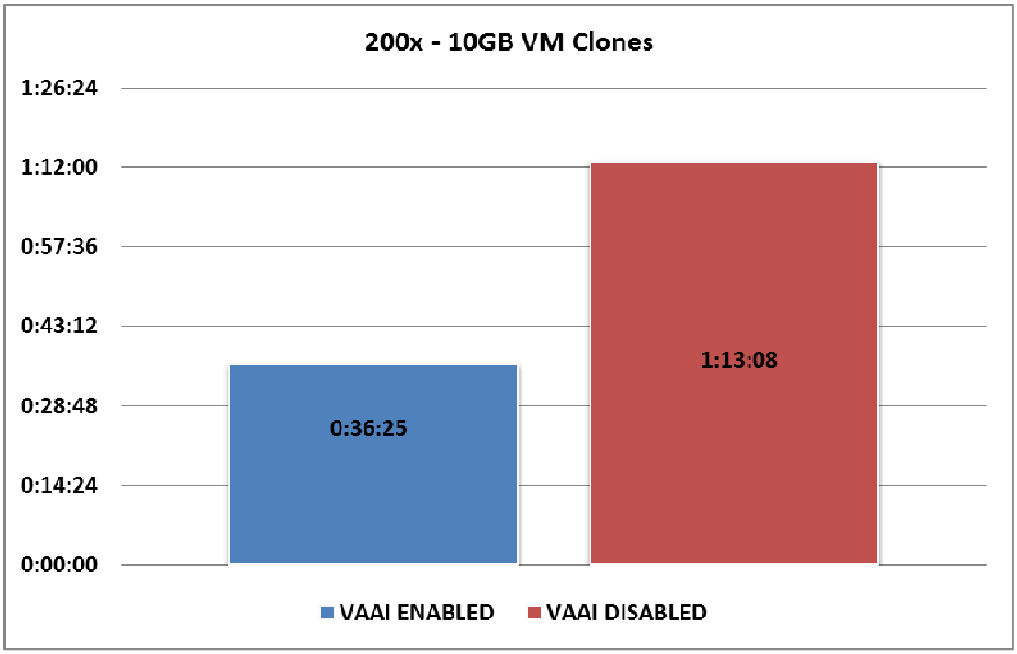
- **Large scenario** - 50 virtual machines – 40GB eagerzeroedthick .VMDK file configured on each virtual machine
- **Small scenario** - 200 virtual machines – 10GB eagerzeroedthick .VMDK file configured on each virtual machine

The metric recorded for this test case was aggregate time to clone all virtual machines. Each configuration was tested three times to establish an average time to clone 50 and 200 virtual machines respectively. Configurations were tested with and without hardware-accelerated Full Copy.

Full Copy validation – Large virtual machine scenario



Full Copy validation – Small virtual machine scenario



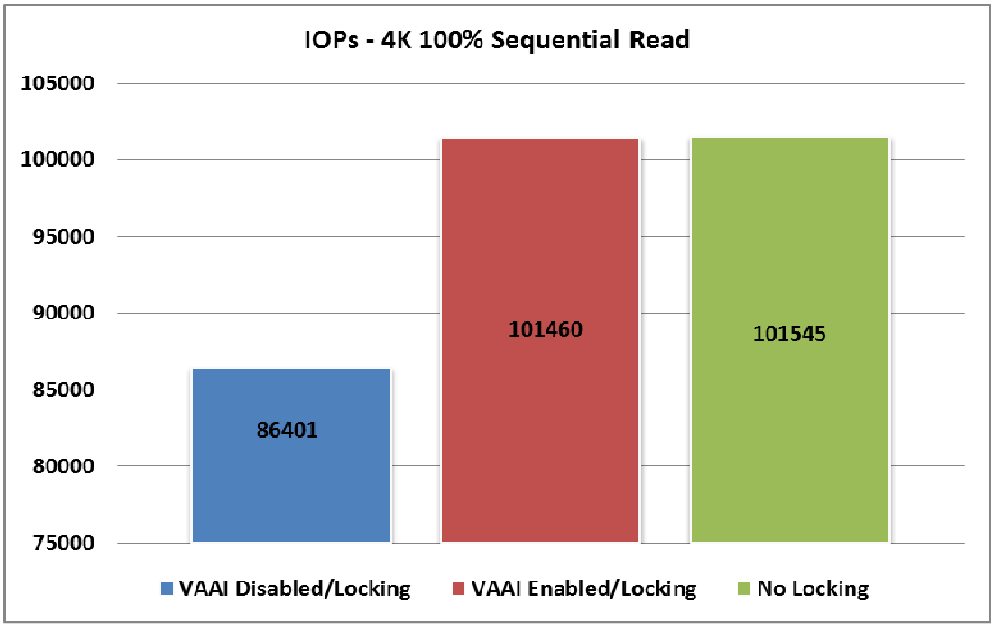
Each test exhibited an improvement of over 100% with hardware-accelerated Full Copy enabled. This improvement in time was also accompanied with an elimination of ESXi host and SAN resources used for the cloning tasks.

Hardware-assisted Locking

Hardware-assisted Locking provides a benefit to nearly every storage-related operation performed by an ESXi host. Operations which generate a high quantity of VMFS locks can easily expose problems when a storage system does not provide support for hardware-assisted Locking. A common operation which can create a high quantity of VMFS locks is the backup of virtual machines on a VMFS data store. Backup software leveraging VMware Data Protection APIs will cause the creation and removal of virtual machine snapshots, which in turn generate VMFS locks. If the VMFS data store does not support hardware-accelerated Locking, as is supported by IBM FlashSystem V840, the performance of other workloads may be impacted.

To validate the benefit provided by hardware-assisted Locking on IBM FlashSystem V840, a custom script was used to repeatedly create and remove virtual machine snapshots on a set of virtual machines sharing a VMFS data store with a running simulated workload. The metric recorded and compared was input/output operations per second (IOPS) of the simulated workload. The test was repeated with hardware-assisted Locking disabled and enabled, and then finally performance without any locking introduced to the VMFS data store.

IOPS comparison with and without hardware-assisted Locking



This test demonstrated that IBM FlashSystem V840 with hardware-assisted Locking prevented any performance impact caused by VMFS locking. However, with hardware-assisted Locking disabled, the

VMFS locks generated by the script caused a 17% impact to the performance of the virtual machine workload.

Summary

Support of VMware vStorage APIs for Array Integration provided by IBM FlashSystem V840 enhances the value of the system for VMware environments. The vStorage APIs for Array Integration functions, such as hardware-accelerated Block Zero and hardware-accelerated Full copy, reduce ESXi host resource consumption otherwise consumed for common tasks, and hardware-assisted Locking aids in large storage scaling and performance. Support of these APIs by IBM FlashSystem V840 increases the overall system performance as well as the efficiency for VMware vSphere environments.

Appendix and Resources:

- IBM FlashSystem V840
  ibm.com/systems/storage/flash/v840/overview.html

- IBM Systems on PartnerWorld
  ibm.com/partnerworld/systems

- VMware website
  www.vmware.com

About the Author:

Rawley Burbridge is a Storage Solution Architect with the IBM FlashSystem Solution Engineering (SE) team, specializing in server, desktop and storage virtualization usage with IBM FlashSystem products.