



# Why You Need VM-Aware Storage

By David Davis, VMware vExpert

Sponsored by

**Western Digital.**

Your virtual infrastructure needs shared storage. You must have it to be able to use advanced virtualization features like high availability and distributed resource scheduling. Up until now, the only storage available was shared storage that knew nothing about virtualization. Similarly, the virtual infrastructure knows little about the storage, other than the basics (LUN, file system, space available, space used). Since the storage and virtual infrastructure are simply two parts of a single whole (the virtual environment), it's unfortunate that they can only communicate in the more rudimentary of ways, thus leading to fractured monitoring. Now, with advancements in storage, all that is changing.

## Storage and Virtualization Working Together

One of the more recent advancements that allow storage and virtualization to work together is VAAI. The vStorage APIs for Array Integration (VAAI) is an Application Programming Interface (API) that storage vendors can support in their arrays. When an array is VAAI enabled and the vSphere hypervisor is aware of this fact, the hypervisor can offload certain storage operations off onto the storage. For example, using VAAI to perform a cloning operation means that the host CPU, memory, and storage fabric not being used for the operation; instead the entire operation is performed within the confines of the array. Essentially, VAAI allows you to offload storage operations onto the array (where they should be). The downside to VAAI is that it isn't going to help you or your array to learn more about your virtual machines and it isn't going to help you to troubleshoot the performance of your array or your virtual infrastructure. Certainly, VAAI is a necessary feature that you need to be using with any array that is providing shared storage for your virtual infrastructure but it isn't the only feature that you need for your storage and virtualization to work together.

What you need is for your array to know about your virtual machines so that it can answer questions like "which virtual machine is demanding the most I/Os per second (IOPS) out of the shared storage?" This way, from a large LUN that stores numerous virtual machines, you can quickly identify any storage

### 5 Reasons You Need Virtualization-Aware Storage

1. Connect your storage and your virtual infrastructure to correlate VM and storage data
2. Learn IOPS, throughput and latency per virtual machine that will allow you to solve problems and plan capacity
3. Plan long-term capacity as you will know the total IO utilization of the virtual infrastructure
4. Understand deduplication and compression rates on a per-VM basis-for your virtual infrastructure storage
5. Simplify Virtualization Storage Configuration, as Tegile's virtualization-aware storage uses simple file-based NFS

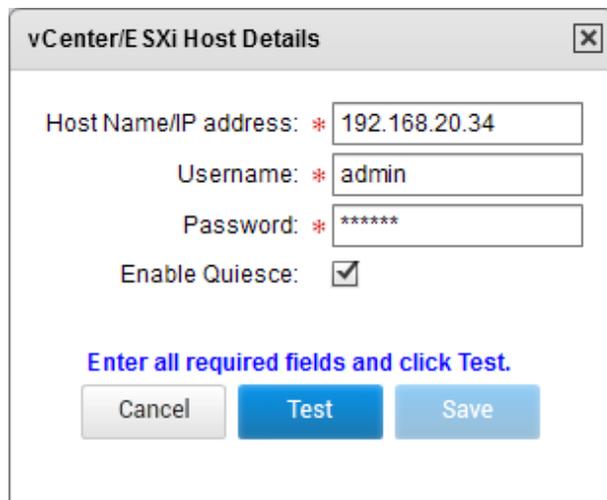
bottleneck down to the virtual machine level so you can more quickly take corrective action. Or questions like “which virtual machine is achieving the best cache hit ratio? Or which virtual machine has the best deduplication rate? The answers to these questions won’t be found in vCenter or in a traditional storage array that isn’t VM-aware.

You need VM-aware storage.

### How Does VM-Aware Storage Work?

VM-Aware storage is storage that talks to the virtual infrastructure and retrieves data about the virtual machines running in the infrastructure, including the virtual machine disk files associated with those VMs, and the location of those virtual disk files. The array then matches up that data with the data it already has to create intelligent, actionable information about how the virtual machines are performing on the array for use by virtualization and storage administrators alike. Finally, the storage and virtual infrastructure can work together.

VM-aware storage works by first connecting your vCenter Server/ESXi Hosts to your storage array, as you see in the graphic below:



The image shows a dialog box titled "vCenter/ESXi Host Details". It has a close button in the top right corner. The dialog contains the following fields and controls:

- Host Name/IP address: \* 192.168.20.34
- Username: \* admin
- Password: \* \*\*\*\*\*
- Enable Quiesce:
- Below the fields, there is a blue text instruction: "Enter all required fields and click Test."
- At the bottom, there are three buttons: "Cancel", "Test", and "Save".

Figure 1 - Connecting vCenter to the array

Once connected, the array shows up within vCenter and allows you to start provisioning storage and also provides performance data that you are used to seeing.

Here's what it looks like in the storage array user interface:

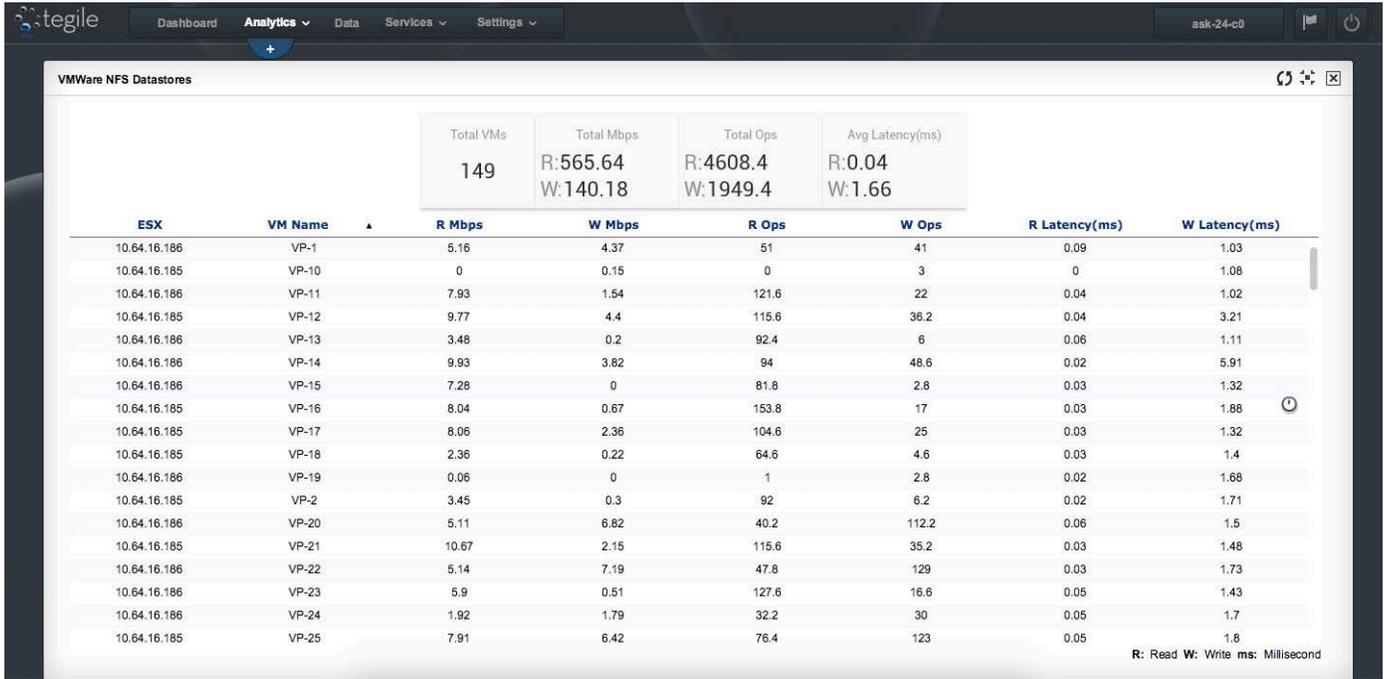


Figure 2 - Real-time virtualization-aware analytics report from array user interface

Here's what it looks like in the vCenter console:



Figure 3 - Real-time virtualization-aware analytics report from vCenter console

From both interfaces, you can see the total number of virtual machines accessing the shared storage array, their ESX host(s), the number of Mbps they are reading and writing, the latency, and the read/write operations per second that they are generating.

## How Virtualization-Aware Storage Helps You

Undoubtedly, better communication between storage and the virtual infrastructure is a “good idea,” but how does it really make the daily life of virtualization and storage admins easier? The best way to explain it is to look at the admin’s life, before and after:

### BEFORE VM-aware storage:

- You use block-based storage and have to understand terms like LUN, volume, block size, stripe width, storage paths, RAID level, WW names, and cache size
- It’s very difficult to determine how many I/Os per second (IOPS) an individual virtual machine or the entire virtual infrastructure
- When there is a problem in the virtual infrastructure, the storage is usually blamed first and is “guilty until proven innocent”.
- There is no way to determine the deduplication or compression rates for individual virtual machines, thus making real world planning impossible.
- Connecting ESXi hosts to datastores requires storage knowledge about iSCSI or Fibre Channel, LUN IDs, and storage pathing (unless you are already using NFS)

### AFTER VM-aware storage:

- You use file-based storage and simply need to know the storage hostname and share name to connect
- The IOPS per VM and for the entire virtual infrastructure are easily visible
- Once you know the IOPS per VM, it’s easy to tell if your storage or virtual infrastructure are to blame when application performance isn’t what it used to be
- It’s easy to tell the deduplication and compression rates per VM
- Connecting ESXi hosts to datastorage is done in a few clicks with only the storage hostname and share name.

With knowledge that virtualization-aware storage provides, both the daily and long-term administration and troubleshooting are made easier.

## In Summary

At too many companies, virtualization stall occurs because of the headaches related to storage complexity and management. By bringing information about virtual machines and VM disk files into the storage array, you'll be able to see the storage resources used by your virtual machines. This knowledge is valuable troubleshooting and capacity management.

My recommendation is to not settle for legacy storage arrays that leave you "flying blind" when it comes to understanding the storage resources used by your virtual infrastructure. Western Digital IntelliFlash arrays are VM-aware, giving you greater visibility into what storage resources are used by each VM. For more information on Western Digital VM-aware storage, visit [www.tegile.com/VM](http://www.tegile.com/VM).

## About the Author

David Davis is a virtualization evangelist at [Pluralsight.com](http://Pluralsight.com), the global leader in video training for IT pros. He holds several certifications including VCP5, VCAP-DCA, and CCIE #9369. He has been awarded the VMware vExpert award 5 years running. Additionally, Davis has spoken at major conferences like VMworld and authored hundreds of articles for websites and print publications, mostly around virtualization, for respected publications like Virtualization Review and ComputerWorld. Follow David on Twitter at [@davidmdavis](https://twitter.com/davidmdavis).

