

How Much Is Enough?

Right-sizing your storage investment

BY ERIC SIEBERT

One of the most vital, and most expensive, resources in a virtualized data center is storage. More than any other resource, inadequate storage can have a negative effect on overall system performance (and on your users' experience) across your virtualized environment.

The key to right-sizing storage is to have enough resources to handle the demand, without overspending on resources that just sit idle.

Because of the nature of virtualized workloads, when storage problems arise the issue is typically not storage capacity, but storage performance—how quickly the storage system can handle the workload requests (IOPS) put to it by the rest of the environment.

Adding more disks can help, but this often results in organizations paying for unused storage capacity when the problem they were trying to fix is storage performance. Fortunately there are solutions such as caching, SSDs, and auto-tiering that can help with storage performance. But again, the challenge is to get the right size solution, and avoid over buying (and overpaying) for excess resources that go unused.

Why storage is such a vital resource in virtualized environments

Virtual machines require four resources to function: CPU, memory, network, and disk. A shortage of any one resource decreases performance and limits your ability to grow your environment. So while you

may have plenty of free host resources available, if your storage (disk) is constrained, you cannot take advantage of those available host resources.

Most virtualized environments use shared storage such as a SAN or NAS device, because shared storage is a requirement for being able to use many of the advanced features that virtualization provides. A shared storage device serves as the foundation that a virtual environment is built upon, and needs to be able to support the entire virtual infrastructure that resides upon it. If your shared storage is not robust enough, it will prevent you from achieving the best possible efficiency and VM density. If shared storage cannot support the workload demands of your virtual machines, performance will suffer, and your attempts to fix the problem can quickly become a money pit.

There are several characteristics that set storage apart as a resource in virtual environments:

- 1. It's the only remote resource**—Of the four resource groups that virtual machines require, shared storage is the only resource that is not local to the host. Virtual machines may run on a host and consume its compute and network resources, but their virtual disks reside on a shared storage device, as shown in the following illustration. This means that each I/O must travel a long way from the host to the shared storage device, with potential bottlenecks along the way that can slow them down.

You can think of it this way: suppose you are in your kitchen making dinner and you have all of your ingredients nearby, except for one key ingredient that is located out back in a storage shed. As you make dinner, having to go out back

constantly to get that one ingredient slows you down and affects when dinner is ready. The same holds true with virtualization and shared storage: Having I/O go from the host (kitchen) to the storage array (shed) affects how well your virtual machines perform.

- 2. It's the slowest resource**—The rate of data transfer to and from traditional spinning disk hard drives is much slower compared to CPU, memory and network resources. Even using much faster solid-state disks, the speeds still don't come close to that of compute resources where access times are measured in nanoseconds instead of milliseconds. Add on the latency of the long journey from the VM to the physical storage device and storage is by far the slowest resource in a virtual environment.

To take the cooking analogy further, if a team of chefs is creating a four-course dinner for you, the time it takes the slowest chef to complete their dish determines when the meal can be served. Likewise, in a virtualized environment, the performance of the slowest resource will dictate the overall performance that a VM can achieve, as it has to constantly wait for the slowest resource (chef) to catch up to the others.

- 3. It's the only shared resource**—Of the four resources that a VM requires, storage is the one resource that is shared among many hosts. All of the other resources are local to a host, which has exclusive access to them and does not have to worry about resource contention from other hosts. Shared storage has a many-to-one relationship, with typically dozens of hosts all accessing the single shared storage device. The end result is a lot of contention for a single resource, which can easily overwhelm a storage device if the workload demands exceed the capabilities and resources of the single storage device.

Back to the food analogy: imagine the line at your favorite coffee shop. As more customers come in for their morning latte, each has to wait longer as the single barista struggles to make shot after shot of espresso to keep up with the demand. The

same holds true in a virtual environment, as the simultaneous workload demands from many hosts increase, the storage device will start to falter and performance will suffer as a result.

How storage affects the bottom line

The classic business case for virtualization is all about increasing efficiency, which results in reduced costs. Cost savings from virtualization include:

- CAPEX savings from purchasing less physical hardware
- OPEX savings from reduced power and cooling
- Other soft savings from reduced management and licensing costs

To gain the most financial benefit from implementing virtualization, you need to increase efficiency as much as you can. This means achieving higher virtual machine densities (more virtual machines per host).

Storage often becomes the limiting factor in achieving higher VM densities because it's a shared resource. When storage is constrained, it impacts the VM density that you can achieve on all of the hosts connected to it. As a result, if you want to achieve the best possible efficiency and VM density in your virtual environment, you need to ensure that your storage does not become the weakest link.

In addition to the business case, there are the administrative benefits to virtualization that make it so popular among IT pros. These benefits include:

- Standardized hardware platform that eliminates OS hardware incompatibilities
- Fast and easy provisioning of new servers
- Centralized server management and easier remote administration
- Improved uptime from virtualization resiliency and high availability features

If storage is constrained, however, and application performance begins to suffer, end user complaints may create pressure to reduce (or at least not expand) the use of virtualization

The challenges with storage for virtualized environments

Clearly, getting storage right is crucial to the success of virtualization projects. But it isn't easy to do. The challenges of implementing shared storage in a virtual environment include:

- Having a properly sized shared storage device to meet all of the workload demands that will be placed on it from the many hosts and VMs that connect to it
- Ensuring your shared storage device can handle peak workloads
- Dealing with the unique demands that virtualization places on a shared storage device

Storage workloads in a virtual environment can be very unpredictable and demanding because storage I/O is all mixed together in a blender-type effect as all of the host workloads are concentrated on a single shared storage device. Virtual environments are also very dynamic, with VMs able to easily travel from host to host or datastore to datastore, which can cause further stress on storage resources.

Storage performance capacity is typically measured in IOPS. Because a storage device can only support a limited number of IOPS based on its architecture and configuration, IOPS becomes an important measurement for sizing storage devices for virtual environments.

Another challenge with shared storage in virtual environments is storage I/O bottlenecks that can severely limit performance. Bottlenecks can occur in different areas on the journey from the VM to the shared storage device, but they typically occur within the shared storage device as it struggles to keep up with the IOPS demands from many hosts. Most bottlenecks form during periods of peak activity, when storage I/O is the greatest, and may exceed the capabilities of the shared storage device to keep up with it.

A good measuring stick for detecting bottlenecks is latency. Latency is the amount of time it takes

a storage I/O request to travel from a guest VM, through a host, to its final destination at the shared storage device. Latency is measured in milliseconds. The higher the latency value, the longer it takes storage I/O to complete its journey, which results in degraded performance. It's normal to have some latency with any storage device, but in general you want latency to be as low as possible which results in the best possible performance.

So how much storage is enough storage?

One of the most challenging aspects of implementing virtualization is figuring out the right amount of storage to meet your VM workload needs. If money were no object, we'd buy tons of storage, but the reality is companies have to minimize costs and stay within budgets. So you need to make smart decisions. If you implement too much storage you're needlessly wasting money, implement too little and you're going to run into performance problems.

Perhaps the trickiest part of sizing storage for virtual environments is trying to achieve a balance of both capacity needs (GB/TB) and performance needs (IOPS). You can end up buying way too much storage capacity (more than you'll ever need) to meet the storage performance requirements for your VM workloads, and end up wasting money.

So how do you know how much storage you'll need for your virtual environment? Start by measuring your current workload and capacity requirements. This is a critical step so don't guess. By making careful measurements and double-checking them you avoid wasting resources.

To do this you need to collect performance data from the servers that you plan on virtualizing. This data needs to be collected over a period of at least 30 days which gives you a better understanding of any performance trends that you may be experiencing that may not happen a regular basis. It is also important to gather metrics during critical business cycles (i.e. weekly payroll processing or a monthly reporting process) where performance may spike. You should note both average and peak resource

usage as you'll need to be able accommodate peak periods or your applications will suffer. You also need to take into account your future needs so you can grow your environment.

There are tools that can help you measure your current workloads and perform capacity planning to ensure that you properly size your shared storage for virtualization. VMware offers a tool called VMware Capacity Planner that automatically collects all the relevant performance metrics and prepares a report that you can use to determine your hardware requirements for your virtual environment. You can also leverage enterprise monitoring tools that capture performance data from each server that you can use to understand your workloads. These tools measure the IOPS that your current servers are generating over time, to calculate your average and peak IOPS usage. You can use this information to size your shared storage for virtualization.

Where sizing storage for virtualization can get complicated is achieving that balance of sufficient capacity and sufficient performance. With the frequent use of thin provisioning, the capacity requirements for virtual environments are typically reduced. This directly affects performance because fewer spindles (hard disks) means fewer available IOPS. As a result, many organizations tend to size their shared storage to meet their capacity requirements but in doing so fall short of the performance they need to handle the combined VM workloads. In addition many organizations also tend to size their shared storage devices based on their combined average IOPS usage, which creates serious issues during peak demand periods, when performance begins to suffer.

The bottom line is that right-sizing storage for virtual environments can be extremely difficult, and many companies suffer from the effects of improperly sized storage, which can degrade performance across the entire virtual environment. Budgets frequently become a limiting factor as well, leaving companies to try and make do with what they can afford. As storage is such a critical resource in a virtual environment this is really one area that you

can't afford to screw up if you want to be successful with virtualization. However, if you take your time and do your homework and complete your due diligence with measuring your workloads it is indeed possible to get this just right.

Oops, I undersized my virtual environment storage. Now what?

Despite your best effort to right-size your shared storage, there's a good chance that budget constraints or unanticipated growth may leave your storage performance struggling.

Because it's so easy to create new VMs, virtual environments tend to grow in a fast and uncontrolled manner in a phenomenon known as VM sprawl. You may not notice storage performance problems right away, but as your virtual environment grows larger and your workload demands increase, it becomes readily apparent. When you start getting those phone calls from your users and application owners that things are running slow, the first place you should check is your shared storage.

Once you find yourself needing to do something about your storage performance problems, where do you start? The traditional solution for this is simple: buy more storage.

To gain more IOPS capability you may think you need to add more spindles. This typically involves adding more shelves to your storage array so you can add more hard disks. If your storage array is maxed out already, you have to buy an entire new array to increase your performance. Either solution can be very expensive and may not be an affordable option.

But there are other methods for increasing the performance capability of your shared storage array including:

- Adding I/O accelerator devices in your servers or storage
- Adding cache in your servers or storage
- Buying expensive array features like auto-tiering
- Adding SSD drives to your servers or storage

These solutions are designed to help handle peak storage I/O periods and to improve overall storage performance, making up for a lack of performance capability in a shared storage device.

While these types of solutions can be a cure for your storage performance problems, that cure typically comes at a cost which can be quite high as you have to add more hardware to your servers and shared storage device. In addition to the budgetary cost, there is also the cost of downtime and disruption to install new hardware.

The challenging aspect of all this is figuring out how much hardware you have to add, and where to add it, to solve your storage performance problems. You could add dozens more hard disks to your shared storage device but will this be enough? Maybe the bottleneck is in the path from the hosts to the shared storage device, or you have a legacy storage array that just plain can't keep up with your workloads.

If you have a dozen virtual servers attached to a single shared storage device you could try adding SSDs or I/O appliances to all of the servers. But that could end up being costly and overkill as you are creating siloes of performance pools that cannot be used by the entire virtual environment. One server may use all of the new SSD capacity, but another may use hardly any, wasting that expensive new hardware. In addition, a virtual environment is very dynamic, with VMs traveling from both host to host and from storage array to storage array; as a result you never know where and how much extra performance capacity you might need as it could change on a daily basis.

Imagine owning a chain of restaurants and you hire extra staff at each restaurant to help ensure that customers get their food faster. If one restaurant gets really busy and service starts to slow down despite the extra staff, having all that additional staff at other restaurants that may not be busy isn't going to help you at that one particular restaurant that needs more help. In addition you

end up paying for a lot of extra help that may not be needed, which ends up being a needless and wasteful expense.

Whatever solution you implement there is always a chance that it won't be enough, or that it's a temporary fix and you'll have to deal with it again later on. No matter what you do it's going to end up being expensive, both in terms of physical hardware and the amount of time and resources you have to spend implementing it.

Infinio Accelerator boosts storage performance in virtual environments

It turns out there is an alternative to hardware-based solutions for taming storage performance issues in virtual environments—Infinio Accelerator.

Imagine if instead of throwing more hardware at your storage performance problem, you could fix the problem with only software. How would you like to solve your problems in hours (and non-disruptively) instead of spending weeks or months and disrupting your whole virtual environment? With Infinio Accelerator you can quickly and easily boost your storage performance and avoid unnecessary expense and wasted time. Infinio Accelerator eliminates the challenge of figuring out how much physical hardware you need to solve your storage performance problem by not requiring any hardware at all.

Infinio Accelerator works as an “offload engine” by becoming a storage proxy and leveraging the RAM that already is present in your hosts. It creates a pooled and deduplicated cache in host memory that uses server-side caching to reduce the I/O traffic load on your shared storage device.

Infinio Accelerator is deployed as a virtual appliance on each host and is transparent to both the host and shared storage device. The end result is that each host contributes a small amount of RAM to the cache pool to create a larger read cache that is shared across all accelerated hosts. By deduplicating the content of the shared cache, the cache size is effectively made even larger.

The impact of this type of solution on storage performance can be dramatic. It allows for both more available IOPS capacity on your shared storage device from the cache offload, and for greatly reduced latency as read requests to local host memory (cache) are much quicker than to the remote storage device.

If we tie the impact that Infinio Accelerator has on storage in a virtual environment back into the analogies that we used earlier to describe some of the challenges with storage it would have the following effect on them:

- 1. Storage as a remote resource**—With Infinio Accelerator, instead of repeatedly going out to the storage shed for the key ingredient for your meal, you have a stockpile of it readily available in the kitchen so you can complete your meal faster.
- 2. Storage as a slow resource**—With Infinio Accelerator the slowest chef becomes one of the fastest and you get to enjoy your dinner that much sooner.
- 3. Storage as a shared resource**—With Infinio Accelerator it's like your favorite coffee shop added more baristas and now you don't have to wait in line as long for your cup of joe.

Even if you don't have a storage problem today you should still consider Infinio. Infinio Accelerator can greatly improve the overall storage performance in your virtual environment, letting you get more from your existing infrastructure. Having Infinio in place also allows you to be proactive in preventing future performance problems as your virtualized environment grows.

Want to find out more and see the benefits firsthand?

Seeing is believing, and the best way to see the effects that Infinio Accelerator will have in your own environment is to give it a try. You can try Infinio Accelerator for free without taking any part of your virtual environment down, and without interrupting operations in any way.

Installation takes about 30 minutes, so your free trial can be up and running quickly. You'll start to see results right away. Infinio Accelerator requires no changes to your existing infrastructure, and it's easily removed (again, non-disruptively) if you aren't completely satisfied with it. Chances are, though, that once you see the benefits first-hand in your own virtual environment, you'll want to keep Infinio Accelerator.