

whitepaper

Storage I/O Bottlenecks in a Virtual Environment

Eric Siebert

Author and vExpert

Storage I/O Bottlenecks in a Virtual Environment

Synopsis

Having a storage I/O bottleneck is perhaps one of the biggest threats to performance in a virtual environment and also one of the more challenging issues to identify and resolve. Bottlenecks can happen as the result of many things, from simple configuration settings to improper storage architecture designs. Storage I/O bottlenecks may not always be obvious, especially if they start small and progressively worsen over time. You may eventually notice that your VMs are starting to slow down, which may cause you to start doing some investigation. But do you know what to look for in order to identify a bottleneck? A virtual environment is like a complicated machine with many moving parts, and there are many different things that can impact the performance of your virtual machines. Knowing the relationships between the many components that make up your virtual environment is important. Even more important is knowing how to interpret the many statistics that are produced so that you can understand the difference between what is normal and what is not, allowing you to identify and correct bottlenecks.

Effects of Storage I/O Bottlenecks

I/O bottlenecks are like being stuck in traffic - the road that you are traveling on cannot accommodate the amount of vehicles on it. This can be because it was designed and built too small with insufficient lanes to handle normal traffic, or because traffic has increased over time to the point that the road can no longer support the volume. The vehicles traveling on the road may have plenty of other resources available to them, such as fuel, but they need all resources to be plentiful to be able to travel at a normal pace. If even one resource becomes constrained, it will slow the vehicle down despite plenty of other available resources.

While a road may be able to handle vehicles traveling at normal speeds most of the time, there are times when traffic volumes may increase to the point where everything slows down. This can be caused by both regular and irregular events. Each day, at a certain time, regular events like rush hour may occur where traffic peaks beyond the roads ability to support the increased volume. Other things like traffic accidents or non-regular events may also occur and slow things down. Roads that are designed to handle average traffic volumes will not be able to handle peak volumes, and as a result, the road will need to be widened or traffic re-directed along other routes to fix the bottleneck that occurs.

The traffic analogy relates well to the storage devices used by your virtual hosts. Bottlenecks can occur at different points in the pipe connecting your storage, and no matter where it occurs, it is going to slow things down. Imagine that your VMs have a single pipe where all storage traffic must travel through to get to the physical storage devices. This pipe may vary in size at different points along its length, with the smallest diameter being the limiting factor to how much storage I/O can flow through it. A bottleneck is just that, you may have an 8 inch pipe along most of the path to the storage device, but if a bottleneck occurs at some point, you may find all of your data trying to funnel through a 1 inch opening. As a result, the total volume of storage I/O that can go through that pipe is greatly reduced and performance will suffer accordingly.



Storage is the slowest component of all host resources due to the mechanical nature of hard disks. So a bottleneck that occurs in the storage subsystem can have a big impact on performance. If a bottleneck occurs, it can severely slow down your virtual machines as the guest operating system waits for disk reads and writes to occur. If a storage I/O bottleneck becomes severe enough, it can cause VMs to experience disk I/O timeouts, which can freeze or cause a Blue Screen Of Death (BSOD) on Windows VMs and cause the disks of Linux VMs to be re-mounted as read-only. A storage I/O bottleneck can effectively cripple or crash much of your virtual environment; therefore, it is critical that you quickly identify the cause and resolve it.

Common Storage I/O Bottleneck Causes

There are three basic things that can cause storage I/O bottlenecks: workload variations, improper configurations/settings and poorly designed storage architectures. Let's take a look at each potential cause in more detail.

Workload Variations

Most storage systems for virtualization are sized to be able to handle known or expected workloads. If the amount of disk I/O increases beyond that workload, the storage system may no longer be able to service the load with the desired response time. Workloads can increase as additional VMs are added to hosts, because of vStorage migrations, changes at the application level or for many other reasons. A storage system has a fixed amount of storage I/O that it can support. This is typically measured in I/O Operations per Second (IOPS). If VM workloads increase so the IOPS requirements of VMs become greater than the IOPS that the storage system can support, slowdowns and bottlenecks start to occur.

Improper Configurations

This can consist of storage related settings that are not set properly on either the host or storage device, improper placement of VMs on datastores and also host misconfigurations. Storage settings, such as queue depth, cache size, and network-specific settings for iSCSI & NFS can cause poor performance if improperly set. A host that has too many VMs with high disk I/O can easily overwhelm the storage adapter on the host. Likewise, a storage device that has too many VMs with high disk I/O can easily overwhelm the storage device. Finally, host-misconfigurations, such as improperly configured multi-pathing policies, can result in poor performance that leads to bottlenecks.

Poorly Design Storage Architectures

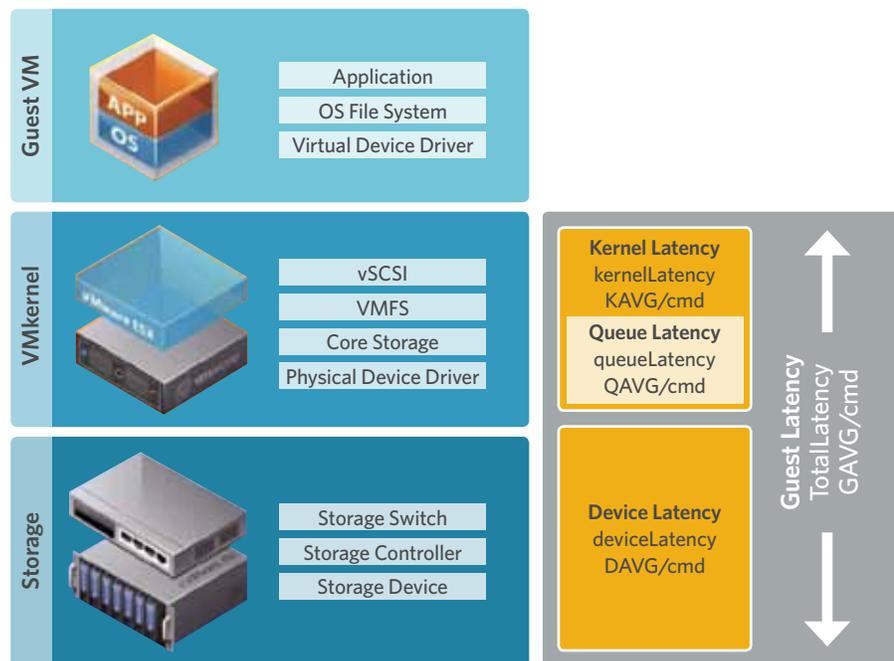
Storage devices that are not architected properly to support the VM workloads that are running on them can cause bottlenecks and poor performance. One common design mistake in this area is improperly sized RAID groups that have too few spindles to spread the workload across. Other poor architecture decisions include

choosing to use under-performing low RPM and/or SATA drives, using a RAID level that has too high of a write penalty, not load balancing storage processors or interfaces, using array caches that are too small, or improper storage network design. Problems can occur in these physical storage subsystems and arrays, so it is critical to understand how virtualization maps into your physical storage such as the underlying LUNs, RAID groups and disks. A properly designed and sized storage system is absolutely critical to ensure good performance and eliminate the potential for bottlenecks to occur.

Identifying Storage I/O Bottlenecks

The key to identifying storage I/O bottlenecks is having a good monitoring and reporting solution in place to alert you when they occur. There are specific performance statistics that are directly related to storage I/O bottlenecks that can be monitored to tell if one exists. SolarWinds Virtualization Manager is equipped to handle this by providing you with detailed visibility into your storage subsystem so you can identify bottlenecks and deal with them. Using storage metrics, such as latency, throughput and IOPS, you can see how your storage is performing and exactly what areas need to be addressed. Virtualization Manager can help you stay on top of your storage I/O and proactively alert you when problems occur across all your storage tiers.

There are several key statistics that should be monitored on your storage subsystem related to bottlenecks, but perhaps the most important is latency. Disk latency is defined as the time it takes for the selected disk sector to be positioned under the drive head so it can be read or written to. Once a VM makes a read or write to its virtual disk, that request must follow a path to make its way from the guest OS to the physical storage device. A bottleneck can occur at different points along that path, but there are different statistics that can be used to help pinpoint where the bottleneck is occurring. The following figure illustrates the path that data takes to get from the VM to the storage device.



The storage I/O goes through the operating system as it normally would and makes its way to the device driver for the virtual storage adapter. From there, it goes through the Virtual Machine Monitor (VMM) of the hypervisor, which emulates the virtual storage adapter that the guest sees. It travels through the VMkernel and through a series of queues before it gets to the device driver for the physical storage adapter that is in the host. For shared storage, it continues out the host onto the storage network and makes its way to its final destination, which is the physical storage device. Total guest latency is measured at the point where the storage I/O enters the VMkernel, up to the point where it arrives at the physical storage device.

The total guest latency (GAVG/cmd as it is referred to in the esxtop utility) is measured in milliseconds and consists of the combined values of kernel latency (KAVG/cmd) plus device latency (DAVG/cmd). The kernel latency includes all the time that I/O spends in the VMkernel before it exits to the destination storage device. Queue latency (QAVG/cmd) is a part of the kernel latency, but also measured independently. The device latency is the total amount of time that I/O spends in the VMkernel physical driver code and the physical storage device. So when I/O leaves the VMkernel and goes to the storage device, this is the amount of time that it takes to get there and return. A guest latency value that is too high is a pretty clear indication that you have a storage I/O bottleneck that can cause severe performance issues. Once total guest latency exceeds 20ms, you will notice the performance of your VMs suffer; and as it approaches 50ms, your VMs will typically become unresponsive.

To identify where latency is occurring, you look at the individual values that make up guest latency.

- **Kernel latency** - Average amount of time, in milliseconds, spent by VMkernel processing each SCSI command. This value should be as close to zero as possible and less than 1ms.
- **Queue latency** - Average amount of time spent in the VMkernel queue, per SCSI command, during the collection interval. This value should be as close to zero as possible and less than 1ms.
- **Device latency** - Average amount of time, in milliseconds, to complete a SCSI command from the physical device. This value should be between 0-5ms for higher speed, 4/8GBps fibre-channel and 10GBps iSCSI/NFS storage devices and between 5-10ms for lower speed, 1/2GBps fibre-channel and 1GBps iSCSI/NFS devices.

All of the latency statistics are further split into two sub-statistics for read and write so you can see exactly which operation latency is occurring. In addition to latency that measures delays in storage I/O delivery, you should also pay attention to IOPS that measure the volume of storage I/O traffic that is occurring. IOPS can show you which VMs are generating the most storage I/O traffic on a host. Total IOPS for a host can also show you the combined total storage I/O for a storage device.

The final statistic that is useful in identifying storage I/O bottlenecks is QUED (queue depth), which is the number of I/O operations that are queued up and require processing, but have not yet been addressed by the VMkernel. Any number higher than zero for QUED can indicate a storage I/O bottleneck.

Virtualization Manager has built-in storage dashboards that allow you to track and monitor all of the key areas that may cause storage I/O latency. Virtualization Manager can alert you when conditions exist that can cause bottlenecks so you can quickly resolve them before they impact your virtual machines. It can also alert you when latency occurs and identify which VMs are generating the most storage I/O. You can also monitor latency on multiple levels and drill down further to get more granular and detailed information. It also provides a trend view so you can see exactly how latency builds up over time, which can help you predict when bottlenecks may start to occur in the future. With Virtualization Manager, you no longer have to waste time trying to sort through statistics in vCenter Server, as you can clearly see exactly what is going on in your environment, quickly at-a-glance.

Solving Storage I/O Bottlenecks

Once you have identified where your storage I/O bottleneck is located, you need to take action to correct it. There are different resolutions to each situation depending on where the bottleneck is occurring. Below are some of the resolutions for each of the common areas where you may experience storage I/O bottlenecks.

- High kernel latency (KAVG/cmd) - it is rare for this to be high (>1ms), but when it is, it usually indicates a problem with queuing as commands sit and wait in the VMkernel. One common cause for this is that the queue depth is set too small, so commands have to wait to enter the queue.
- High queue latency (QAVG/cmd) - it is rare for this to be high (>1ms) as well; and if so, it also indicates a queuing problem as commands are spending too much time in the queue. This, too, can be the result of a queue depth that is set too small.

- High device latency (DAVG/cmd) - if this value is high, it indicates a problem with the storage device. This value is the time that it takes for a command to leave the storage adapter on a host, get processed by the storage device and then return to the storage adapter. High device latency can be a result of an over-loaded or improperly architected storage device that is having difficulty keeping up with all the storage I/O that is occurring. Keep in mind that multiple hosts can access the same storage devices concurrently, so the combined I/O between all the hosts may be too much for it to handle.
- High queue value (QUED) - if this value is greater than zero, it can indicate that the queue depth is set too small and may need to be increased. Refer to VMware KB article #1267 for detailed instructions on how to properly increase this.
- High IOPS (commands) - if this value is high, it can indicate that the amount of I/O that is occurring might be greater than the storage device can handle. There are a number of factors (i.e. cache) that determine the amount of IOPS that a storage device can handle, but it is mainly determined by the rotational speed of a drive multiplied times the number of drives in a RAID group. The higher the speed of a drive, the more IOPS it can handle. So a 15,000 rpm drive can handle roughly twice as many IOPS as a 7,200 rpm drive. A RAID group consisting of six 15K drives that typically handle about 175 IOPS each would be capable of around 1,050 IOPS. Keep in mind, there is also a RAID penalty for disk writes that slightly decreases the amount of IOPS available. The greater the level of RAID protection, the higher this penalty is. Re-arranging your workloads or re-architecting your storage can help ensure that the number of IOPS that your VMs are generating does not exceed the number that your storage device is capable of.

Solving storage I/O bottlenecks can be complicated, as there are many factors that can contribute to their cause. The performance statistics may tell you where to look, but not exactly what to fix. Some bottlenecks can easily be fixed by changing settings or by balancing storage I/O across datastores. Other bottlenecks can be more difficult to resolve and may require adding additional storage capacity or upgrading components.

Summary

Bottlenecks are not always obvious and you may not know you have one until you actually take the time to look for it. SolarWinds Virtualization Manager has you covered in this area and can help you stay informed on what is going on in your whole virtual environment. The dashboards in Virtualization Manager can provide you with a wealth of critical information in a single pane of glass so you can quickly see the health of your environment. The dashboards can also be customized to meet your specific needs and you can drill-down through them to get even more detailed information on specific areas.

Furthermore, Solarwinds Storage Manager compliments SolarWinds Virtualization Manager by allowing you to see beneath the virtual layer to the physical storage subsystems and arrays underneath, allowing you to pinpoint I/O issues or contention that could be occurring at the LUN or underlying RAID groups. This provides you with a comprehensive end-to-end view of your virtualization assets and their physical storage dependencies.

Storage I/O bottlenecks are perhaps the biggest threat in virtualization; you can't afford to let them choke off the resources that your VMs require. You're in the driver's seat of your virtual environment, so don't let your VMs get stuck in traffic -- let Virtualization Manager and Storage Manager keep you informed of the road conditions so you can keep your VMs running smoothly and without any delays.

About the Author

Eric Siebert is an IT industry veteran, author and blogger with more than 25 years of experience, most recently specializing in server administration and virtualization.

Siebert has published books, including his most recent, *Maximum vSphere* from Pearson Publishing and has authored training videos in the Train Signal Pro series. He also maintains his own VMware information web site, vSphere-land.com, and is a regular blogger and feature article contributor on TechTarget's SearchServerVirtualization and SearchVMware web sites. Siebert has presented at VMworld in 2008 and 2010 and has been recognized as a vExpert by VMware in 2009 and 2010.

About SolarWinds

SolarWinds, a leading provider of IT management software to more than 97,000 customers worldwide, has acquired Virtualization Manager (Hyper9). Virtualization Manager's products add optimization, capacity planning, configuration and chargeback reporting for virtualized server environments to the SolarWinds portfolio.

SolarWinds (NYSE: SWI) provides powerful and affordable IT management software to more than 97,000 customers worldwide - from Fortune 500 enterprises to small businesses. In 2011, SolarWinds acquired Virtualization Manager (Hyper9) adding optimization, capacity planning, configuration and chargeback reporting for virtualized server environments to the SolarWinds product portfolio. The company works to put its users first and remove the obstacles that have become "status quo" in traditional enterprise software. SolarWinds products are downloadable, easy to use and maintain, and provide the power, scale, and flexibility needed to address users' management priorities. SolarWinds online user community, <http://thwack.com>, is a gathering-place where tens of thousands of IT pros solve problems, share technology, and participate in product development for all of the company's products. Learn more today at <http://solarwinds.com>.

For additional information, please contact SolarWinds at 866.530.8100 or e-mail sales@solarwinds.com. To locate an international reseller near you, visit http://www.solarwinds.com/partners/reseller_locator.aspx

Did you like this white paper? Tweet about it.  <http://www.twitter.com>



We are a leading IT Solutions provider specialising in Network and Systems Management monitoring solutions for small/medium and enterprise networks. Prosperon Networks provides a range of Professional Services to support its portfolio and is recognised as an authority within the UK for designing and installing management solutions from SolarWinds, Prosperon Networks markets its solutions across the UK in all major verticals.

Key Facts:

- Founded in 2006, has seen 100% growth year on year
- Home Office Security Clearance to work on sensitive Government sites
- Platinum Partner Status for SolarWinds with SCP Accredited Engineers
- 1st & 2nd Line telephone and email technical support offered to customers
- Provides more training and consultancy than any other SolarWinds partner
- Offsite classroom training facilities, full range of onsite courses offered
- Installation, Customisation and post Consultancy SolarWinds services provided.

Technical Support Information:

Telephone Contact Hours – 9:00 – 17:30
Support Email – support@prosperon.co.uk
Telephone Support Number: 0845 8331185

Contact Information:

Prosperon Networks Limited
55 Chapel Road
Worthing, West Sussex
BN11 1EE
Tel: 0845 8331185
International: +44 1903 239504
Email: sales@prosperon.co.uk
www.prosperon.co.uk

