

- [Home](#)
- [About](#)
- [Resume](#)
- [Virtualization Jobs](#)
- [Virtualization Resources](#)

VMtoday

VMware News, Views, & How-To's from Josh Townsend

« [Update: SVGA Drivers on Windows 2008 R2 and Windows 7](#)

[Washington DC VMware User Group – April 29th](#) »

Storage Basics – Part VI: Storage Workload Characterization

April 8th, 2010 | Author: [Joshua Townsend](#)

Most of what I covered in Storage Basics Parts 1 through 5 was at a very elementary level. The math I used to do [IOPS calculations](#), for example, is only true under very certain conditions. [RAID](#) controllers implement [caching](#) and other techniques that skew the simple math that I provided. I mentioned that the type of [interface](#) that you ought to use on your storage array should not be randomly chosen. In fact, choosing the right array with the appropriate components and characteristics can only be done when you enlighten your decision with a characterization of workloads it will be running.

[6](#)
[tweets](#)
[retwee](#)

The character of your storage workload can be broken down into several traits – random vs. sequential I/O, large vs. small I/O request size, read vs. write ratio, and degree of parallelism. The traits of your particular workload dictate how it interacts with the components of your storage system and ultimately determine the performance of your environment under a given configuration. There is an excellent whitepaper available from VMware entitled “[Easy and Efficient Disk I/O Workload Characterization in VMware ESX Server](#)” that is authoritative on this subject. If you want to get down and dirty with the topic, it’s a good read. I’m aiming for something a bit less academic. With that said, let’s break down workload characterization a bit so as to better understand how it will impact your real-world systems.

Random vs. Sequential Access

In [Part II](#) of this series we looked at the formula for calculating IOPS capabilities for a single disk. That formula goes something like this:

$$\text{IOPS} = 1000 / (\text{Seek Latency} + \text{Rotational Latency})$$

You’ll recall that we divide into 1000 to remove milliseconds from the equation, leaving (Seek Latency + Rotational Latency) as the important part of the equation. Rotational latency is based on the spindle speed of the disk – 7.2k, 10k, or 15k RPM for standard server or SAN disks. If we consider [the same Seagate Cheetah 15k drive from Part II](#), we see that rotational latency is 2.0ms. The only way to change rotational latency is to buy faster (or slower) disks. This essentially leaves seek latency as the only variable that we can “adjust”. You’ll also recall that seek latency was the larger of the latencies (3.4ms for read seeks, and 3.9ms for write seeks) and counts more against IOPS capability than does rotational latency. Seeking is the most expensive operation in terms of performance.

It is next to impossible to adjust seek latency on a disk because it is determined by the speed of the servos that move the heads across the platter. We can, however, send workloads with different degrees of randomness to the platter. The more sequential a workload is, the less time that will be spent in seek operations. A high degree of sequentiality ultimately leads to faster disk response and higher throughput rates. Sequential workloads may be candidates for slower disks or RAID levels. Conversely, workloads that are highly randomized ought to be placed on fast spindles in fast RAID configurations.

You’ll notice that I said it was next to impossible to adjust seek latency on a disk. While not common, some storage administrators employ a method know as ‘short stroking’ when configuring storage. Short stroking uses less than the full capacity of the disk by placing data at the beginning of the disk where access is faster, and not placing data at the end of the disk where seeks times are greater. This results in a smaller area on the disk platter for heads to travel over, effectively

reducing seek time at the expense of capacity.

While not applicable to all workloads, storage arrays, or file systems, fragmentation can cause higher degrees of randomness leading to degraded performance. This is the prime reason some vendors recommend that you regularly defragment your file system. It should be noted that a VMware VMFS file system is resilient against the forces of fragmentation. Whereas a Windows NTFS partition may hold hundreds, thousands or tens of thousands of files of different sizes, accessed randomly throughout the system's cycle of operations, a VMFS datastore typically holds no more than a couple hundred files. Additionally, most of the files on a VMFS datastore are created contiguously if you are using thick-provisioned virtual disks (VMDK). Thin-provisioned VMDK's are slightly more susceptible to fragmentation, but do not typically suffer a high enough degree of fragmentation to register a performance impact. See this VMware whitepaper for more on VMFS fragmentation: [Performance Study of VMware vStorage Thin Provisioning](#).

Examples of sequential workloads include backup-to-disk operations and the writing of SQL transaction log files. Random workloads may include collective reads from Exchange Information Stores or OLTP database access. Workloads are often a mix of random and sequential access, as is the case with most VMware vSphere implementations. The degree to which they are random or sequential dictates the type of tuning you should perform to obtain the best possible performance for your environment.

I/O Request Size

I/O request size is another important factor in workload characterization. Generally speaking, larger reads/writes are more efficient than smaller I/O to a certain point. The use of larger I/O requests (64KB instead of 2KB, for example) can result in faster throughput and reduced processor time. Most workloads do not allow you to adjust your I/O request size. However, knowing your I/O request size can help with appropriate configuration of certain parameters such as array stripe size and file system cluster size. Check with your storage vendor for more information as it pertains to your specific configuration.

If you are in a Windows shop, you can use perfmon counters such as Avg. Disk Bytes/Read to determine average I/O size.

If you are running a VMware-virtualized workload, you can take advantage of a great tool – vscsiStats – to identify your I/O request size. More on vscsiStats later in this article.

Read vs. Write

Every workload will display a differing amount of read and write activity. Sometimes a specific workload, say Microsoft Exchange, can be broken down into sub-workloads for logging (write-heavy) and reading the database (read-heavy).

Understanding the read-to-write ratio may help with designing the underlying storage system. For example, a write-heavy workload may perform better on a RAID10 LUN than a RAID5 array due to the write penalty associated with RAID5. The ratio of read:write may also dictate caching strategies. The read:write ratio, when combined with a degree of randomness measure, can be quite useful in architecting your storage strategy for a given application or workload.

Parallelism/Outstanding I/O's

Some workloads are capable of performing multi-threaded I/O. These types of workloads can place a higher amount of stress on the storage system and should be understood when designing storage, both in terms of IOPS and throughput.

Multipathing may help with multi-threaded I/O workloads. A typical VMware vSphere environment is a good example of a workload capable of queuing up outstanding I/O.

Measuring the Characteristics of Your Workload

So how do we actually characterize storage workloads? Start with the application vendor – many have published studies that can shed light on specific storage workloads in a standard implementation. If you are interested in measuring your own for planning/architecture reasons, or performance troubleshooting reasons, read on.... There are several tools to measure storage characteristics, depending on your operating system and storage environment. Standard OS performance counters, such as Windows Performance Monitor (perfmon) can reveal some of the characteristics. Array based tools such as NaviAnalyzer on EMC gear can also reveal statistics on the storage end of the equation.

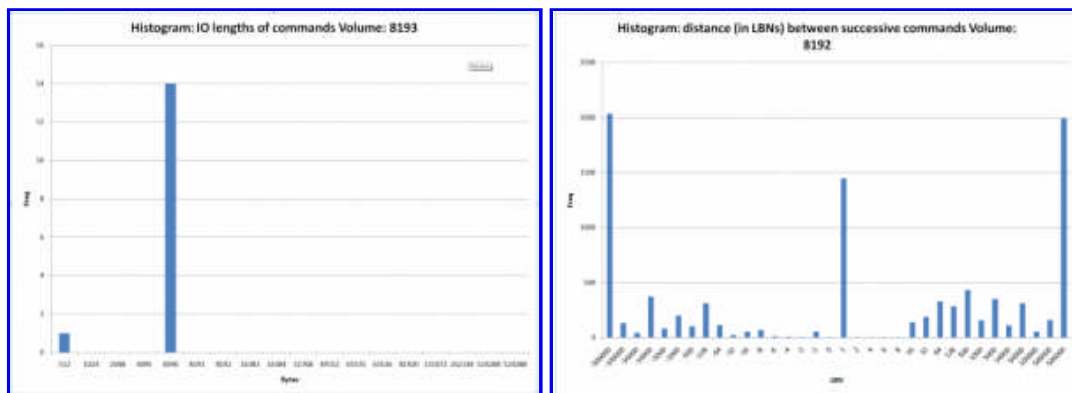
One of the most exciting tools for storage workload characterization comes from VMware in the form of *vscsiStats*.

vscsiStats is a tool that has been included in VMware ESX server since version 3.5. Because all I/O commands pass through the Virtual Machine Monitor (VMM), the hypervisor can inspect and report on the I/O characteristics of a particular workload, down to a unique VM running on an ESX host. There is a ton of great information on using vscsiStats, so I won't re-hash it all here. I recommend starting with [Using vscsiStats for Storage Performance Analysis](#) as it contains an overview

and usage instructions. If you want to dig a bit deeper into vscsiStats, read both [Storage Workload Characterization and Consolidation in Virtualized Environments](#) and [vscsiStats: Fast and Easy Disk Workload Characterization on VMware ESX Server](#).

vscsiStats can generate an enormous amount of data which is best viewed as a histogram. If you're a glutton for punishment, the data can be reviewed manually on the COS. To extract vscsiStat output data, use the -c option to export to a .csv file.

From there you can analyze the data and create histograms using Excel. Paul Dunn has a nifty Excel macro for analyzing and reporting on vscsiStats output [here](#). Gabriele van Zanten more detailed instructions for using Paul's macro [here](#). Here are a couple histogram examples that I just generated from a test VM.



vscsiStats is only included with ESX, not ESXi. However, Scott Drummond was kind enough to post a download of vscsiStats for ESXi on his Virtual Pivot blog: <http://vpivot.com/2009/10/21/vscsistats-for-esxi/>. Using vscsiStats on ESXi requires dropping into Tech Support Mode (unsupported) and enabling ESXi for scp to transfer the binary to the ESXi server.

VMware *esxstop* can display some information but is limited in scope and does not currently support NFS. A [community-supported python script](#) called nfstop can parse vscsiStats data and display esxstop-like data per VM on screen.

Experiment

If you are interested in generating workloads with various characteristics, check out [Iometer](#) and [Bonnie++](#). These tools will allow you to generate I/O that you can monitor with the tools I covered in this article.

Put it to Use

If you are provisioning a new workload or expanding an existing, invest some time in understanding your storage workload characteristics and convey those characteristics to your storage team. A request for storage that includes the workload characteristics I discussed here, as well as expected IOPS requirements, will go much further in ensuring performance for your applications – physical or virtual – than simply asking for a certain capacity of disk.

Related posts:

1. [Storage Basics – Part V: Controllers, Cache and Coalescing](#)
2. [Storage Basics – Part II: IOPS](#)
3. [Storage Basics – Part VII: Storage Alignment](#)
4. [Storage Basics – Part I: An Introduction](#)
5. [Storage Basics – Part III: RAID](#)

Posted in [Storage](#), [Storage Basics](#), [VMware](#) | Tags: [esxstop](#), [I/O](#), [IOPS](#), [performance](#), [SAN](#), [Storage](#), [VMware](#), [vscsiStats](#)

 [4 Tweets](#)

4 Responses to “Storage Basics – Part VI: Storage Workload Characterization”



- [joshuatownsend:](#)
[April 8, 2010 at 11:34 am](#)

New VMtoday.com post: Storage Basics – Part VI: Storage Workload Characterization <http://bit.ly/9hMGcU> #I/O #IOPS #SAN #vmware

This comment was originally posted on [Twitter](#)

[Reply](#) [More from author](#)



- [PlanetV12n:](#)
[April 8, 2010 at 12:21 pm](#)

Storage Basics – Part VI: Storage Workload Characterization (VMtoday.com) <http://bit.ly/9p0e7N>

This comment was originally posted on [Twitter](#)

[Reply](#) [More from author](#)

- [Top 5 Planet V12n blog posts week 14 / VMvisor:](#)
[April 11, 2010 at 8:02 pm](#)

[...] Townsend – Storage Basics – Part VI: Storage Workload Characterization You’ll recall that we divide into 1000 to remove milliseconds from the equation, leaving (Seek [...])

[Reply](#) [More from author](#)

- [cmdln.org \(a sysadmin blog\) » Blog Archive » Analyzing I/O performance in Linux:](#)
[April 22, 2010 at 2:01 pm](#)

[...] you want to use sysstat like I did, you might find this Nagios plug-in that I wrote helpful check_sar_perf. I use it with Zenoss, but it could be tied into any NMS that records the performance data from a [...]

[Reply](#) [More from author](#)

Leave a Reply

Name (required)

Mail (will not be published) (required)

Website

[Submit Comment](#)

Additional comments powered by [BackType](#)

Recent Posts

- [VMware vExpert](#)
- [EMC Virtual Storage Integrator Update](#)
- [Storage Basics – Part VII: Storage Alignment](#)
- [Free SAN Monitor for DS3300, MD3000i and others](#)
- [Washington DC VMware User Group – April 29th](#)
- [Storage Basics – Part VI: Storage Workload Characterization](#)
- [Update: SVGA Drivers on Windows 2008 R2 and Windows 7](#)

Tags

[3.5](#) [2009](#) [best practices](#) [client](#) [compatibility](#) [disk](#) [DL380](#) [ecosystem](#) [EMC](#) [ESX](#) [esxi](#) [esxtop](#) [fusion](#) [G3](#) [HP](#) [Hyper-V](#) [I/O](#) [IOPS](#) [memory](#) [Microsoft](#) [NetApp](#)
[network](#) [Patch](#) [performance](#) [powershell](#) [reporting](#) [SAN](#) [security](#) [sizing](#) [sql](#) [statistics](#) [Storage](#) [Update 3](#) [vcenter](#) [VI3](#) [virtual](#) [virtual center](#)
[virtualization](#) [Virtual Machine](#) [VM](#) [VMUG](#) [VMware](#) [vmworld](#) [vsphere](#) [windows](#)

Categories

- [Documentation](#) (2)
- [EMC](#) (2)
- [General IT](#) (10)
- [Issues & Troubleshooting](#) (18)
- [Microsoft](#) (8)
- [NetApp](#) (3)
- [News](#) (6)
- [Personal](#) (3)
- [Product Releases](#) (2)
- [Storage](#) (8)
- [Storage Basics](#) (3)
- [Uncategorized](#) (3)
- [VMUG](#) (2)
- [VMware](#) (46)
- [VMware How To](#) (18)
- [VMworld](#) (3)

[Recent Tweets](#)

- [@Virtustream360](#) Thanks for the follow - nice to see you guys on Twitter!
- [@DellServerGeek](#) YUM!
- RT [@ericsiebert](#): @TrainSignal's new Pro Series video - my experience - <http://bit.ly/brIE35> <- Nice - I look forward to watching yours!
- [@stu](#) thanks for including me in your VMworld 2010 list! - <http://tweepml.org/VMworld-2010-San-Francisco/>
- RT [@jasonboche](#): North Korea is an idiot. <- Ha!
- RT [@cshanklin](#): I hope that if, one day, I accidentally post something to twitter that it's totally awesome.
- [@tscalzott](#) Quality.
- [@tscalzott](#) That was some impressive rain today!
- Please, sir. May I have another?
- RT [@veeam](#): Veeam FastSCP 3.0.3 with vSphere 4.1 support is available for download (and still FREE): <http://bit.ly/aaV3gc> <- time 2 thinapp!

[Recent Bookmarks](#)

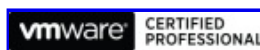
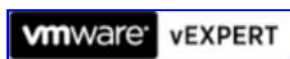
- [Windows Guest Defragmentation, Take Two « Pivot Point](#)
- [Dan Weiss Blogsite » NetApp performance testing engineer tells it like it is](#)
- [Understanding how storage design has a big impact on your VDI! - Ruben Spruijt - BrianMadden.com](#)
- [Performance Troubleshooting VMware vSphere – Storage | virtual insanity](#)
- [PowerCLI – Where do I start ?! « Virtu-Al](#)

- [CPU/MEM Reservation Behavior » Yellow Bricks](#)
- [Refreshing Test Dev SQL Environments using NetApp Technologies · Rajeev Karamchedu](#)
- [Finding thin-provisioned virtual disks with PowerShell | VCritical](#)
- [ESXi4 Console - Secret Commands for Emergency Maintenance](#)
- [VMware Server on a Domain Controller \(DC\) « Derek's Musings](#)
- [How to Enable SSH in ESXi – VMware « Code Ghar](#)
- [VMware ESX, EMC CLARiiON Arrays, and Multiple Protocols - blog.scottlowe.org - The weblog of an IT pro specializing in virtualization, storage, and servers](#)
- [PVSCSI and Low IO Workloads « Pivot Point](#)
- [VDI.com - Why should I provision VDI VM's with less RAM and other resources than my physical machines?](#)
- [The old pain in the DAS « StorageTexan's Blog](#)

About Me



Hello, and thank you for visiting VMtoday. My name is Josh Townsend. I am a technology professional with a strong background in VMware Virtualization, Storage, and Microsoft technologies. I am a Sr. Systems Administrator at [Tiber Creek Consulting](#) in Fairfax, VA, and hold several technical certifications, including VMware Certified Professional. I am also a 2010 VMware vExpert.



I am also leader of the [Washington DC Metro Area VMware User Group \(VMUG\)](#).



The opinions expressed on this site are my own and may not reflect the views of my employer, VMware, or any other party unless otherwise stated.

Please feel free to follow me on Twitter

[@joshuatownsend](#)

Archives

- [July 2010](#) (2)
- [June 2010](#) (1)
- [May 2010](#) (1)
- [April 2010](#) (2)
- [March 2010](#) (3)
- [February 2010](#) (1)
- [January 2010](#) (4)
- [December 2009](#) (3)
- [November 2009](#) (3)
- [September 2009](#) (7)
- [August 2009](#) (2)
- [July 2009](#) (1)
- [June 2009](#) (6)
- [March 2009](#) (4)
- [January 2009](#) (4)
- [December 2008](#) (2)
- [November 2008](#) (2)

VMware Related Sites

- [boche.net – VMware Virtualization Evangelist](#)
- [Chris Wolf](#)
- [Gages Virtual World](#)
- [Gerben's Blog On Virtualization](#)
- [IPMer](#)
- [Ken's Virtual Realty](#)
- [Mike Laverick's RTFM Eduction](#)
- [Planet V12n](#)
- [Rene Jorissen's Booches.nl](#)
- [Run-Virtual](#)
- [Scott Lowe's Blog](#)
- [The VMguy](#)
- [vinternals](#)
- [Virtu-AI](#)
- [Virtualization.info](#)
- [VM /ETC](#)
- [VM Guru](#)
- [VMtoday Virtualization Resources](#)
- [VMware](#)
- [VMware Land](#)
- [VMware vLaunchpad](#)
- [vSphere Land \(Eric Siebert\)](#)

Virtualization Jobs

Virtualization Jobs

[Server Administrator](#)

[New Albany, IN - Floyd Memorial
Hospital & Health Services](#)

[DAS - SDCOE CLOUD](#)

[Infrastructure Architect](#)

[Louisville, KY - IBM](#)

[DRA - Advisory \(SCON\) IT
Architect](#)

[Louisville, KY - IBM](#)

[DRA - Sr. \(SCON\) IT Architect](#)

[Louisville, KY - IBM](#)

[DRA - Executive \(SCON\) IT
Architect](#)

[Louisville, KY - IBM](#)

[System x FTSS](#)

[Louisville, KY - IBM](#)

[Custom Migration Windows
Architect](#)

[Louisville, KY - IBM](#)

[Field Technical Sales Specialist
\(Must Reside In the Boston /
Hartford Area\)](#)

[Louisville, KY - IBM](#)

[Midrange Storage Virtualization
Solutions Technical Consultant](#)

[Louisville, KY - IBM](#)


[Server Systems Administrator
Jeffersonville, IN - Heartland
Payment Systems](#)

[See All Jobs](#)


[Post a job for \\$25](#)

Jobs by **SimplyHired** [+ Get Widget](#)


Virtualization Resources




[Windows Server 2008 H...](#)
Rand Morimoto, Jeff Guillet...
\$32.84
★★★★☆




[Windows Server 2008 H...](#)
Robert Larson, Janique Ca...
\$37.79
★★★★★




[Windows Server 2008 H...](#)
John Kelbley, Mike Sterlin...
\$31.49
★★★★★




[Win Svr Std with o Hyp...](#)
Microsoft Software
\$1,169.99




[POWERCHUTE NETWORK...](#)
APC
\$86.99



[virtualization.info](#)
Alessandro Perilli
\$0.99




[MCTS: Windows Server ...](#)
William Panek (Paperback ...
\$40.49
★★★★★



[Microsoft Virtualization...](#)
Jason Kappel, Anthony Vel...
\$42.74

1 2 3 4 >



[Privacy](#)

Powered by [WordPress](#) and [WordPress Theme](#) created with Artisteer by [Josh Townsend](#).