

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

## VMtoday

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

« [Washington, DC VMUG – March 10, 2010](#)

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

### Storage Basics – Part V: Controllers, Cache and Coalescing

March 23rd, 2010 | Author: [Joshua Townsend](#)

My Storage Basics series has been neglected for some time (sick kids, snow storms, VMware Upgrades, SAN implementations and some Cisco switch upgrades took all my free time), so let's jump right in to Part V – Cache, Controllers, and Coalescing. Between the alliteration and fancy words, it might seem like I am about to tell a tale of international espionage. Unfortunately, my introductory treatment of these aspects of a storage system will probably not keep you on the edge of your seat – but I'll try to keep it interesting.

[2](#)  
[tweets](#)  
[retwee](#)

Throughout this series, we've been working our way from the basic building block of any storage system – the disks – outwards towards the brains of the operation – the controller. You'll recall that in [Part II](#) I introduced IOPS and the math that goes into calculating the IOPS capacity of a disk array. In [Part III](#) we considered a RAID implementation's impact on performance and availability. And most recently in [Part IV](#) we looked at the common interface types when dealing with storage arrays. If we put the previous parts together we still don't have a functional storage system. The missing piece is the controller. Simply put, the storage controller is the hardware adapter between the disks and the servers that connect to the storage. The controller has a specific '[interface](#)' type, is responsible for RAID operations, and handles advanced storage functionality. A controller can be as simple as the Dell PERC or HP Smart Array add-in card on your server, or as complex as the Storage Processor in an enterprise class Storage Area Network (SAN) such as an EMC CLARiiON or NetApp FAS.

#### **Controllers**

As we look at controllers and the advanced features they provide we'll see that some of the earlier performance equations start to break down. The simplest controllers take disk read/write commands from the operating system and send commands down to the disk(s) attached to be read or written. This gets data onto the disk, but often does not do so in an efficient or reliable manner. RAID-capable controllers take on the added responsibility of configuring disks in the desired RAID level, calculating & writing parity data, and writing the data in disk-spanning stripes or mirrors depending on the RAID level.

#### **Cache**

To increase performance and improve reliability, storage vendors implement a caching system on their controllers. Cache is memory that acts as a buffer for disk I/O, and is usually battery-backed to prevent data loss in the event of a power failure. Because of the exponentially greater speed of RAM over spinning magnetic disks, cache can improve performance by orders of magnitude. Cache can operate on both reads and writes to disk.

When dealing with writes, the controller cache is typically used in one of two ways: write-through or write-back. In write-through mode, data is written to volatile cache and then to disk, and only acknowledged as written once the data resides on the non-volatile disk. Write-back mode allows the controller to acknowledge the data as having been written as soon as it is held in cache. This allows the cache to buffer writes quickly and then write them to the slower disk when the disk has cycles to accept I/O. The greater your cache size, the more data that can be buffered, ultimately resulting in better performance as measured in both IOPS and throughput. [This graph](#) from my [article on troubleshooting write performance on an IBM DS3300 iSCSI array](#) shows how throughput increased and latency decreased when enabling write cache. The extent to which cache increases performance is highly dependent on the workload characteristics (I/O size, randomness, and ratio of reads:writes).

Read-cache acts as a buffer for reads in a couple ways. First, some controllers attempt to ‘read-ahead’, anticipating future read requests from the operating system and buffering what it expects to be the next blocks of desired data. Some entry-level controllers simply buffer the next physical chunk of data and fill cache memory with it, while more advanced controllers may attempt to predict the right block of data based on previous requests (you just asked for 3 blocks in a row, I’m guessing you’ll come asking for the 4th next so I’ll just buffer it in fast cache for you now). Secondly, read cache holds data that has been previously read, regardless of any pre-fetching the controller may have done. This allows for much faster subsequent access of the same data because it is held in the faster cache, eliminating the need for the controller to go to disk for the data again. Just like with write cache, the extent to which cache increases performance is highly dependent on the workload characteristics.

A given storage array controller only has so much cache to work with. A Dell PERC5/E, for example, has 256MB of cache that can be used for both read and write. While this may be enough for a direct-attached storage array, SAN’s serving multiple systems demand more cache. In contrast, an EMC CLARiiON [CX4-960](#) has 32GB. Some storage vendors, such as NetApp, are getting creative with cache. NetApp’s [Performance Acceleration Module \(PAM\)](#) is an add-in card that provides up to a whopping [512GB](#) of Layer 2 cache to the storage system.

Caching mechanisms can dramatically influence performance under the right conditions. With healthy cache in place, IOPS calculations become skewed. However, cache can be exhausted or may not hold the data you are interested in. If cache is insufficient to satisfy read requests, or has reached its high-water mark for writes, performance can drop off. When cache is exhausted, the backing disk must be able to satisfy the I/O workload or performance will be unacceptable. This is where the IOPS calculations kick in, and where having the right disk type and configuration really matters.

## Queuing & Coalescing

Advanced storage systems introduce additional features to reduce I/O contention and improve cache utilization. I won’t go into all of the features here because they vary by storage vendor. However, I will point out two common techniques – queuing and coalescing.

Queuing refers to the ability of a storage system to queue storage commands for later processing. Queuing can take place at various points in your storage environment, from the HBA to the storage processor/controller. A little queuing may be OK depending on your workload, but too many outstanding I/Os can negatively impact performance (this is measured in latency). Queue depths can be adjusted on many components in your storage and VMware landscape, but check with your vendor’s support group before you make changes to these settings.

Coalescing is performed by some storage systems to modify the character of the workload. To better understand coalescing, picture a bunch of random write activity. Without cache in place, the disk heads will be bouncing all over the platters trying to get the data on to disk. A little write cache will allow the storage array to acknowledge the write for the OS, but the array still needs to de-stage that data from cache to disk quickly to prevent cache exhaustion. The back-end disks will still be doing the chicken dance, bouncing around trying to write the random workload.... Now picture an intelligent system that re-orders the random writes that are held in cache and writes them to the disk in nice sequential stripes. The disk heads will be less prone to jumping around the platter and the behavior will start to look more like a nice waltz than the funky chicken dance. Coalescing is used for writes, not reads, so not all workloads benefit.

## Wrap-up

With this article on Controllers, Cache, and Coalescing we’ll end our look at the basic building blocks of a storage array. Before we end the Storage Basic series I am planning a few more articles on Storage Workload Characterization (which has been mentioned, but not directly addressed in this and previous articles), Identifying a Stressed Storage System, and Best Practices for Storage Performance in a VMware Environment.

If you are interested in more reading on Controllers, Cache, and Coalescing, I recommend the following:

### Additional Reading:

- [Impact of cache on the performance of the HP StorageWorks XP12000 Disk Array white paper](#)
- [Performance impact of controller cache: SQL Server read only workloads](#)
- [IOPS?](#) – Dig into the article’s comments for some great dialog between some people who really know their stuff!
- [Storage Performance for SQL Server](#)
- [Storage Caching 101 – Chuck Hollis \(EMC\)](#)
- [Improving Performance with Interrupt Coalescing for Virtual Machine Disk IO in VMware ESX Server](#)

**Related posts:**

1. [Storage Basics – Part VI: Storage Workload Characterization](#)
2. [Storage Basics – Part III: RAID](#)
3. [Storage Basics – Part I: An Introduction](#)
4. [Storage Basics – Part IV: Interface](#)
5. [Storage Basics – Part VII: Storage Alignment](#)

Posted in [EMC](#), [General IT](#), [NetApp](#), [Storage](#), [Storage Basics](#), [VMware](#) | Tags: [disk](#), [EMC](#), [ESX](#), [I/O](#), [NetApp](#), [performance](#), [SAN](#), [Storage](#), [VMware](#)

 [4 Tweets](#)

**2 Responses to “Storage Basics – Part V: Controllers, Cache and Coalescing”**

- [joshuatownsend](#):  
[March 24, 2010 at 7:48 am](#)

New post on VMtoday.com: Storage Basics – Part V: Controllers, Cache and Coalescing – <http://bit.ly/d4Xg4C>

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

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

- [Tweets that mention Storage Basics – Part V: Controllers, Cache and Coalescing / VMtoday -- Topsy.com](#):  
[March 24, 2010 at 10:08 am](#)

[...] This post was mentioned on Twitter by Todd Scalzott, joshuatownsend. joshuatownsend said: New post on VMtoday.com: Storage Basics – Part V: Controllers, Cache and Coalescing – <http://bit.ly/d4Xg4C> [...]

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

**Leave a Reply**

Name (required)

Mail (will not be published) (required)

Website

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](#)

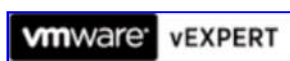
- [Windows Guest Defragmentation, Take Two « Pivot Point](#)
- [Dan Weiss Blogsite » NetApp performance testing engineer tells it like it is](#)
- [Understanding how storage design has a big impact on your VDI! - Ruben Spruijt - BrianMadden.com](#)
- [Performance Troubleshooting VMware vSphere – Storage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

keywords	location
<input type="button" value="Go"/>	
Jobs by <a href="#">SimplyHired</a> <a href="#">+ Get Widget</a>	
<h3>Virtualization Resources</h3>	
	<a href="#">Windows Server 2008 H...</a> John Kelbley, Mike Sterlin... <b>\$31.49</b> ★★★★★
	<a href="#">Windows Server 2008 H...</a> Rand Morimoto, Jeff Guillet... <b>\$32.84</b> ★★★★☆
	<a href="#">Windows Server 2008 H...</a> Robert Larson, Janique Ca... <b>\$37.79</b> ★★★★★
	<a href="#">Win Svr Std with o Hyp...</a> Microsoft Software <b>\$1,169.99</b>
	<a href="#">virtualization.info</a> Alessandro Perilli <b>\$0.99</b>
	<a href="#">MCTS: Windows Server ...</a> William Panek (Paperback ... <b>\$40.49</b> ★★★★★
	<a href="#">POWERCHUTE NETWORK...</a> APC <b>\$86.99</b>
	<a href="#">Microsoft Virtualization...</a> Jason Kappel, Anthony Vel... <b>\$42.74</b>
1 2 3 4 >	
<div> <a href="#">Privacy</a></div>	

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