TechRepublic : A ZDNet Tech Community

Calculate IOPS in a storage array

Date: February 12th, 2010 Author: Scott Lowe Category: storage Tags: Disk, Performance, I/O, Storage Array, Calculation, Scott Lowe, RAID 6, Transport Choice, RAID, Storage

What drives storage performance? Is it the iSCSI/Fiber Channel choice? The answer might surprise you. Scott Lowe provides insight into IOPS.

When it comes to measuring a storage system's overall performance, Input/Output Operations Per Second (IOPS) is still the most common metric in use. There are a number of factors that go into calculating the IOPS capability of an individual storage system.

In this article, I provide introductory information that goes into calculations that will help you figure out what your system can do. Specifically, I explain how individual storage components affect overall IOPS capability. I do not go into seriously convoluted mathematical formulas, but I do provide you with practical guidance and some formulas that might help you in your planning. Here are three notes to keep in mind when reading the article:

Published IOPS calculations aren't the end-all be-all of storage characteristics. Vendors often measure IOPS under only the best conditions, so it's up to you to verify the information and make sure the solution meets the needs of your environment.

IOPS calculations vary wildly based on the kind of workload being handled. In general, there are three performance categories related to IOPS: random performance, sequential performance, and a combination of the two, which is measured when you assess random and sequential performance at the same time.

The information presented here is intended to be very general and focuses primarily on random workloads.

IOPS calculations

Every disk in your storage system has a maximum theoretical IOPS value that is based on a formula. Disk performance — and IOPS — is based on three key factors:

Rotational speed (aka spindle speed). Measured in revolutions per minute (RPM), most disks you'll consider for enterprise storage rotate at speeds of 7,200, 10,000 or 15,000 RPM with the latter two being the most common. A higher rotational speed is associated with a higher performing disk. This value is not used directly in calculations, but it is highly important. The other three values depend heavily on the rotational speed, so I've included it for completeness.

Average latency. The time it takes for the sector of the disk being accessed to rotate into position under a read/write head.

Average seek time. The time (in ms) it takes for the hard drive's read/write head to position itself

over the track being read or written. There are both read and write seek times; take the average of the two values.

To calculate the IOPS range, use this formula: Average IOPS: Divide 1 by the sum of the average latency in ms and the average seek time in ms (1 / (average latency in ms + average seek time in ms).

Sample drive:

Model: Western Digital VelociRaptor 2.5" SATA hard drive Rotational speed: 10,000 RPM Average latency: 3 ms (0.003 seconds) Average seek time: 4.2 (r)/4.7 (w) = 4.45 ms (0.0045 seconds) Calculated IOPS for this disk: 1/(0.003 + 0.0045) = about 133 IOPS

So, this sample drive can support about 133 IOPS. Compare this to the chart below, and you'll see that the value of 133 falls within the observed real-world performance exhibited by 10K RPM drives.

However, rather than working through a formula for your individual disks, there are a number of resources available that outline average observed IOPS values for a variety of different kinds of disks. For ease of calculation, use these values unless you think your own disks will vary greatly for some reason.

Below I list some of the values I've seen and used in my own environment for rough planning purposes. As you can see, the values for each kind of drive don't radically change from source to source.

Average IOPS/drive				
Drive speed	Source 1	Source 2	Source 3	
7200 RPM	75-100	75	30-80	
10K RPM	100-130	125	140	
15K RPM	150-190	175	180	

Sources:

http://blog.aarondelp.com/2009/10/its-now-all-about-iops.html http://www.yellow-bricks.com/2009/12/23/iops/ http://www.tomshardware.com/forum/251893-32-raid-raid

Note: The *drive type* doesn't enter into the equation at all. Sure, SAS disks will perform better than most SATA disks, but that's only because SAS disks are generally used for enterprise applications due to their often higher reliability as proven through their mean time between failure (MTBF) values. If a vendor decided to release a 15K RPM SATA disk with low latency and seek time values, it would have a high IOPS value, too.

Multidisk arrays

Enterprises don't install a single disk at a time, so the above calculations are pretty meaningless unless they can be translated to multidisk sets. Fortunately, it's easy to translate raw IOPS values from single disk to multiple disk implementations; it's a simple multiplication operation. For example, if you have ten 15K RPM disks, each with 175 IOPS capability, your disk system has 1,750 IOPS worth of performance capacity. But this is only if you opted for a RAID-0 or just a bunch of disks (JBOD) implementation. In the real world, RAID 0 is rarely used because the loss of a single disk in the array would result in the loss of all data in the array.

Let's explore what happens when you start looking at other RAID levels.

The IOPS RAID penalty

Perhaps the most important IOPS calculation component to understand lies in the realm of the write penalty associated with a number of RAID configurations. With the exception of RAID o, which is simply an array of disks strung together to create a larger storage pool, RAID configurations rely on the fact that write operations actually result in multiple writes to the array. This characteristic is why different RAID configurations are suitable for different tasks.

For example, for each random write request, RAID 5 requires many disk operations, which has a significant impact on raw IOPS calculations. For general purposes, accept that RAID 5 writes require 4 IOPS per write operation. RAID 6's higher protection double fault tolerance is even worse in this regard, resulting in an "IO penalty" of 6 operations; in other words, plan on 6 IOPS for each random write operation. For read operations under RAID 5 and RAID 6, an IOPS is an IOPS; there is no negative performance or IOPS impact with read operations. Also, be aware that RAID 1 imposes a 2 to 1 IO penalty.

The chart below summarizes the read and write RAID penalties for the most common RAID levels.

I/O Impact				
RAID level	Read	Write		
RAID 0	1	1		
RAID 1 (and 10)	1	2		
RAID 5	1	4		
RAID 6	1	6		

Parity-based RAID systems also introduce other additional processing that result from the need to calculate parity information. The more parity protection you add to a system, the more processing overhead you incur. As you might expect, the overall imposed penalty is very dependent on the balance between read and write workloads.

A good starting point formula is below. This formula does not use the array IOPS value; it uses a workload IOPS value that you would derive on your own or by using some kind of calculation tool, such as the Exchange Server calculator.

(Total Workload IOPS * Percentage of workload that is read operations) + (Total Workload IOPS * Percentage of workload that is read operations * RAID IO Penalty)

Source: http://www.yellow-bricks.com/2009/12/23/iops/

As an example, let's assume the following:

Total IOPS need: 250 IOPS

Read workload: 50% Write workload: 50% RAID level: 6 (IO penalty of 6)

Result: You would need an array that could support 875 IOPS to support a 250 IOPS RAID 6-based workload that is 50% writes.

This could be an unpleasant surprise for some organizations, as it indicates that the number of disks might be more important than the size (i.e., you'd need twelve 7,200 RPM, seven 10K RPM, or five 15K RPM disks to support this IOPS need).

The transport choice

It's also important to understand what is not included in the raw numbers: the transport choice iSCSI or Fibre Channel. While the transport choice is an important consideration for many organizations, it doesn't directly impact the IOPS calculations. (None of the formulas consider the transport being used.)

If you want more proof that the iSCSI/Fibre Channel choice doesn't necessarily directly impact your IOPS calculations, read this article on NetApp's site.

The transport choice *is* an important one, but it's not the primary choice that many would make it out to be. For larger organizations that have significant transport needs (i.e., between the servers and the storage), Fibre Channel is a good choice, but this choice does not drive the IOPS wagon.

Summary

In order to intricately understand your IOPS needs, you need to know a whole lot, including specific disk technicalities, your workload breakdown as a function of read vs. write, and the RAID level you intend to use. Once you implement your solution, you can use tools that are tailor-made to IOPS analysis, such as Iometer, to get specific, real-time performance values. This assumes that you have a solution in place that you can measure.

If you're still in the planning stages or a deep level of analysis simply isn't necessary for your needs, the generalities presented in this article will help you figure out your needs.

Want to keep up with Scott Lowe's posts on TechRepublic?

Automatically sign up to the Servers and Storage newsletter Subscribe to the Servers and Storage RSS feed Follow Scott Lowe on Twitter



Scott Lowe has spent 15 years in the IT world and is currently the Chief Information Officer for Westminster College in Fulton, Missouri. He is also a regular contributor to TechRepublic and has authored one book, Home Networking: The Missing Manual (O'Reilly), and coauthored the Microsoft Exchange Server 2007 Administrator's Companion (MS Press). Read his full bio and profile.

People who read this, also read...

Extend the capabilities of STATSPACK to measure Oracle disk I/O Discover the benefits of java.nio's new features Database administration: Causes of Oracle buffer busy waits i/o baseplate for Asrock P4VM800 HP DL580 G5 I/O

Print/View all posts Comments on this blog

Time for IO != Latency *plus* Seek thattommyhall@... | 02/15/10 RE: Calculate IOPS in a storage array - numerous flaws david@... | 02/15/10 Errors do not have that great an impact zackers | 02/15/10The two are independent zackers | 02/15/10 mmmm NEW thattommyhall@... | 02/16/10 "General Information" "focused" on random I/O micheldufrenoy@... | 02/15/10 You need to consider what happens at the storage level zackers | 02/15/10General workloads Scott Lowe | 02/15/10 Mean-time between failure misunderstanding micheldufrenoy@... | 02/15/10 Unrecoverable error instead? Scott Lowe | 02/15/10 RE: Calculate IOPS in a storage array The Igneous Group | 02/15/10RE: Calculate IOPS in a storage array jekidd@... | 02/15/10 Elevator operations NEW oldbaritone | 02/16/10

My Updates My Contacts

Popular on CBS sites: College Signing Day | March Madness | Lost | iPhone | Cell Phones | Video Game Reviews | Free Music

About CBS Interactive | Jobs | Advertise | Mobile | Site Map

© 2010 CBS Interactive Inc. All rights reserved. | Privacy Policy (updated) | Terms of Use