PowerVault MD1200/MD1220 Storage Solution Guide for Applications

A Dell Technical White Paper

Dell PowerVault Storage Systems

Joe Noyola
Systems Performance Analysis Group

Chuck Colburn
Storage Advanced Engineering

Vamsee Kasavajhala
Dell PowerVault Technical Marketing

May 2011

ABSTRACT

One of the most important factors to consider when implementing a storage solution is the type of application workload that will utilize the storage infrastructure. To gain the maximum benefit from your investment in Dell's PowerVault™ MD1200/1220 storage solution, it is important to optimize this solution for the application workload it will be serving.

Selecting the right components and design for the PowerVault MD1200/1220 will not only impact the performance and effectiveness of your application, but it will also improve the efficacy of the entire IT infrastructure that relies on that storage. The goal of this white paper is to provide application workload-based recommendations and guidelines on the storage components to use when implementing PowerVault MD1200/1220 DAS solutions.

Contents

Introduction

There are several factors to consider when selecting an enterprise storage solution. These include reliability, availability, manageability, and price. However, an equally important factor is the application utilizing the storage system and its performance requirements, such as input/output (I/O) load and data usage patterns. It is important to ensure that the selected storage system meets these criteria.

Components such as disk type (hard disk drive, or HDD, versus solid state drive, or SSD), RAID configuration, and adapter parameter settings determine the performance of a storage system. When implementing a storage system, therefore, it is important to quantify the performance impact of different components and select those that match the performance requirements of the application.

To help customers determine the optimal PowerVault MD1200/1220 configuration for each application type, Dell Labs has tested various solution configurations by simulating I/O loads from many widely used applications. Results from these tests clearly show that different combinations of PowerVault enclosure, drive type, RAID level, and PERC H800 cache configuration provide different levels of I/O performance for each application type, depending upon its I/O characteristics.

Table 1 provides a description of the hardware configurations tested in this study.

Table 1. Hardware Configurations

PowerEdge Server	PowerVault Enclosure (Supported Drive Form Factor)	PERC Adapter	Traditional Hard Drives				Solid State Storage		
			7.2K RPM NL-SAS (# Drives)	15K RPM SAS (# Drives)	PERC H800 Cache Settings		SSD* (# Drives)	PERC H800 Cache Settings	
					Read	Write		Read	Write
R610	MD1220 (2.5in)	H800	500GB (24)	146GB (24)	Adaptive Read Ahead	Write Back	149GB (8)	No Read Ahead	Write Through
	MD1200 (3.5in)		2TB (12)	600GB (12)					

*Solid State Storage measured in this study was Dell's Enterprise Level SLC Solid State Drives. Performance of this drive type far exceeds consumer branded MLC SSD drives.

Storage Solution Recommendations

At Dell Labs, the hardware configurations shown in Table 1 were subject to extensive IOmeter testing under different simulated application loads. Performance data was collected from various simulation runs of RAID 5, RAID 6, and RAID 10 on the PERC H800 adapter directly attached to either an MD1200 or MD1220 expansion enclosure with 3 disk drive types: 7.2K RPM nearline SAS (NL-SAS), 15K RPM SAS, and 149GB Pliant SSD.

To maximize performance with traditional hard drives, each RAID implementation used the maximum number of drives that could fit in each PowerVault enclosure, resulting in 12 HDDs per volume (virtual drive) in MD1200 and 24 HDDs per volume (virtual drive) in MD1220. In contrast, only eight SSDs per volume (virtual drive) were required in both enclosures to achieve maximum performance. In order to record maximum performance, we collected the performance metrics from virtual drives left uninitialized/not formatted as NT file systems.

In Table 2, we have grouped application types that share similar I/O characteristics, and listed each category's I/O load profile and the metric by which we measured that profile.

Table 2. Application Categories Based on I/O Workload Profile and Their I/O Metric

Application Category	Examples	I/O Workload Profile	I/O Metric*
A	Web Hosting	Small Random Reads	IOPS
B	Database Server Small Departmental E-mail Server	Random Reads and Writes	IOPS
C	Web and Database Logs	Sequential Writes	MBPS
D	Video On Demand	Large Random Reads	MBPS
E	Streaming Media	Large Sequential Reads	MBPS
*Metric measurements: IOPS = I/O operations per second MBPS = Megabytes per second			

We used IOmeter simulation runs to compare the performance of various storage components such as enclosure, drive type, and RAID level for each application category listed in Table 2. Pertinent storage solution information about RAID level, enclosure type, PERC adapter settings, and drive type was noted for the storage solution that maximized the recommended I/O metric. The following sections describe our findings and recommendations for each application category listed in Table 2.

Application Category A: Small Random Reads

Applications in this category include web hosting, where multiple users request multiple web pages simultaneously. These applications usually have a large number of small random read operations in their I/O load profile, so IOPS is the I/O metric to maximize. The traffic handled by Web hosting applications ranges from intranet sites with minimal internal company traffic to global internet portals that receive page requests from hundreds of thousands of users daily. Large global web sites usually have several web servers responding to client requests, which results in redundancy of PERC H800 adapters and data-paths. Different RAID levels also provide this level of redundancy. Moreover, small to mid-size web hosting applications may not require significant disk capacity scalability, since typical

website content is usually small—less than a few terabytes (TBs)—and relatively static, with only minor content modifications. Table 3 provides the typical I/O workload profile for this application category.

Table 3. Typical Small Random Reads I/O Workload Profile

Application Category A	I/O Workload Profile		
	Read/Write Percentage	Random/Sequential Percentage	Size in KB
Web File Server	95%/5%	75%/25%	4KB,8KB,16KB

Figure 1. Simulation Results for Application Category A: Small Random Reads

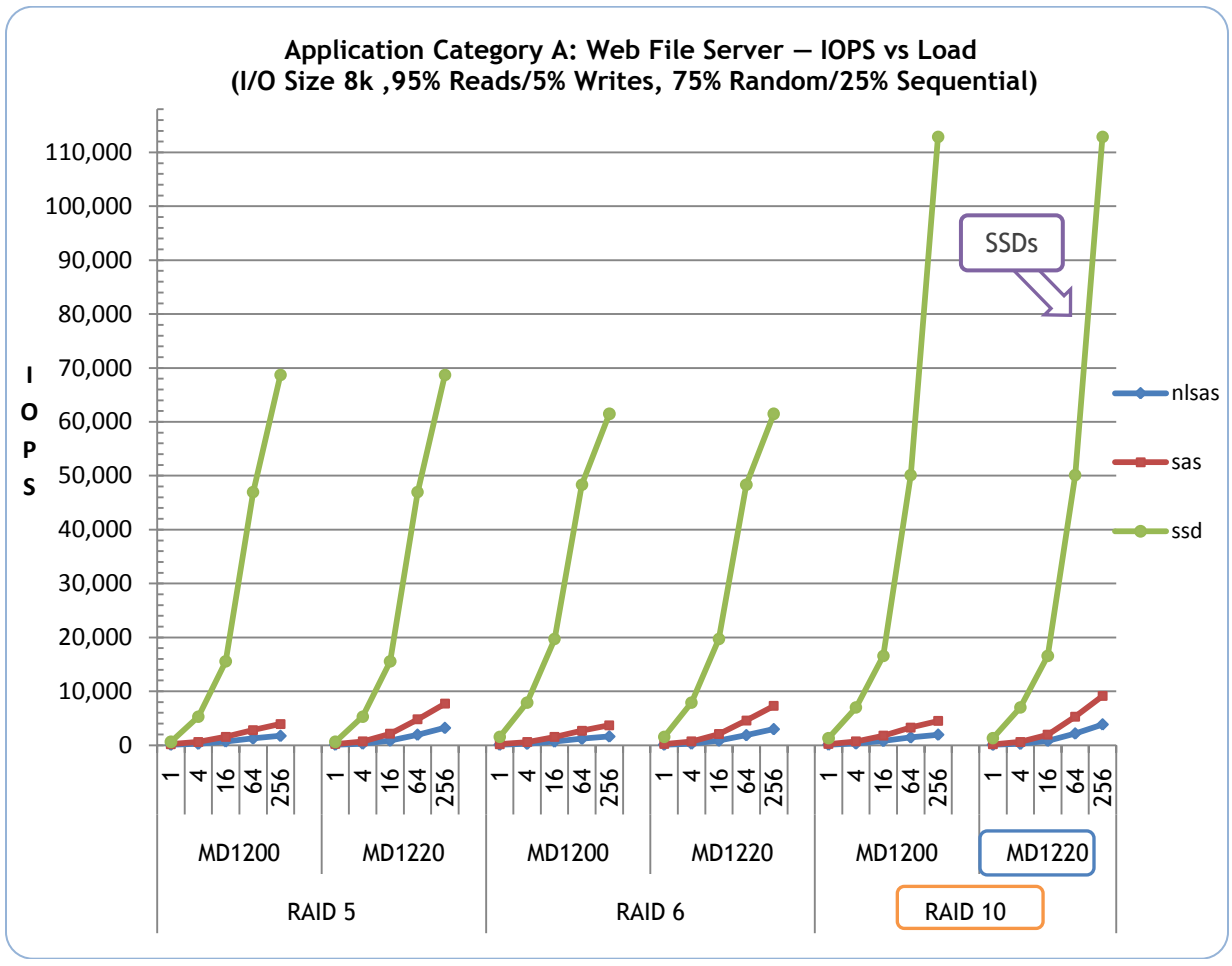

Table 4 lists the IOmeter simulation results of application category A graphed in Figure 1 for all components (RAID, enclosure, drive type) at a maximum load of 256 (queue depth).

Table 4. Simulation Results for Application Category A: Small Random Reads at Maximum Load of 256

RAID Type	Enclosure	Load (Queue Depth)	Max IOPS (8KB Block Size)		
			NL-SAS (# Drives)	SAS (# Drives)	SSD (# Drives)
RAID 5	MD1200	256	1727 (12)	3942 (12)	68691 (8)
	MD1220	256	3207 (24)	7719 (24)	68691 (8)
RAID 6	MD1200	256	1621 (12)	3701 (12)	61486 (8)
	MD1220	256	2979 (24)	7278 (24)	61486 (8)
RAID 10	MD1200	256	1955 (12)	4512 (12)	112889 (8)
	MD1220	256	3859 (24)	9116 (24)	112889 (8)

Based on these results, the following are the guidelines and recommendations for selecting the enclosure, drive type, and RAID level for small random reads, application category A.

Application Category A (Small Random Reads) Recommended Storage Configuration		
Component	**Recommendation**	**Rationale**
Enclosure	MD1220 Maximum Spindles Best Performance	For maximum performance, it is beneficial to have more drives, since data can be spread across them and data access times are reduced. The Powervault MD1220 supports up to 24 drives, while the MD1200 supports up to 12. The MD1220's performance (the number of IOPS) is almost twice that of the MD1200 (see Table 4).
Drive Type	SSDs* Maximum Performance	The advantage of SSDs for random workloads is that they achieve tremendous performance with fewer drives. Depending upon the RAID configuration, 8 SSDs provided up to 12 times better performance (number of IOPS) than 24 HDDs in the same MD1220 enclosure (see Table 4).
RAID Level	10	RAID 10 offers the maximum performance for small random reads, application category A (see Table 4). However, the end user can choose the appropriate RAID level depending on the number of drives and the redundancy level requirements of data stored on the drives. Refer to Appendix B: RAID Guide for more details.
*Users should factor in the higher cost of SSDs when considering a SSD-based solution.		

Application Category B: Random Reads and Writes

Applications in this category include e-mail and database servers.

E-mail server storage requirements can vary depending on the size and the number of e-mails and the type and number of users. While a storage solution with a small capacity and limited features may work well for small departmental e-mail servers, large corporate e-mail servers normally require greater storage capacity, higher availability, and better performance and scalability. I/O profiles will vary, depending on the number and type of users and the number and size of e-mails and their attachments.

Database servers can range from simple workgroup databases, like Microsoft® Access™ with a few hundred users, to critical enterprise databases, like Oracle or SQL Server with thousands of users. Generally, the more critical the database, greater the need for data protection. Additionally, performance requirements increase in proportion to the number of users accessing the database. Many industries use online transaction processing (OLTP) applications for entry and retrieval of business information. Such application servers are business critical and require a combination of performance (IOPS) and data redundancy. Table 5 provides the typical I/O workload profile for this application category.

Table 5. Typical Random Reads and Writes I/O Workload Profile

Application Category B	I/O Workload Profile		
	Read/Write Percentage	Random/Sequential Percentage	Size in KB
OLTP Database	70%/30%	100% Random	8KB
Small Departmental Exchange E-mail	70%/30%	100% Random	4KB

Figure 2. Simulation Results for Application Category B: Random Reads and Writes

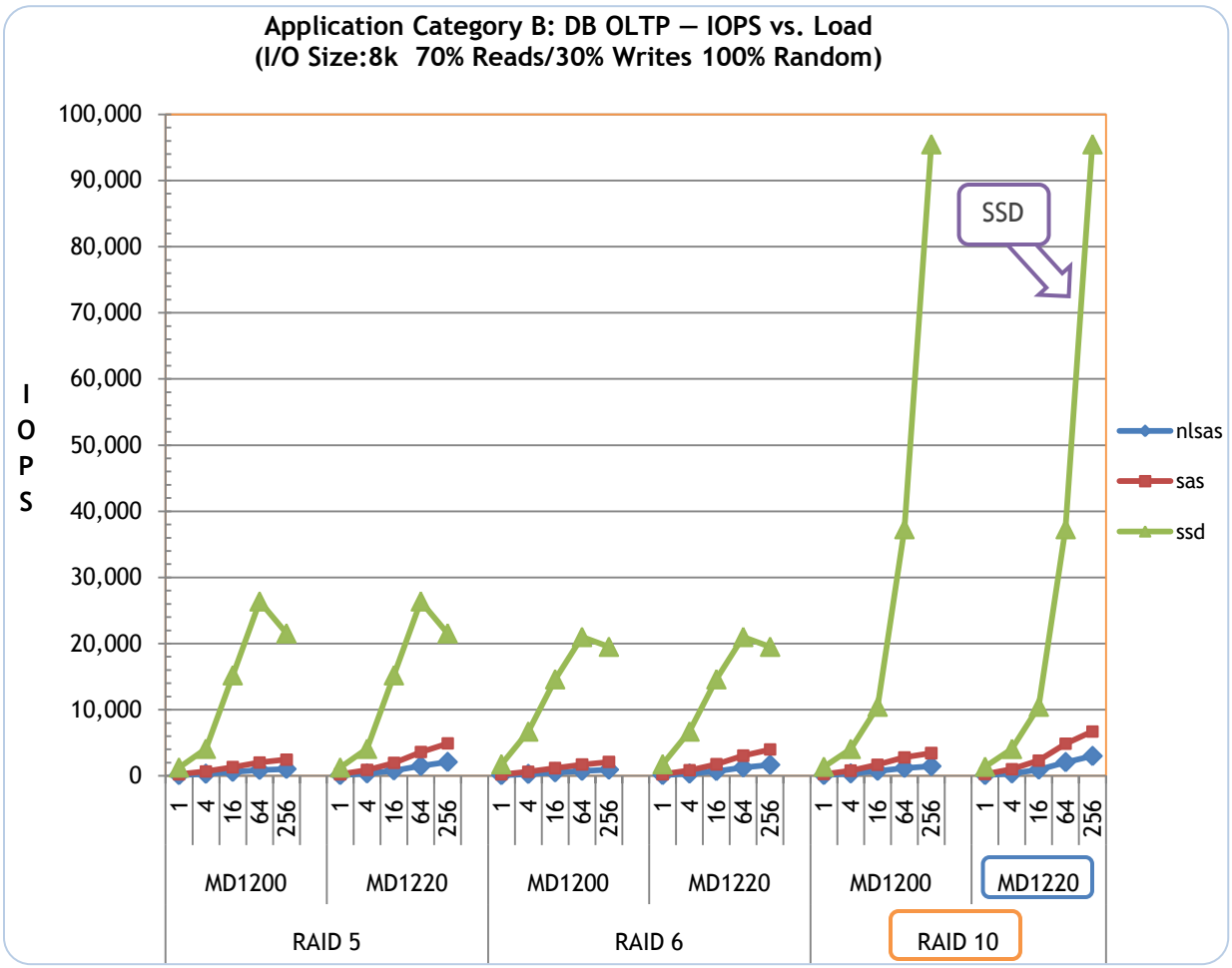

Table 6 lists the IOmeter simulation results of application category B graphed in Figure 2 for all components (RAID, enclosure, drive type) at maximum I/O load of 256 (queue depth).

Table 6. Simulation Results for Application Category B: Random Reads and Writes at Maximum Load of 256

RAID Type	Enclosure	Load (Queue Depth)	Max IOPS (8KB Block Size)		
			NL-SAS (# Drives)	SAS (# Drives)	SSD (# Drives)
RAID 5	MD1200	256	1031 (12)	2460 (12)	21549 (8)
	MD1220	256	2090 (24)	4907 (24)	21549 (8)
RAID 6	MD1200	256	927 (12)	1994 (12)	19524 (8)
	MD1220	256	1665 (24)	3983 (24)	19524 (8)
RAID 10	MD1200	256	1453(12)	3446 (12)	95514 (8)
	MD1220	256	2991 (24)	6704 (24)	95514 (8)

Based on these results, the following are the guidelines and recommendations for selecting the enclosure, drive type, and RAID level for random reads and writes, application category B.

Application Category B (Random Reads And Writes) Recommended Storage Configuration		
Component	**Recommendation**	**Rationale**
Enclosure	MD1220	For e-mail and database servers where performance and power efficiency are high priority, the Powervault MD1220 is our recommended enclosure. The MD1220's performance (number of IOPS) is approximately twice that of the MD1200. For e-mail servers where large mailbox capacity is a priority, Powervault MD1200 is a good fit, since it can provide up to 24 TBs of storage.
Drive Type	SSDs* Maximum Performance	For random access based applications, SSD offers significant performance improvements over traditional hard drives. However, they are currently capacity limited. Therefore, for small capacity database and e-mail applications, we recommend SSDs. SSDs can perform up to 40 times better than average HDDs (RAID 10). Exchange mail and database server applications are predominantly random read operations (70% reads) with small packet sizes. These characteristics are a perfect fit for SSDs.
RAID Level	RAID 10	RAID 10 is the recommended RAID level. RAID 10 read performance is similar to that of RAID 5, but RAID 10 has better write performance and enterprise-level redundancy, so multiple drive failures can occur without data loss. Refer to Appendix B: RAID Guide for more information.
*Users should factor in the higher cost of SSDs when considering a SSD based solution.		

Application Category C: Sequential Writes

Applications in this category include web and SQL logs. Web and SQL log servers are mission critical systems that require availability and redundancy to be able to recover from power failure or hardware malfunction.

Web and SQL log servers typically require a large storage capacity to write events from daily transactions. For instance, a database server writes to its log device before serving a transaction. These logs are used to rebuild the database after an event like a database crash or user error. Without the logs, the database may no longer be recoverable and data loss may occur for those transactions that were in flight during the crash or after issuing the user error. Table 7 provides the typical I/O workload profile for this application category.

Table 7. Typical Sequential Writes I/O Workload Profile

Application Category C	I/O Workload Profile		
	Read/Write Percentage	Random/Sequential Percentage	Size in KB
Web Server Log	100% Writes	100% Sequential	8KB
SQL Server Log	100% Writes	100% Sequential	64KB

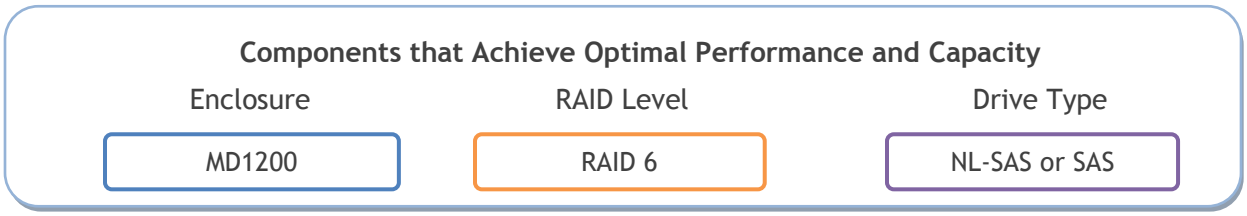

Figure 3. Simulation Results for Application Category C: Sequential Writes

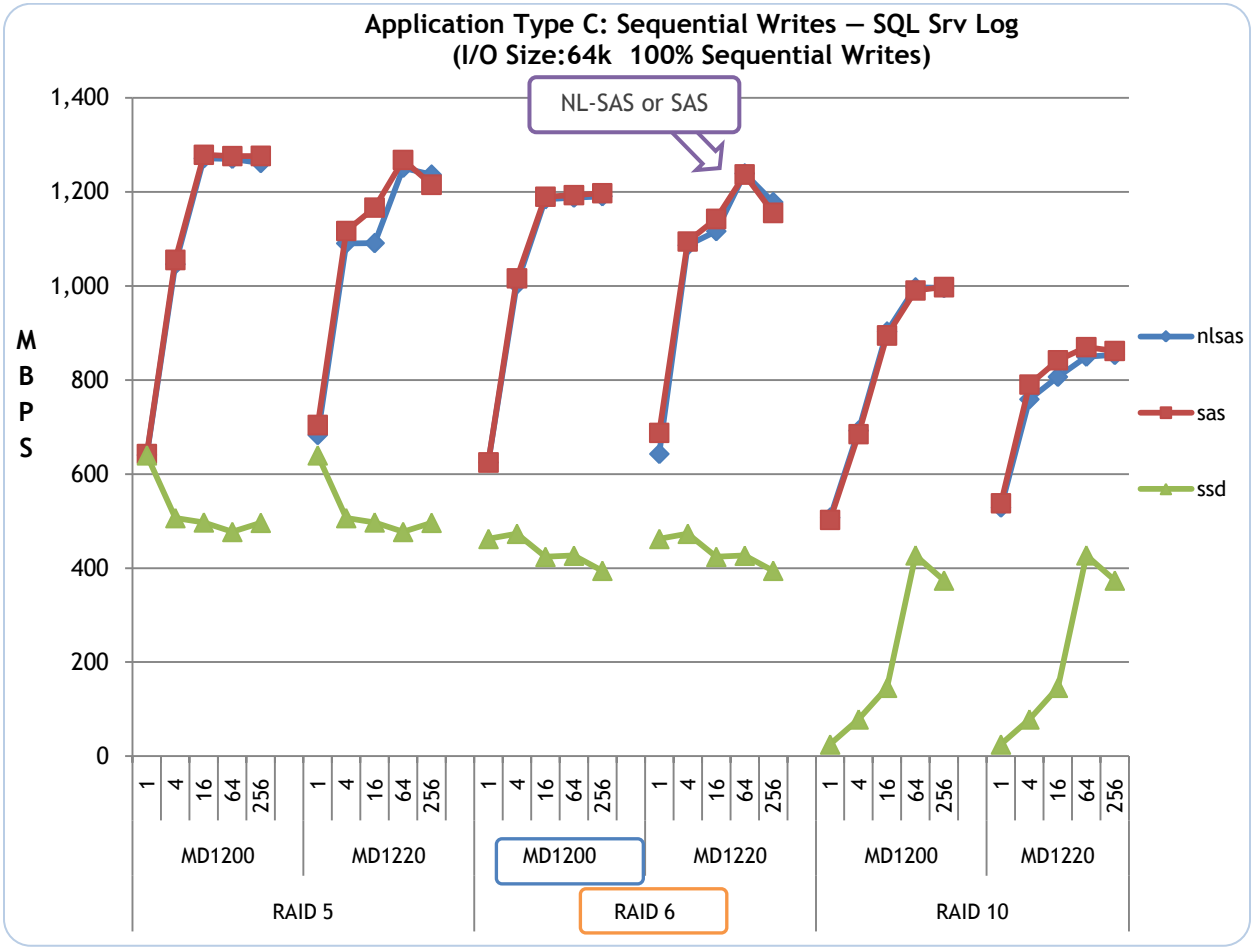

Table 8 lists the IOmeter simulation results of application category C graphed in Figure 3 for all components (RAID, enclosure, drive type) at maximum I/O load of 256 (queue depth).

Table 8. Simulation Results for Application Category C: Sequential Writes at Maximum Load of 256

RAID Type	Enclosure	Load (Queue Depth)	Max IOPS (64 KB Block Size)		
			NL-SAS (# Drives)	SAS (# Drives)	SSD (# Drives)
RAID 5	MD1200	256	1261 (12)	1276 (12)	496 (8)
	MD1220	256	1236 (24)	1215 (24)	496 (8)
RAID 6	MD1200	256	1190 (12)	1197 (12)	394 (8)
	MD1220	256	1178 (24)	1155 (24)	394 (8)
RAID 10	MD1200	256	995 (12)	997 (12)	373 (8)
	MD1220	256	853 (24)	861 (24)	373 (8)

Based on these results, the following are the guidelines and recommendations for selecting the enclosure, drive type, and RAID level for sequential writes, application category C.

Application Category C (Sequential Writes) Recommended Storage Configuration		
Component	**Recommendation**	**Rationale**
Enclosure	MD1200 Large Capacity and Best Streaming Performance —**OR**— MD 1220 Typical Capacity and Performance	For servers where capacity is a priority, Powervault MD1200 is our recommended enclosure. For servers where performance and power efficiency are priorities, the Powervault MD1220 is our recommended enclosure.
Drive Type	HDDs SAS or NL-SAS Combination of Performance and Capacity	Web and SQL server log workloads are sequential in nature. For sequential applications, traditional hard drives provide a potent combination of performance and capacity that cannot be matched by SSDs. HDD performance is twice that of SSD in this category.
RAID Level	RAID 6	If drive count per array is small, RAID 10 is best. If array spreads across six drives or more, RAID 6 will deliver optimal performance. For more detail, refer to Appendix B: RAID Guide.

Application Category D: Large Random Reads

Applications in this category include video on demand (VOD) and decision support system (DSS) servers.

A typical VOD server must be capable of streaming several different movies at the same time. An example is the Netflix online streaming service. DSS is used by individuals who gather information from a variety of inputs to solve problems and make decisions. An example of DSS is the loan processing system used by a bank to verify the credit of the applicant. In both VOD and DSS servers, bandwidth is of prime importance.

Table 9 provides the typical I/O workload profile for this application category.

Table 9. Typical Large Random Reads I/O Workload Profile

Application Category D	I/O Workload Profile		
	Read/Write Percentage	Random/Sequential Percentage	Size
Video On Demand (VOD)	100% Reads	100% Random	512KB
Decision Support Systems (DSS)	100% Reads	100% Random	1MB

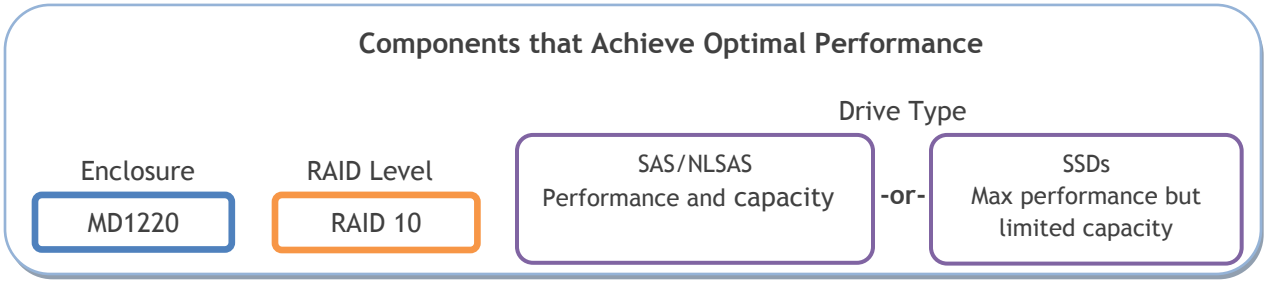

Figure 4. Simulation Results for Application Category D: Large Random Reads

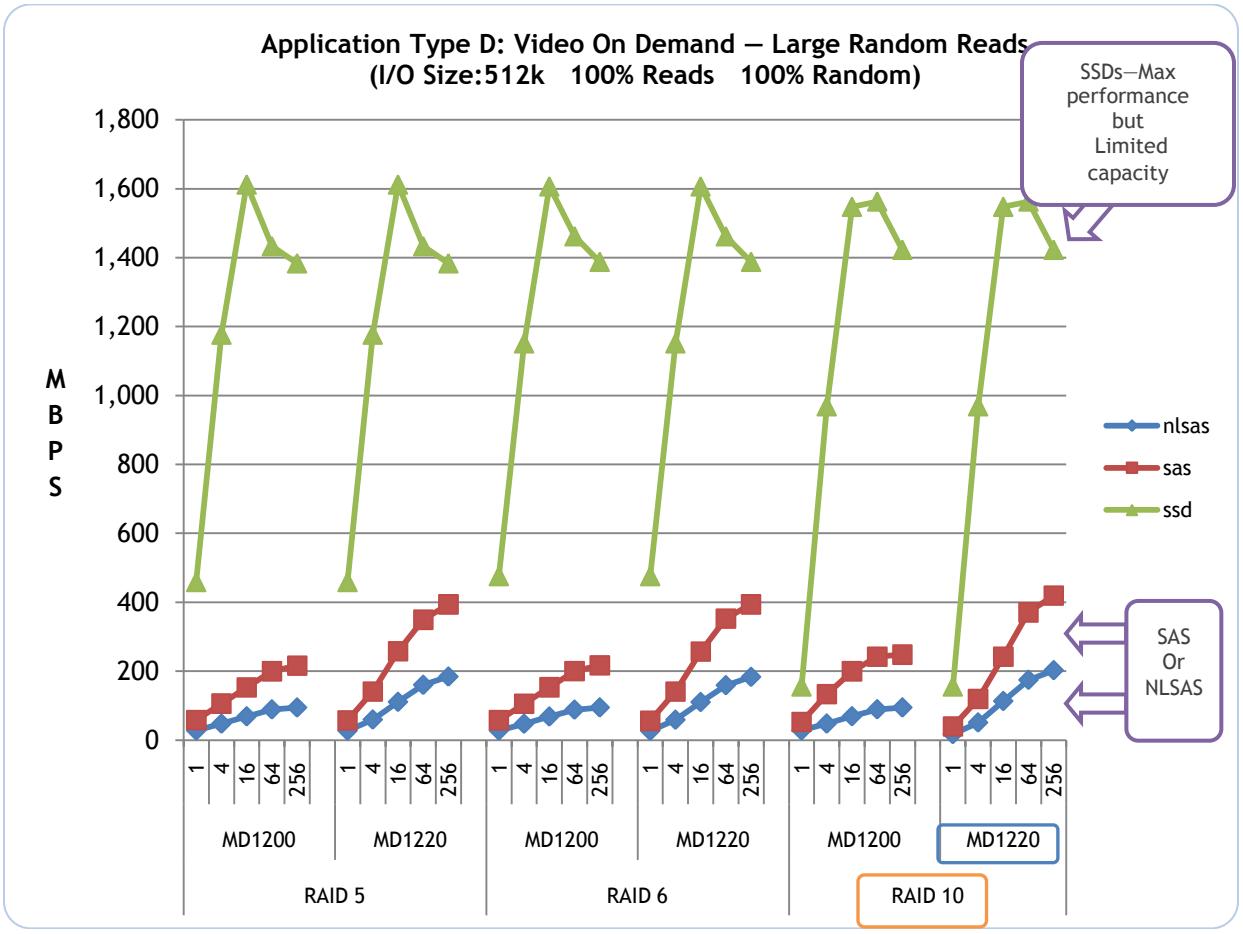

Table 10 lists the IOmeter simulation results of application category D graphed in Figure 4 for all components (RAID, enclosure, drive type) at maximum I/O load of 256 (queue depth).

Table 10. Simulation Results for Application Category D: Large Random Reads at Maximum Load of 256

RAID Type	Enclosure	Load (Queue Depth)	Max IOPS (512 KB Block Size)		
			NL-SAS (# Drives)	SAS (#Drives)	SSD (# Drives)
RAID 5	MD1200	256	94 (12)	216 (12)	1383 (8)
	MD1220	256	184 (24)	394 (24)	1383 (8)
RAID 6	MD1200	256	94 (12)	216 (12)	1388 (8)
	MD1220	256	183 (24)	393 (24)	1388 (8)
RAID 10	MD1200	256	94 (12)	248 (12)	1422 (8)
	MD1220	256	202 (24)	419 (24)	1422 (8)

Based on these results, the following are the guidelines and recommendations for selecting the enclosure, drive type, and RAID level for large random reads, application category D.

Application Category D (Large Random Reads) Recommended Storage Configuration		
Component	**Recommendation**	**Rationale**
Enclosure	MD1220	Since performance is the priority in VOD/DSS servers, spreading data across more spindles improves random performance.
Drive Type	HDDs (NL-SAS or SAS) Combination of Performance and Capacity **-OR-** SSDs* Maximum Performance but Limited Capacity	Video on demand and decision support system servers require a large amount of storage due to huge payload sizes (512KB and 1MB respectively). • NL-SAS/SAS HDDs provide a combination of storage capacity and two distinct levels of performance that may make them a better solution. • SSDs provide the best performance (up to 15 times better than some HDDs) with limited storage capacity.
RAID Level	RAID 10	Since workloads are 100% reads, RAID 5, 6, and 10 deliver similar performance. For more detail, refer to Appendix B: RAID Guide.
*Users should factor in the higher cost of SSDs when considering a SSD based solution.		

Application Category E: Large Sequential Reads

Applications in this category include streaming media servers, which are systems that provide web-casting, video conferencing, Internet entertainment such as Internet TV or radio, and multimedia services. These systems require a balance between storage capacity, availability, redundancy, and performance. Like web servers, they may also be part of a group of systems that work together to provide content. Table 11 provides the typical I/O workload profile for this application category.

Table 11. Typical Large Sequential Reads I/O Workload Profile

Application Category E	I/O Workload Profile		
	Read/Write Percentage	Random/Sequential Percentage	Size in KB
Streaming Media	98/2	100% Sequential	64KB

Components that Achieve Optimal Performance and Capacity

Enclosure	RAID Level	Drive Type
MD1220	RAID 6	NLSAS or SAS

Figure 5. Simulation Results for Application Category E: Large Sequential Reads

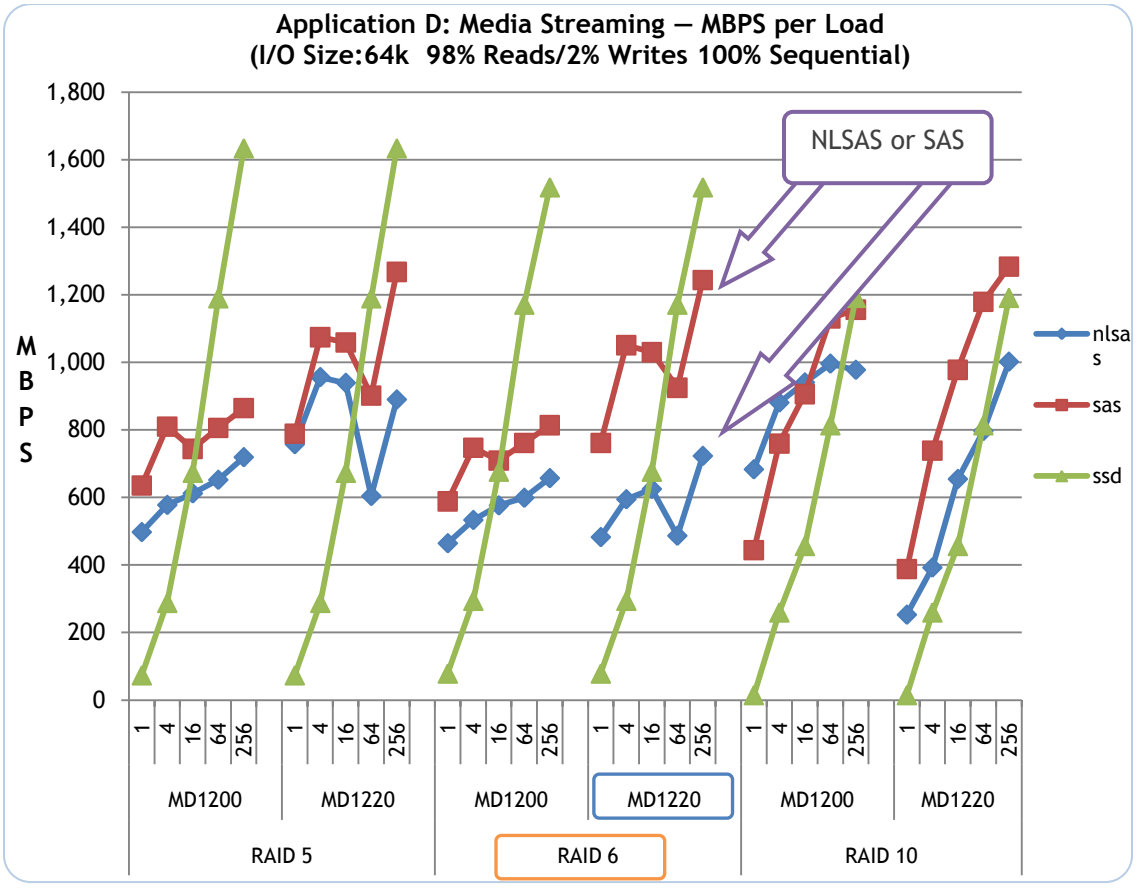

Table 12 lists the IOmeter simulation results of application category E graphed in Figure 5 for all components (RAID, enclosure, drive type) at maximum I/O load of 256 (queue depth).

Table 12. Simulation Results For Application Category E: Large Sequential Reads at Maximum Load of 256

RAID Type	Enclosure	Load (Queue Depth)	Max IOPS (64 KB Block Size)		
			NL-SAS (# Drives)	SAS (# Drives)	SSD (# Drives)
RAID 5	MD1200	256	718 (12)	864 (12)	1633 (8)
	MD1220	256	889 (24)	1267 24)	1633 (8)
RAID 6	MD1200	256	656 (12)	813 (12)	1518 (8)
	MD1220	256	722 (24)	1243 (24)	1518 (8)
RAID 10	MD1200	256	977 (12)	1155 (12)	1190 (8)
	MD1220	256	1001 (24)	1282 (24)	1190 (8)

Based on these results, the following are the guidelines and recommendations for selecting the enclosure, drive type, and RAID level for large sequential reads, application category E.

Application Category E (Large Sequential Reads) Recommended Storage Solution		
Component	**Recommendation**	**Rationale**
Enclosure	MD1220 Maximum Performance but Limited Capacity	Spreading data across more drives increases throughput, since more spindles are accessed concurrently. MD1220 offers up to 24 drives per enclosure.
	MD1200 Capacity And Performance	MD1200 offers fewer drives (12) but provides greater storage capacity with larger 3.5'' form factor drives capable of up to 2 TBs per drive.
Drive Type	HDDs (NL-SAS or SAS) Capacity and Performance	For streaming media applications, the focus is on storage capacity. SSDs may perform better than HDDs at larger queue depths, but cannot match the capacity of NL-SAS or SAS drives. Until the capacity of SSDs increase, HDDs are a better option.
RAID Level	RAID 6	RAID 6 is the recommended RAID level based on performance; however, RAID 5 also offers similar performance but with less redundancy protection. More details can be found in Appendix B: RAID Guide.

Summary

Performance metrics were collected from various PowerVault storage solutions and were configured using different RAID levels, drive and enclosure types, and PERC H800 adapter settings. The metrics show that each solution can achieve different performance levels depending on application type. Both traditional HDDs and recently introduced enterprise SSDs offer unique benefits depending upon the I/O load profile. Each drive type offers a different combination of storage performance and capacity. If the application type is 100% random I/O, combining the PERC H800 with either the MD1220 or MD1200 matched with SSDs will usher in a new era of performance that hard drives cannot match. Alternatively, if the application type is 100% sequential I/O, the 7.2K RPM NL-SAS and 15K RPM SAS drives deliver a powerful combination of performance and capacity (up to 2 TBs per 3.5'' drive w/MD1200 enclosure) that current generation SSDs cannot match.

Table 13 summarizes the MD1200/1220 direct-attached storage (DAS) solution recommendations for achieving maximum I/O performance for the application categories outlined in Table 2.

Table 13. Solution Recommendation Based on Application Category

Application		Storage Recommendations			
Category *Example*	**Key Focus**	**Enclosure**	**Drive Type**	**Controller Cache Settings**	**RAID Level**
A: Small Random Reads *Web Server*	Performance	MD1220	SSDs**	Write Through No Read Ahead	5, 6, 10*
B: Random Reads and Writes *DB OLTP and Exchange Mail*	Performance and Scalability	MD1220	SSDs**	Write Through No Read Ahead	5, 6,10*
C: Sequential Writes *SQL/Web Logs*	Capacity and Performance	MD1200 Capacity and Performance -**OR**- MD1220 Typical Capacity and Performance	HDDs NL-SAS or SAS	Write Back Adaptive Read Ahead	5, 6*,10*
D: Large Random Reads *Video On Demand and Decision*	Capacity and Performance	MD1200	HDDs Combination of Performance and Capacity	Write Back Adaptive Read Ahead	5, 6, 10*

Application		Storage Recommendations			
Category ***Example***	**Key Focus**	**Enclosure**	**Drive Type**	**Controller Cache Settings**	**RAID Level**
Support Systems			SSDs** Maximum Performance; Limited Capacity	No Read Ahead Write Through	
E: Large Sequential Reads *Streaming Media Servers*	Capacity and Performance	MD1220 Maximum Spindle and Performance -**OR**- MD1200 Maximum Capacity and Performance	HDDs NL-SAS or SAS	Write Back Adaptive Read Ahead	5, 6*,10
***** RAID level that achieved maximum performance. ******Users should factor in the higher cost of SSDs when considering a SSD based solution.					

Appendix A: The Hardware Used

PERC H800 SATA/SAS Adapter

The PERC H800 controller is the 3rd generation SATA/SAS controller designed for a mature 6Gbps end-to-end solution with increased flexibility, scalability, and I/O performance capabilities when attached to the PowerVault MD1200 and MD1200 expansion array enclosures.

Key Feature List

- PCI-Express Gen2.0 support
- 6Gb/s SAS (SAS 2.0) host interface
- Doubles throughput performance capability from previous generation of PERC
- Connectivity to a mixture of both PowerVault MD1200 and MD1220 storage enclosures

For a complete overview of the H800's features, see the User's Guide:
http://support.dell.com/support/edocs/storage/Storlink/H700H800/en/UG/PDF/H700H800.pdf

http://www.dell.com/content/topics/topic.aspx/global/products/pvaul/topics/en/us/raid_controller?c=us&cs=555&l=en&s=biz

PowerVault MD1200/MD1220 Expandable Array Enclosure

PowerVault MD1200 and MD1220 array enclosures are the 2nd generation of Dell's energy-efficient, small-form-factor (SFF) drive expansion enclosures. The MD1200 supports a combination of twelve (12) 3.5-inch or 2.5-inch drives while the MD1220 supports twenty-four (24) 2.5-inch drives in the same chassis. Both the enclosures support 3.0Gbps and 6.0Gbps drive types. The modular design of the PowerVault MD1200 and MD1220 enables scalability. Both enclosures can be daisy-chained together for a combination of up to eight enclosures on one PERC H800 adapter. Up to six PowerVault MD1220 24 mixed drive expansion enclosures can be daisy-chained together on one PERC Series 7 H800 controller, giving you up to 144 mixed disk drives options. To simplify management, a single PERC H800 RAID controller-equipped PowerEdge server can connect and control both PowerVault enclosures.

Key Feature List

- 6Gb/s SAS: 35% increase in IOPS and twice the throughput of last generation arrays
- Solid State Drives: Support for solid state drives (SSD) that have 3x the IOPs performance

For a complete overview of the features of the PowerVault MD1200 and MD1220, see the User's Guide:

http://support.dell.com/support/edocs/systems/md1200/multlang/GSG/DAO_BCC/K356M.pdf

http://www.dell.com/us/en/enterprise/storage/sas/cp.aspx?refid=sas&s=biz&cs=555

http://support.dell.com/support/edocs/systems/md1200/en/index.htm

Disk Drive Technologies Measured

The 7.2K NL-SAS and 15K SAS were the first two drive types measured. Both are traditional HDDs, built on the same internal design composed of rotating platters, arms, and mechanical components. SSDs are the third drive type measured. Internally, SSDs are composed of memory chips versus the electro-mechanical parts found in the first two drive types, and introduce a new era of I/O performance.

7.2K NL-SAS drive

The 7.2K NL-SAS drive is Dell's enterprise class SATA drive, engineered to withstand the business operation cycle while leveraging large storage capacities from the consumer grade version. Of the three drive types, the NL-SAS offers the best dollar per gigabyte ($/GB) ratio. It is a popular choice for server applications that require both a large capacity and good I/O performance.

15K SAS drive

The 15K SAS drive is the second HDD measured inside both PowerVault enclosures. Similar in design to the 7.2K NL-SAS disk drives, the internal platters of the 15K SAS disk drive rotate twice as fast. This translates to a 20-30% improvement in I/O performance over the 7.2K NL-SAS drive. This drive type is suited for mission critical applications requiring the highest reliability and the best I/O performance out of the traditional drive types offered.

Solid State Drives (SSDs)

The newest drive technology tested is the 2.5-inch SSD drive, Dell's first enterprise level flash drive. SSDs deliver the best I/O performance because I/O latencies are now comparable to memory access times. Results collected from applications that issue random I/Os show that SSDs are able to achieve the best I/O performance of all the drive types measured. Additionally, the number of SSDs required to achieve this level of performance is only a fraction of the number of traditional drives needed.

Appendix B: RAID Guide

Which RAID Level is Best for My Application Type?

Selecting the appropriate RAID level to implement across a storage sub-system can be challenging. Each of the PERC H800-supported RAID levels offers a different level of redundancy, available/useable capacity, and I/O performance.

Table 14 is a guide to selecting the best RAID level, based on the number of drives that can fail before data is lost, storage capacity availability after array creation, and application type.

Table 14. RAID GUIDE

RAID Level	Available Capacity (N = total number of disks)	Number of Drives That Can Fail	Sequential Reads (MB/s)	Sequential Writes (MB/s)	Random Reads (IOPS)	Random Writes (IOPS)
RAID 0*	N	0	Best	Best	Best	Best
RAID 5	N-1	1	Best	Better	Best	Good
RAID 6	N-2	2	Better	Better	Better	Good
RAID 10	N/2	Multiple	Good	Good	Better	Better
RAID 50	N-2	2	Best	Better	Best	Good
RAID 60	N-4	4	Better	Better	Better	Good
***While RAID 0 offers the best MB/sec and IOs/sec performance, it offers no fault tolerance.**						

RAID I/O Amplification: To help understand RAID performance, it is useful to know how many physical reads and/or writes must occur when a logical read or write is issued to the RAID array. As we know, RAID 0 offers no data integrity; one (1) logical read or write maps to one (1) physical read/write. Logical reads on RAID 5, RAID 6, RAID 10, RAID 50, and RAID 60 arrays all translate to one (1) physical read as no data is being written to the arrays. Once a single write is issued to the same RAID levels, their performance is impacted by the parity calculation. Table 15 summarizes RAID I/O amplification for RAID levels measured in this paper.

Table 15. RAID I/O Amplification

	RAID 0	RAID 5	RAID 6	RAID 10	
1 Logical Read	1	1	1	1	Physical Read I/O
	0	0	0	0	Physical Write I/O
1 Logical Write	0	2	3	0	Physical Read I/O
	1	2	3	2	Physical Write I/O

RAID Levels Measured and Analyzed

The PERC H800 offers hardware RAID support with the following RAID levels: RAID 0, 1, 5, 6, 10, 50, and 60. All performance metrics collected in these tests were with virtual drives implementing RAID levels 0, 5, 6, and 10.

RAID 0 — RAID 0 offers the highest throughput rates of all RAID levels but no fault tolerance, since data is only striped across the disk array. The group reliability decreases as more disks are added to the RAID 0 disk array. User should consider RAID 0 only when performance is critical and data loss is acceptable. The throughput of the array will be the aggregate transfer capacity of all the disks, limited only by the throughput of the PERC Adapter.

RAID 5 — RAID 5 offers fault tolerance by generating and using block level parity information. All drives participate and store parity information. If a disk should fail in the array, parity with the surviving disks' data blocks can reconstruct the lost data on the fly; I/O continues to the array seamlessly, though with some performance degradation. Once the user replaces the failed drive, array rebuilding can begin; when complete, the performance degradation disappears.

RAID 6 — RAID 6 provides protection against double disk failures as well as failures while a single disk is rebuilding, by using striping in combination with parity information; each stripe maintains two disk blocks with parity. RAID 6 is not as efficient as RAID 5 or 10 when used with a small number of drives, since a significant amount of storage capacity is lost. However, as arrays become bigger and contain more drives, the loss of storage capacity is not as great. In the case where there is only one array, it may make more sense to use RAID 6 rather than have a dedicated hot spare disk.

RAID 10 — RAID 10 helps ensure fault tolerance and availability with a striped and mirrored disk pair sub-system. Since every disk has a mirrored copy, half the drives in the array are used for mirroring. Therefore, RAID 10 needs a greater number of disks to provide redundancy than other RAID levels. Unlike RAID 0, RAID 10 volumes must write data twice due to mirroring, but these writes are simultaneous.

Appendix C: PERC H800 Cache Settings

PERC H800 Cache Settings

The PERC controller uses a small amount of high performance cache memory to hold data for both reads and writes. The controller offers both read and write cache modes that can help increase application I/O performance.

Adaptive Read Ahead vs. No Read Ahead

In a data read from a disk with Adaptive Read Ahead enabled, the controller transfers data to its cache from the logical drive equal to the stripe size. I/O performance is best when the workload is mainly sequential reads and the I/O request size is larger than the stripe size. In No Read Ahead mode, the controller reads the I/O request size *without* reading ahead to the end of the stripe. Random read workloads or I/O requests smaller than the stripe size benefit the most from this mode.

Write Through vs. Write Back

In Write Through mode, the data bypasses cache memory, and the controller must wait for completion of the data write to the physical disk before returning the complete command to the operating system. This is a slow process that affects I/O performance, since the server thread must wait for the entire write operation to take place synchronously. In Write Back mode, the RAID controller responds with the complete command as soon as the data is in the cache. The server can continue processing additional commands instantly while the controller completes the write asynchronously. When the workload is write intensive and mostly random, Write Back mode usually results in increased performance.

For all IOmeter measurements with rotational media, we used PERC H800 default cache modes of Write Back with Adaptive Read Ahead. This was to show the out-of-the-box performance results achieved with the PERC H800, the PowerVault enclosures, and the 7.2K NL-SAS and 15K SAS drive types. Due to the high performance results of SSDs, the controller's default cache modes were set to No Read Ahead with Write Through to bypass all caching logic. Results showed that with caching logic enabled on SSDs, overall I/O performance was reduced.

Appendix D: Hardware Description

Test Platform – PowerEdge R610	
Processor	(2) E5530 2.4Ghz, 8M Cache, Turbo, HT
System BIOS	1.0.4
Memory	16GB
OS	Microsoft Windows Server 2008 R2
IOmeter Version	2006.07.27

PowerEdge RAID Controller H800	
Raid on chip	LSI 2108
Data Transfer Rate	Up to 6Gb/s per port
Cache Memory Size	512MB
Caching Methods	Read: No/Always/Adaptive Write: Write Back/Write Through
Battery Backup	Yes
Firmware Version	12.3.0-0031
Driver	Dell 4.23.0.64

Hard Drives used in PowerVault MD1220			
	7.2K NL-SAS	15K SAS	SSD
Manufacturer	Seagate	Seagate	Pliant
Model	ST9500430SS	ST9146852SS	LB150
Drive Firmware	DS63	HT62	D011
Capacity	500GB	146GB	150GB

Hard Drives used in PowerVault MD1200			
	7.2K NL-SAS	15K SAS	SSD
Manufacturer	Seagate	Seagate	Pliant
Model	ST32000444SS	ST3600057SS	LB150
Drive Firmware	KS68	ES62	D011
Capacity	2TB	600GB	150GB