

Save

ULTAMUS™ RAID

Guide to Performance & Benchmarking



A Word about Benchmarks

Benchmarks can be valuable tools for understanding and comparing the performance characteristics of RAID products. There are, however, so many different tools available, so many variables and so many interpretations possible that it is important to understand what can be expected as the result of any benchmark test.

This guide will help the reader understand what is actually being measured by various benchmarking tools, how to interpret the results those tools produce, what is important and what is not and what to stay away from.

Caution: RAID performance benchmarks are notorious for being difficult to reproduce and even more difficult to compare. There are so many variables and tuning parameters involved and so many optimizations possible, that it is very difficult to predict exactly how any RAID system will perform under any given workload.

Among the terms that enter into discussions about RAID performance are:

- Interface type and speed
- Controller cache
- Type of disk drives (SAS, SATA, FC, SCSI, etc.)
- Drive writeback cache
- I/O queue depth
- Number of outstanding I/O requests
- Ratio of sequential to random I/O
- Server bus type and speed
- Host Bus Adapter card type and settings
- Host Bus Adapter cache and DMA coordination with OS if applicable
- RAID levels
- Command queuing
- Command reordering
- Number of simultaneous streams
- Shared device response to non-deterministic requests
- Raw I/O vs. filesystem I/O (and filesystem type)
- Etc.

The most accurate way to understand exactly how a RAID array will respond to any given environment is to test the product in the environment with the actual workload expected.

Generally, two performance criteria are measured during benchmark testing. The first is a measure of the system's speed in transferring data measured in megabytes per second and abbreviated as MB/s or MB/sec. The second is a measure of a systems ability to handle transactional data measured in input/output operations per second and abbreviated as "IOPS".

These two benchmarks stress the system in completely different ways and it is impossible to achieve maximum performance in both dimensions simultaneously. In general, MB/s will be the greatest when the host computer is requesting large blocks of data. Intuitively this makes sense since once a location is found on the disk, reading data from it or writing data to it can occur at very high speed.

The opposite is true for IOPS. Maximum IOPS are generally achieved with very small request sizes. This also makes sense since the number of transactions that can occur in a given time frame increase when the size of the transaction is reduced. It's as if you're standing in line at the bank. If each person in front of you only requests a small transaction (for example, withdraw \$20), the line will move quickly. However, if each person in front of you requests a large transaction (for example, open an account), the line will move slowly.

The performance charts in this document will show the request size of the data shown along the horizontal (X) axis of the graph. Request sizes begin at 512 bytes and increase up to one megabyte (1 MB). IOPS will be shown on the left hand vertical axis (Y) and MB/sec will be shown on the right hand vertical axis (Y').

Figure 1 is an example of a RAID benchmark characterization chart.

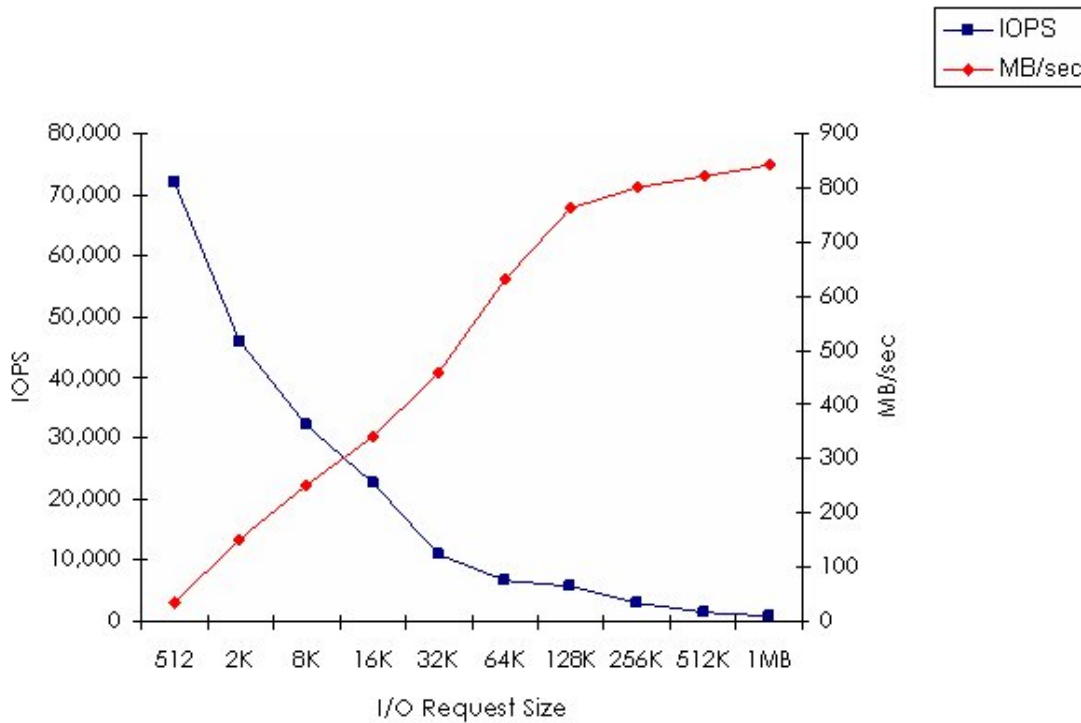


Figure 1. Sample RAID Characterization Graph

Figure 1 shows the performance characteristics of a RAID product that is capable of producing about 75,000 IOPS at a 512 byte request size and 750 to 850 MB/sec using 128K to 1 MB request sizes.

Whether this performance characterization is “good” or “bad” is entirely dependent upon the application.

Producing and publishing benchmark data allows the manufacturer, VAR partner and buyer to better understand the performance characteristics of the RAID product under investigation. This process is referred to as “RAID performance characterization.”

Characterization is important to RAID sellers and buyers alike, because both parties are interested in selecting the product that most closely matches the buyer's application needs. For example, you wouldn't buy a RAID product that is optimized for large block streaming video as the storage for a SQL database application. Nor would you buy a RAID product optimized for small block I/O for a streaming video server.

In most cases, however, buyers are looking for RAID systems with performance characteristics that match a number of workloads and applications. A good example is the storage manager looking for RAID storage for an Exchange server.

A typical Exchange server will have the following performance characteristics:

- Random 4K reads to the main database files
- Random 16K writes to the main database files
- Sequential 8K writes to the database log file
- Sequential 64K writes to the backup subsystem
- Sequential 64K reads from the main database files

Clearly, RAID storage optimized for either small or large block I/O would not be ideal for the Exchange server workload. What is needed is RAID storage that delivers balanced performance over the range of I/O sizes encountered.

A Special Note on Benchmarks Involving Filesystems

RAID sub-system characterization demands that testing focuses on the storage device. While this sounds obvious, benchmarks that involve filesystems typically produce results that obscure actual RAID performance in favor of overall system response performance.

Servers and storage are typically purchased with an application in mind. Over time, the application mix frequently changes. Server vendors tune and optimize the operating system and cache to perform well for the applications most typically encountered by their buyers.

Although this process optimizes the I/O for those few applications, it is likely that the system may not perform well for other applications that were not targeted for optimization. An example is database optimization. Many operating systems have filesystems and cache configured to work well with databases. While the database users are happy, other users may not be quite as happy as the entire system may be allocating all of the system resources to the database users at the expense of non-database users.

Over time the load may shift from random reader applications (database) to applications that demand optimized sequential read or write performance. When this happens, users may perceive that the machine is very slow when running this new application and become dissatisfied.

An example of extreme application optimization is the Overland Storage REO Series of disk-based backup, recovery and VTL appliances. REO Protection OS™

is tuned to maximize large block write performance. This is because the backup application is one that primarily writes data to disk-based virtual tape.

As REO has evolved, its performance characteristics have evolved as well. Now, in addition to very high write performance, REO also offers equally high read performance. Read performance is important to maximize the transfer of backup data from disk (virtual tape) to physical tape.

The changes recently implemented in REO Protection OS v3.0 improved read and write performance in selected cases by up to 6X – on the same hardware and with the same storage sub-system!

Clearly, the effects of operating systems and filesystems can have tremendous influence on overall system performance.

As a final note, consider the operations required to write a new file. When a new file is written, both the actual data and certain overhead information (referred to as “filesystem metadata”) are committed to permanent media. Metadata consists of the directory information, the space allocation and other data associated with a file that is not part of the data itself.

Benchmark tests that use filesystems should be avoided if the objective is to better understand the performance characteristics of a RAID sub-system.

Whether an Art or a Science – Why Bother?

Here's what Microsoft says about storage performance analysis:

“Analyzing storage subsystem performance is an art, not a science. Every rule has an exception; every system designer or system administrator has a different combination of hardware configurations and software workloads to accommodate. To fully understand the behavior of storage subsystems, one must understand every component along the path, from the user applications to the file systems to the storage driver stack to the adapters, controllers, buses, caches, and hard disks. Few people have the time and resources to accomplish this heroic task, so one is left with rules of thumb and general guidelines that inevitably need tweaking to maximize storage performance on any given system.”

Disk Subsystem Performance Analysis for Windows, Microsoft, March 2004

The reality of buying and selling RAID is that human nature tells us that “simpler is better”. And reducing a complex set of performance variables down to one or two numbers is what most often happens. The RAID buyer needs to know the right questions to ask and the RAID seller needs the right answers. This document will supply some of those questions and answers.

Interpreting Vendor Performance Claims

According to the Wikipedia:

“Your mileage may vary is a truism originating from the early 1970s when the United States Environmental Protection Agency (EPA) began to rate passenger automobiles on their fuel economy.

The first EPA fuel economy ratings were based on measurements of tailpipe emissions in laboratories, not on road testing, and consumers often observed significant differences between the EPA mileage figures and their experiences. Automakers whose products had favorable EPA numbers bragged of them in their television commercials, but an announcer would usually disclaim, "Your mileage may vary," because it is impossible for laboratory tests to simulate all driving conditions and driver behavior.

The phrase is frequently abbreviated in e-mails and Usenet postings as YMMV.”

When reviewing RAID vendor performance claims, be sure to remember that YMMV.

Actually, the EPA lab tests produce results that are more comparable than the claims of RAID manufacturers. The EPA tests all cars the same way. Run the car under test on a standard dynamometer at 48 MPH (76.8 km/h) with a peak speed of 60 MPH (96 km/h) and calculate the amount of fuel consumed in an hour. This can't be said about RAID performance testing. Each vendor and manufacturer sets the performance specifications that they believe best represents the capability of the product.

Some vendor specifications stop at the interface rate. For instance, a 2 Gb/sec FC interface has a theoretical maximum bandwidth of 206 MB/s. Round to 200 MB/s and that becomes the performance specification. OK, I guess, but it doesn't help the RAID buyer or seller determine whether the product is “right” for the application.

Regardless of how any other vendors rate the performance of their RAID products, Overland Storage has chosen to provide “transparent” performance information for the ULTAMUS RAID products.

Overland's published performance specifications are meant to help our partners and their customers select the product and configuration that best matches the performance requirements of the applications that will be using the storage.

Overland's specifications are repeatable. We publish our test methodologies and our results. We want the buyer to decide.

The Tools

Although far from exhaustive, the tools listed below are typically encountered in RAID benchmark testing:

| Tool name | Description and typical use |
|-------------------|--|
| "dd" | A Linux (and Unix) command that can be used to simulate I/O at various block sizes. |
| Bonnie v2.0.6 | An open systems benchmark tool that can be compiled under different UNIX flavors. |
| IOmeter | Originally developed by Intel, now in the public domain and available at www.sourceforge.net , IOmeter does for a computer's I/O subsystem what a dynamometer does for an automobile engine: it measures performance under a controlled load. Available for Windows 2000, Linux, and Solaris, IOmeter is perhaps the most widely used I/O performance measurement tool. |
| Jetstress | Used to predict the number of email users an Exchange server storage subsystem can support with "acceptable" response times |
| Postmark | A new benchmark tool to measure performance of e-mail, netnews, and e-commerce applications. Simulates heavy small-file system loads. Postmark can be compiled under Solaris, Digital Unix and Win32 environments. |
| SCSI Toolbox | Most often used by device manufacturing and integrators to run tests on many devices at once. Good reputation for repeatability of tests and detailed logging. |
| SiSoftware Sandra | A 32- and 64-bit Windows system analysis tool. |
| XDD | A comprehensive benchmarking tool developed by Tom Ruwart, a well known academic and consultant to DOD and DOE supercomputing centers. Available for Windows, Linux and most Unix flavors without charge at www.ioperformance.com |

Test Setups

Each of the tools described in the list above was used to benchmark the performance of the ULTAMUS RAID 1200. A standard server configuration was used in all test cases. The array(s) were connected to the server using 4 Gb/sec FC, directly attached and without an intervening switch.

The server specifications used to perform each test were as follows:

| Component | Type |
|------------------|---|
| Processor | Dual Intel Xeon 3.0 GHz |
| DDR-SDRAM | 4 x 1 GB DDR, 400 MHz, PC3200R-30330 |
| System drive | 36.4 GB, 15K RPM, Ultra 320 SCSI |
| FC Cards | Emulex 11002 4 Gb dual-port |
| Operating System | Windows 2003 Enterprise Edition, Service Pack 1 |

Internal Monitoring, Statistics Collection & Performance Reporting

ULTAMUS RAID includes an embedded performance monitoring tool that scans commands and data transfer operations to produce real-time performance reports and statistics. Monitoring data and statistics are recorded for both reads

and writes and can be used by the storage manager to tune the operating system and the I/O sub-system for optimum performance.

Details on how to best use the information collected and presented by the ULTAMUS RAID Manager are available in the Overland publication titled: **“ULTAMUS RAID Manager Users Guide”**.

The statistics monitored include:

- Data Transfer Performance
- Command Count
- Command Alignment
- Command Size
- Read-Ahead Statistics
- Write Clustering Statistics
- Write Statistics

Each RAID controller maintains individual access information for all logical drives, controllers, and ports. This information may also be exported to a comma delimited file (CSV file) for use in third-party software products. Figure 2 below illustrates the basic statistical data presentation screen.

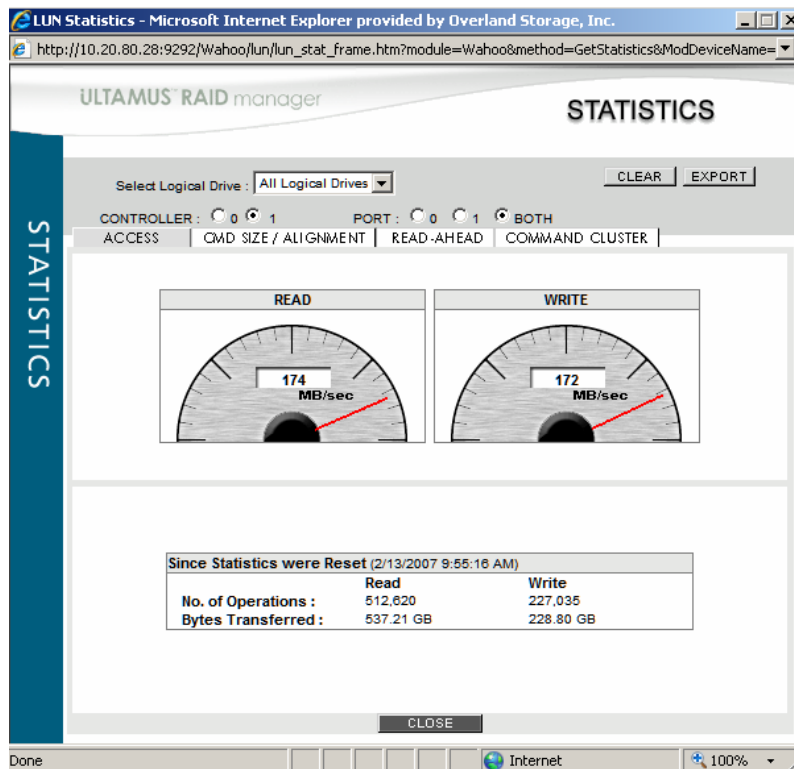


Figure 2. ULTAMUS RAID “Built In” Performance Monitor

Field Description

- Reads: The average number of megabytes of data transferred in the last few seconds from the logical drives, controllers or ports. This value is expressed in MB/sec.
- Writes: The average number of megabytes transferred in the last few seconds to the logical drives, controllers or ports. This value is expressed in MB/sec.
- Number of Operations: The total number of read and write accesses that have occurred since these statistics were reset, or the controller was last powered on.
- Bytes Transferred: The total number of bytes read and written since these statistics were reset, or the controller was last powered on.

Command Size and Alignment Statistics

Command size statistics express the percentage of commands whose size is as specified. The alignment statistic is the percentage of commands whose addresses aligned on the specified address boundary. For optimal performance, a write of one chunk of data should reside exactly within a chunk on one disk.

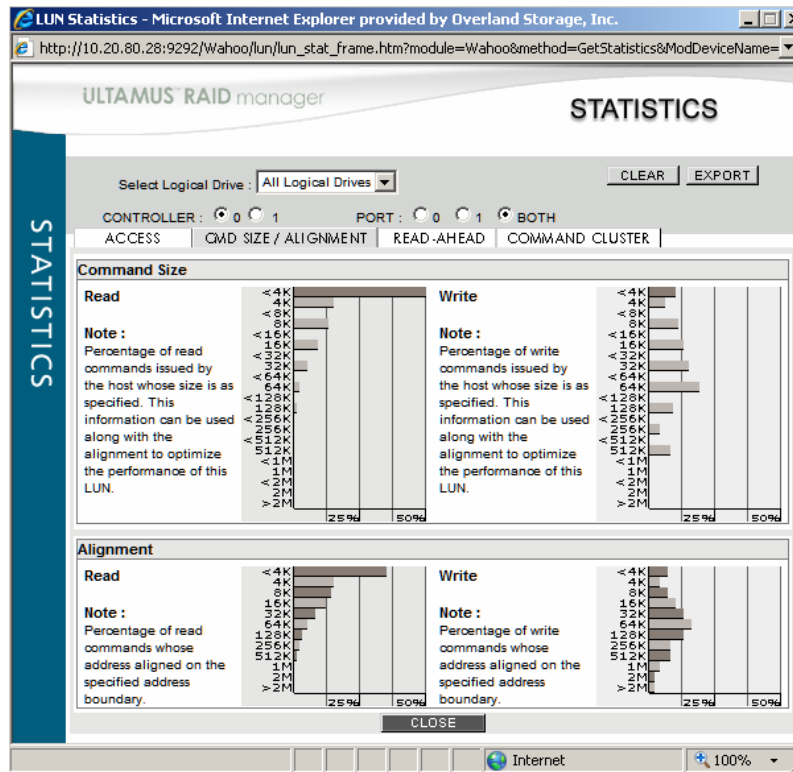


Figure 3. ULTAMUS RAID Command Size and Alignment Tab

Read-Ahead Statistics

When sequential read commands are sent to the RAID controller, it assumes that the commands which follow are also sequential. The RAID controller then reads ahead before the host requests the data. This improves performance, particularly for smaller reads. The size of the read-ahead is calculated based on the original command size, so the controller does not read too much data.

The controller maintains statistics for all read-ahead commands performed.

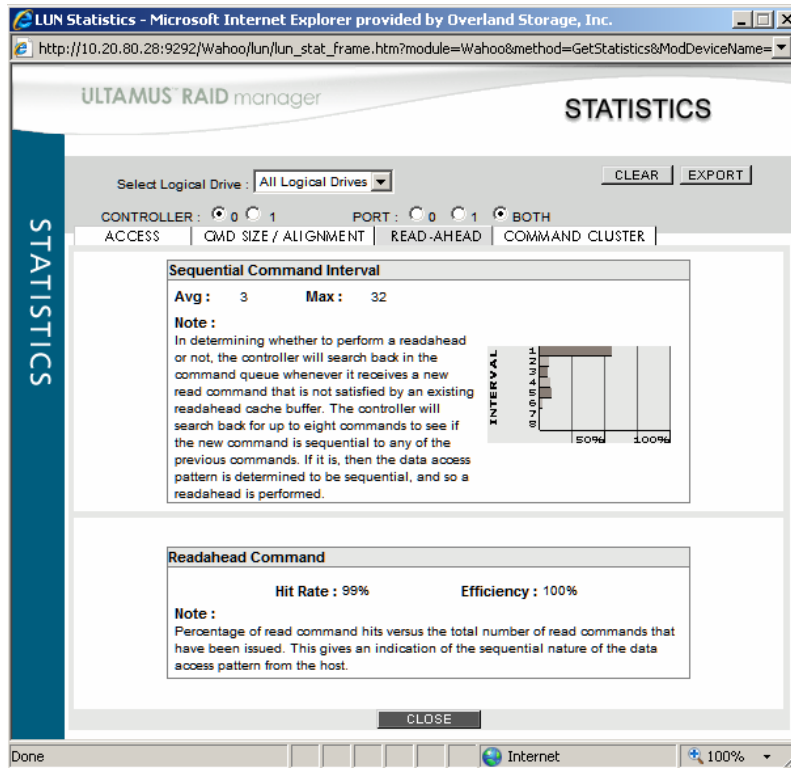


Figure 4. ULTAMUS RAID Read Ahead Tab

Command Cluster Statistics

To increase performance, the RAID controller aggregates (or “clusters”) sequential write commands together to create larger write commands. This results in fewer commands being sent to the disk drives. Additionally, if sufficient data is aggregated by the controller, it can then perform a full stripe write which can significantly improve performance. In cases where the host does not issue a sufficient number of outstanding write commands, RAID controller writeback cache can be used to delay the write to disk, increasing the likelihood of aggregating more data.

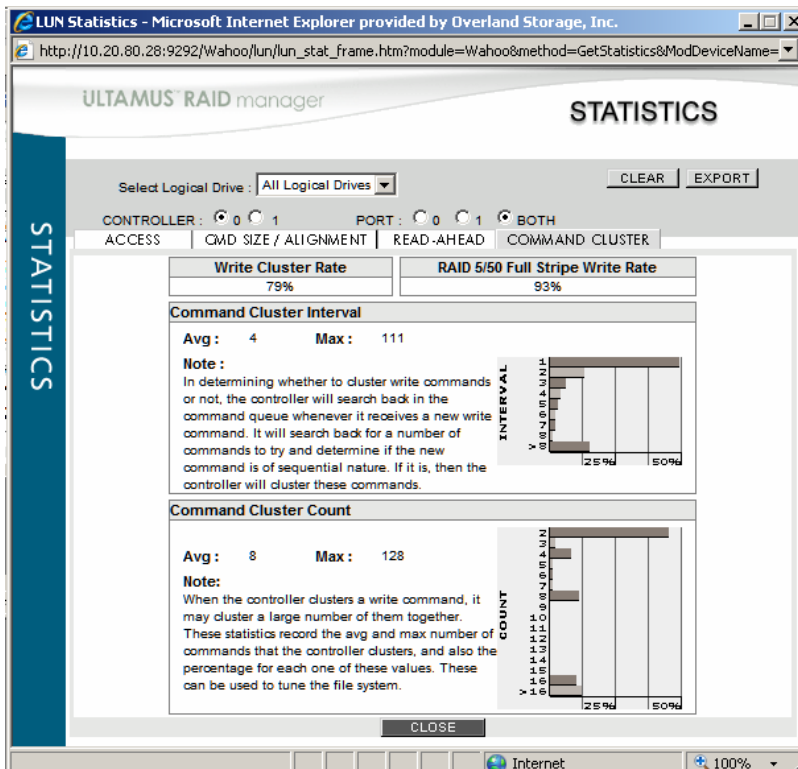


Figure 5. Command Cluster Tab

Performance Measurement Recommendations

Notes

ULTAMUS RAID allows end users to specify the capacity settings for RAID stripe depth, write-back cache and read-ahead cache through the ULTAMUS RAID manager. Our benchmarks have shown that general best practices for these settings are to use 256K stripe depth per RAID set, 256MB write-back cache and the “automatic” setting for read-ahead cache. These values were used to achieve the performance documented in the “Results” section of this document.

Note on SATA drives: SATA drives use "write back cache" to speed up write performance. In the event of a power failure (for a storage system not running on a UPS) the write cache memory on the disk drive itself is not committed to permanent storage (the disk drive itself). Filesystems rely on knowing when a write has completed to the disk. If writes are not properly committed to permanent storage, filesystem and data corruption occur. It is a best practice to power SATA RAID using a UPS or disable write back cache to ensure data integrity. Disabling write back cache reduces performance but preserves data integrity.

General Settings for all tests

Number of sectors - choose a value for the number of sectors that is at least 4 times the write cache size. This will prevent testing from running totally out of cache which can give artificially high results. Using a value equal to the number of sectors in the LUN will give true performance since it will include all the latencies of the system including disk seeks.

Application-Specific Settings

When studying SAN RAID performance, be sure to include tests that employ multiple servers accessing RAID storage simultaneously to observe the effects of "non-deterministic" access patterns. Single servers with a single file system produce the most deterministic access patterns but do not reflect the reality of SAN RAID access.

File and Print Servers and Email Servers (Exchange)

File and Exchange server user requests tend to be scattered across the disks in the RAID sets. "Hot spots" are frequently present. Hot spots, which typically contain file information structures, open-file updates, page swap space and other frequently accessed data, benefit from effective array caching.

RAID storage for file and Exchange serving must have the ability to handle a great number of user requests to achieve the lowest response times.

Thread count - separate test should be run with 2 and 4 threads. Two host systems should be used with 2 threads each to a clustered device.

Access Method - Random

Write Read Ratio - the write/read ratio should be set at 20% write, 80% read to emulate unstructured reference data and email traffic.

Block Size - separate tests should be run for 4K, 8K, 16K, 32K and 64K. All current file systems running today write some minimum number of sectors (i.e. cluster for windows, Inode for UNIX and Linux). To support high capacity, disk drives have increased the size of the sector cluster to 32 or 64 sectors.

Exchange servers will typically adhere to the following access characteristics:

- Random 4K reads to the main database files
- Random 16K writes to the main database files
- Sequential 8K writes to the database log file
- Sequential 64K writes to the backup subsystem
- Sequential 64K reads from the main database files

“Jetstress”, available from Microsoft, is a tool used to predict the number of email users an Exchange server storage subsystem can support with “acceptable” response times.

Jetstress can be downloaded here:

<http://www.microsoft.com/downloads/details.aspx?FamilyID=94B9810B-670E-433A-B5EF-47054595E9C&displaylang=en>

HD Video Content

Thread count - Single thread

Access Method - Sequential

Write Read Ratio - the write / read ratio should be set at 25% write, 75% read

Block Size - separate test should be run using the most efficient block sizes possible, typically 512KB to 1 MB.

| Video type | Format | Minimum RAID Throughput Required |
|------------------------------|-----------------|----------------------------------|
| Compressed High Definition | DVCPRO HD | 5.49 MB/sec |
| Uncompressed High Definition | 720p 24 fps | 42.19 MB/sec |
| | 720p 30 fps | 52.73 MB/sec |
| | 720p 60 fps | 105.47 MB/sec |
| | 1080 24p 8-bit | 94.92 MB/sec |
| | 1080 24p 10-bit | 126.56 MB/sec |
| | 1080i 8-bit | 118.65 MB/sec |
| | 1080i 10-bit | 158.2 MB/sec |

Large File Content

Thread count - separate test should be run with 2, and 4 threads. Two host systems should be used with 2 threads each to a clustered device.

Access Method - Sequential

Write Read Ratio - the write/read ratio should be set at 10% write, 90% read to emulate large content reference data.

Block Size - separate test should be run for 256K, 512K, 1M, and 2M to simulate CAD/CAM and other engineering drawing environments.

Non-HD Video Content

Thread count - Single thread

Access Method - Sequential

Write Read Ratio - the write / read ratio should be set at 25% write, 75% read

Block Size - separate test should be run using the most efficient block sizes possible, typically 256KB to 1 MB.

| Video type | Format | Minimum RAID Throughput Required |
|---------------------|------------------------|----------------------------------|
| Standard Definition | MiniDV | 3.43 MB/sec |
| | DVCAM | 3.43 MB/sec |
| | DVCPRO | 3.43 MB/sec |
| | DVCPRO 50 | 6.87 MB/sec |
| | Uncompressed SD 8-bit | 20.02 MB/sec |
| | Uncompressed SD 10-bit | 26.7 MB/sec |

OLTP Servers

RAID systems storing online transaction data, including database applications, web-based e-commerce, decision support and data warehousing, must efficiently respond to transaction processing and database query I/O requests. These requests tend to be small, generally less than 16KB, requiring a read of the original value, followed by an update, write, and verify operation.

OLTP servers are highly sensitive to I/O service and response times. The ability to preserve data integrity by maintaining in-order execution of requests in a mixed read/write workload environment with low latency is necessary to achieve high OLTP performance.

Thread count - separate tests should be ran with 2, 4, 8, and 16 threads. Two Systems should be used, each with 1- 8 threads to a clustered device.

Access Method - Random

Write Read Ratio - the write / read ratio should be set at 20% write, 80% read to better emulate write intensive operations of a data base.

Block Size - separate tests should be run for 4K, 8K, 16K, 32K and 64K, Oracle and other general purpose DBMS typically run in the 4K/8K range while SAP runs in the 64K range.

Web Servers

The typical web server workload is similar to that of the OLTP server. Web servers service multiple simultaneous requests with the lowest possible latency. Web server workloads are almost exclusively read requests with random I/O patterns. Approximately 84% of web server request sizes fall below 16KB, while the remaining 16% is evenly distributed from 32KB to 1MB (video, audio, images).

Optimal web server performance is achieved with RAID storage that distributes I/O requests across all the drives in a RAID set rather than a single DAS device. RAID in the web serving application significantly increases performance while improving availability and reliability.

Thread count - separate tests should be ran with 2, 4, 8 and 16 threads. Two Systems should be used, each with 1 - 8 threads to a clustered device.

Access Method - Random

Write Read Ratio - the write / read ratio should be set at 20% write, 80% read to better emulate write intensive operations of a database.

Block Size - separate tests should be run for 4K, 8K, 16K, 32K and 64K, Oracle and other general purpose DBMS typically run in the 4K/8K range while SAP runs in the 64K range.

Workstations

The workstation workload represents single users running multiple applications. Workstation workloads are predominantly a mix of random accesses to blocks with a sequential component of subsequent reads and writes. Since users are waiting for disk operations to complete before being able to continue, low latency is very important.

Thread count - separate test should be run with from 1 to 4 threads. A single host system should be used with 2 threads each to a clustered device.

Access Method – 80% sequential, 20% random

Write Read Ratio - the write/read ratio should be set at 20 % write, 80% read .

Block Size - separate tests should be run from 2K to 16K.

The table that follows summarizes Overland's guidance for setting up performance tests based on the access characteristics of selected business applications.

| Workload | Thread count | Access method | Block size | I/O characteristics | Typical applications |
|------------------------------------|---|----------------------------|-----------------------|--|--|
| File and print; Exchange server | 2 and 4 threads. Two systems with 2 threads each to the RAID device | Random | 4K – 64K | 20% write, 80% read; 70% of requests are for 4K and 64K block sizes | File and print servers, email (Exchange) servers; decision support systems |
| HD video | 1 | Sequential | 512K – 1M | 25% write, 75% read | HDTV video |
| Large file content | 2, 4 threads | Sequential | 256K, 512K, 1M, 2M | 10% write, 90% read | Backup to disk; CAD/CAM; engineering drawing |
| Non-HD video | 1 | Sequential | 256K - 1M | 25% write, 75% read | Digital video editing; telecine; digital television |
| OLTP server | 2, 4, 8, 16 threads. Two systems, 1 – 8 threads to the RAID device | Random | 4K, 8K, 16K, 32K, 64K | 20% write, 80% read; Oracle and other DBMS run in 4K/8K range; SAP runs in 64K range | SAP, Oracle, SQL, MySQL |
| Web server | 2, 4, 8, 16 threads. Two systems, 1 – 8 threads to the RAID device | Random | 512B – 512KB | 5% write, 95% read; large number of concurrent requests during peak periods | Apache; IIS; Sun Java System Web Server; Zeus |
| Workstation | 1, 2, 3, 4 threads | 80% sequential; 20% random | 2K – 16K | 20% write, 80% read | General business; scientific; engineering |

SATA II Drive Performance Test Results

IOmeter

Version 2004.07.03

32 Outstanding I/Os

2 workers

Raw filesystem

Originally developed by Intel and now available at no charge through SourceForge, IOmeter is one of the most popular and well respected I/O benchmarking tools available. Because of its popularity, a more thorough discussion of this benchmarking tool is warranted.

IOmeter's test agent is called a "Worker". Workers emulate an application sequentially reading/writing blocks of data starting from a certain sector. By increasing the number of requests coming from one Worker (from 1 to 8, stepping by

1) the observer can examine the ability of the RAID controller to reorder I/O requests (to group several requests for sequential data into a single request). By increasing the number of the Workers, IOmeter puts a heavier load on the disk subsystem – as in real world environments where several applications are competing to access the drives. Each Worker works with its own data (i.e. the data addresses requested by the Workers are different).

Instructions for running IOmeter:

1. Navigate to <http://sourceforge.net/projects/iometer>. Click on 'Download IOmeter', then 'iometer-stable'.
2. Click on 'iometer-2004.07.30.win32.i386-setup.exe' to download the version for 32-bit Windows. Download and run this file to install IOmeter on your system.
3. Go to Start / All Programs / IOmeter 2004.07.30 / IOmeter to start IOmeter.
4. Click the 'Disk Targets' tab. On the left part of the IOmeter screen, you will see a small picture of a computer with two workers attached to it. Click on that small computer. In the middle of the screen you will see logical drive letters with the disk volumes in your system. Check the box next to the logical drive letter corresponding to the 3ware array that you want to measure performance on.
5. Click the 'Access Specifications' tab. By default nothing is in the 'Assigned Access Specification' area. In the 'Global Access Specification' area, highlight '64K: 100 Read, 0% random'. This will measure sequential read performance on the array with a 64K block size. Click the 'Add' button to add this Access Specification. It will now show up in the 'Assigned Access Specification' area.
6. Click on the 'Results Display' tab. Move the 'Update Frequency (seconds)' slider bar to '1'.

NOTE: One time instructions - You know that the disk is unprepared because the small yellow disk drive image has a red line through it. After performing steps 3-5 above, click the green flag at the top of the screen. Click 'Save' to save the results file. It will say 'Preparing Drives' in the bottom of the screen. Wait about a minute until IOmeter creates a small test file called iobw.tst on the array. Click the Stop button, then the X to exit IOmeter. When you restart IOmeter, you will see that the small icon of the disk drive no longer shows the read line through it. You can now repeat steps 3-5.

7. Click the green flag at the top of the screen. Click 'Save' to save the results file. Since the drive has already been prepared, IOmeter will run and measure performance. The 'Total MBs per Second' is the Megabytes per second that are being transferred to the array.
8. By default IOmeter uses 1 outstanding I/O, simulating a single user load. Click on the 'Disk Targets' tab and look in the '# of Outstanding I/Os' box to see that value that is currently being used. To test with a different value, click the Stop button to stop the current test. Put a different value in this box (any integer between 1 and 256). Good numbers to test are 1, 16, 64, and 256.

SAS vs. SATA II Drive Performance Test Results

SAS drives use 10,000 - 15,000 RPM rotational speeds, which are well-suited for handling random database type transactions such as email, databases, data warehousing and OLTP (Online Transaction Processing).

SAS enables a direct, unshared 3.0 Gb/sec link from each drive to the RAID controller as opposed to SCSI that delivers a maximum bandwidth of 320 MB/sec that must be shared across every drive on the same bus.

There are currently two SAS disk drive form factors: the 3.5-inch form factor and the emerging 2.5-inch form factor. 3.5-inch form factor drives offer significant advantages over the emerging 2.5-inch disk drives, including nearly 30% higher performance (a total of five 2.5-inch drives is needed to match the performance of four 3.5-inch 15,000 RPM disk drives).

A 3.5-inch SAS disk drive can store as much as 300GB. A 2.5-inch drive can store up to 73GB.

SATA II disk drives are well-suited for less I/O performance-intensive applications that require large storage capacities at the lowest cost per GB. Current SATA II disk capacities range up to 750 GB each with 1 TB drives nearing production (approximately June 2007).

Disk drive performance metrics can be difficult to translate directly into RAID system performance characteristics. In the chart shown below, note the differences in latency and seek time between the SAS and SATA II drives. It is these key performance metrics that come into play to deliver the significant random I/O performance differential between SAS and SATA disk drives in RAID configurations

| Performance metric | 750 GB SATA II | 300 GB SAS |
|---|----------------|-------------|
| Maximum internal transfer rate (MB/sec) | 1,030 | 685 – 1,142 |
| Maximum sustained (MB/sec) | 78 | 58 – 96 |
| Cache (MB) | 16 | 8 |
| Average latency (msec) | 4.16 | 2.0 |
| Spindle speed (RPM) | 7,200 | 15,000 |
| Average read/write seek (msec) | 8.5/9.5 | 3.5/4.0 |
| Track to track read/write (msec) | 0.8/1.0 | 0.2/0.4 |

In October 2006, Network Computing Magazine published a SAS vs. SATA RAID performance comparison¹. This analysis documented the differences in performance that were observed when using different block sizes, I/O modes (read/write) & disk drive technologies. The results of those tests are shown below.

¹ <http://www.networkcomputing.com/gallery/2006/1012/1012rolloutb.jhtml>

| Drive type | Sequential I/O | IOPS | MB/sec | Performance Difference* |
|------------|----------------|--------|--------|-------------------------|
| SAS | 64K read | 4,964 | 310 | 38% |
| SATA | | 3,597 | 225 | |
| SAS | 64K write | 2,739 | 171 | 15% |
| SATA | | 2,391 | 149 | |
| SAS | 1MB read | 330 | 330 | 46% |
| SATA | | 226 | 226 | |
| SAS | 1MB write | 158 | 158 | 1% |
| SATA | | 156 | 156 | |
| SAS | 512B reads | 51,559 | 25 | 7% |
| SATA | | 48,340 | 24 | |
| SAS | NWC custom | 4,402 | 114 | 115% |
| SATA | | 2,049 | 53 | |

*SAS increase in performance over SATA – [vendor product]

Here is what Network Computing Magazine had to say about the test:

“We tested the [vendor product] in our Green Bay, Wis., Real-World Labs®. For comparison, we also ran tests on [vendor product] SATA-based 10.5-TB [vendor product]. We used Iometer and four Opteron-based servers with multiple [I/O] ports on both arrays. In all cases, the SAS unit outperformed its SATA competition. Overall, the SAS system averaged 48 percent faster in reads and 8 percent faster in writes than the high-capacity SATA enclosure. Our 512-byte IOPS test showed only a 7 percent difference between the two systems. The SAS [vendor product] also aced our NWC Custom test—a torture test that hammers systems with a mix of small random reads and writes. The SAS array moved more than twice as much data as its SATA counterpart.”

The test was repeated in the Overland Storage Applications Test Lab using two servers and a similar storage configuration. The results were as follows:

| Drive type | Sequential I/O | IOPS | MB/sec | Performance Difference* |
|------------|----------------------|--------|--------|-------------------------|
| SAS | 64K sequential read | 5,059 | 316 | -19% |
| SATA | | 6,252 | 391 | |
| SAS | 64K sequential write | 4,697 | 294 | 72% |
| SATA | | 2,725 | 170 | |
| SAS | 1MB sequential read | 587 | 587 | 35% |
| SATA | | 434 | 434 | |
| SAS | 1MB sequential write | 315 | 315 | 82% |
| SATA | | 173 | 173 | |
| SAS | 512B sequential read | 63,951 | 31 | -3% |
| SATA | | 65,820 | 32 | |

*SAS difference in performance compared to SATA – Overland Storage ULTAMUS RAID 1200

Overland's testing suggests that the difference in sequential write performance between SATA II and SAS drives is observable in I/O's using 64K and 1MB request sizes. Read performance differences are somewhat less pronounced.

Random I/O Performance Comparison - SAS/SATA II

Random I/O performance demonstrates a far more significant and consistent performance differential between SAS and SATA drives than observed in sequential operations, confirming the IOPS performance superiority of SAS drives over SATA drives for database, Exchange server and other random access applications.

| Drive type | Random I/O | IOPS | MB/sec | Performance Difference* |
|--|------------------------|-------|--------|-------------------------|
| SAS | 512B random read/write | 3,544 | 1.7 | 244% |
| SATA | | 1,030 | 0.5 | |
| SAS | 4K random read/write | 3,442 | 13.4 | 242% |
| SATA | | 1,006 | 3.9 | |
| SAS | 16K random read/write | 3,144 | 49.1 | 235% |
| SATA | | 939 | 14.7 | |
| SAS | 64K random read/write | 2,346 | 146.6 | 214% |
| SATA | | 747 | 46.7 | |
| SAS | 1MB random read/write | 396 | 395.5 | 164% |
| SATA | | 183 | 183 | |
| *SAS difference in performance compared to SATA – Overland Storage ULTAMUS RAID 1200 | | | | |

Overland's testing suggests that the performance difference in a random access, mixed read/write environment between SAS and SATA II disk drives is significant. Across the entire range of tested I/O – from 512B to 1MB request sizes - SAS drives significantly outperformed SATA II drives in random read/write environments.

The conclusion is that in applications in which random I/O access patterns predominate (like Oracle, MySQL, SQL databases and Exchange email systems), the use of SAS drives rather than SATA II drives can significantly improve application performance and reduce response time.

SAS vs SATA II Summary

SAS and SATA II drives have arrived as the state-of-the-art in disk drive interface technology. SAS joins Fibre Channel as a premier high performance storage interface technology, essentially obsoleting SCSI drives. SATA II emerges as the new price/performance standard for capacity storage. The table shown below describes the differences among the popular storage interfaces.

| | ATA/IDE | SATA II | SCSI | SAS | FC |
|-------------------------------------|--------------|----------------|---------------|--------------------------------------|--------------------------------------|
| Number of devices | 2 | 1 | 16 | 16,384 | 16 million |
| Maximum distance | 0.5 meters | 1 meter | 3 – 25 meters | 10 meters | 10 km |
| Cable type | Copper | Copper | Copper | Copper | Copper or fiber optic |
| Interface type | Parallel | Serial | Parallel | Serial | Serial |
| Form factor | 3.5 in | 3.5 in | 3.5 in | 2.5/3.5 in | 3.5 in |
| Transfer speeds (MB/second) | Up to 133 | 300 | Up to 320 | 300/600 | 100, 200, 400 |
| I/O rates (IOPS) | Up to ~50 | Up to ~130 | Up to ~140 | Up to ~210 | Up to ~210 |
| Drive rotational speed | Up to 7,200 | Up to 10,000 | 10,000/15,000 | 10,000/15,000 | 10,000/15,000 |
| Fault isolation architecture | N/A | Point-to-point | Shared bus | Point-to-point | Looped |
| Device compatibility | ATA/IDE only | SATA only | SCSI only | SAS and SATA | FC only |
| High reliability features | N/A | Port selector | N/A | Dual active porting; multi-initiator | Dual active porting; multi-initiator |

The table shown below describes the suitability of each popular disk drive technology for typical applications.

| | High-End Applications (OLTP, HA, random write and retrieval) | Mid-Range Applications (OLTP, HA, random write and retrieval) | Low-End Applications (Nearline, disk-based backup, sequential reads and writes) | Available Drive Capacities | Typical Applications and Characteristics |
|-----------------------------------|---|---|---|-----------------------------------|--|
| Fibre Channel (FC) | YES (Suitable for mission-critical data in large enterprises, e.g. SAN) | NO (Relatively costly unless FC infrastructure already in place) | NO (Too costly) | 18 – 300 GB | Transaction data (databases) <ul style="list-style-type: none"> • Maximum IOPS for OLTP • Calculation intensive files • Small block random reads/writes • Reservation systems • Billing systems |
| Serial Attached SCSI (SAS) | YES (Suitable for mission-critical data in large enterprises, e.g. SAN) | YES (Suitable for mission-critical data in SMB and mid-range enterprises) | NO (Not as cost-effective for high capacity applications as SATA) | 36 - 300 GB | Transaction data (databases) <ul style="list-style-type: none"> • Maximum IOPS for OLTP • Calculation intensive files • Small block random reads/writes Throughput data <ul style="list-style-type: none"> • High MB/sec, large data intensive files |
| Serial ATA (SATA II) | NO (Not suitable for OLTP/database applications) | YES (Suitable for mission-critical data in SMB and mid-range enterprises with RAID 6) | YES (Highly cost-effective for high capacity applications) | 80 – 750 GB (1TB soon) | Reference data <ul style="list-style-type: none"> • Fixed content • Archival data for secondary/nearline applications • Image storage • Backup to disk |

RAID 5 vs. RAID 6 Performance Test Results

RAID 6 is a close relative of RAID 5. However, RAID 6 provides the ability to recover from 2 drive failures in the same RAID set. The extra parity block is different from the first in the way it's generated and calculated. Generating parity blocks can slow down data transfer speeds, and because RAID 6 performs this process twice, it becomes that much slower.

ULTAMUS RAID offers hardware accelerated RAID processing. The ULTAMUS RAID storage processor ASIC accelerates parity calculations in silicon for all RAID levels, including RAID 6. Because of dual drive failure protection, RAID 6 requires twice as many parity calculations as RAID 5 when writing data. This level of parity calculation can stress competing products, many of which perform RAID 6 parity calculations using a general-purpose processor and software. The ULTAMUS RAID storage processor delivers the performance needed to support RAID 6 with a "worst case" read performance penalty of 12% when compared to RAID 5 and a "worst case" write performance penalty of 22% while delivering dual drive failure protection.

A comparison of RAID 5 vs. RAID 6 read performance is tabulated below. This test was conducted using 750 GB disk drives.

| Read Access Specification | RAID 6 MB/sec | RAID 5 MB/sec | RAID 6 Performance Penalty |
|----------------------------------|----------------------|----------------------|-----------------------------------|
| 512B Sequential Read | 32.14 | 30.84 | N/A |
| 4K Sequential Read | 166.07 | 168.45 | 0.6% |
| 8K Sequential Read | 228.08 | 234.88 | 2.9% |
| 16K Sequential Read | 299.11 | 320.83 | 6.8% |
| 64K Sequential Read | 386.00 | 438.85 | 12.0% |
| 128K Sequential Read | 410.69 | 451.67 | 9.1% |
| 1M Sequential Read | 434.77 | 463.20 | 6.1% |

A comparison of RAID 5 vs. RAID 6 write performance is tabulated below. This test was conducted using 750 GB disk drives.

| Write Access Specification | RAID 6 MB/sec | RAID 5 MB/sec | RAID 6 Performance Penalty |
|-----------------------------------|----------------------|----------------------|-----------------------------------|
| 512B Sequential Write | 4.92 | 4.86 | N/A |
| 4K Sequential Write | 38.37 | 39.44 | 2.7% |
| 8K Sequential Write | 72.79 | 75.49 | 3.3% |
| 16K Sequential Write | 133.36 | 139.36 | 4.3% |
| 64K Sequential Write | 162.54 | 208.34 | 21.9% |
| 128K Sequential Write | 163.00 | 203.82 | 20% |
| 1M Sequential Write | 167.41 | 194.81 | 14.1% |

500 GB vs. 750 GB SATA II Drive Performance Test Results

750 GB disk drives are now available and shipping in ULTAMUS RAID. The 750 GB drives and their predecessor 500 GB drives share many characteristics. These tests were performed to see how the performance of the drives compares in a typical RAID 6 environment.

Our findings were that the 750 GB drives typically performed slightly better when performing sequential read and write operations, but were from 16% to 19% slower than 500 GB drives when performing random access operations.

The drives used in these tests were the same Seagate Enterprise SATA II drives that ship with the ULTAMUS RAID products.

| Access Specification | 500 GB Drives | | 750 GB Drives | | 750GB Performance/ 500 GB Performance |
|-----------------------|---------------|---------------|---------------|---------------|--|
| | IOPS | MB/sec | IOPS | MB/sec | |
| 512B Sequential Read | 61,599.37 | 30.08 | 63,149.05 | 30.84 | 2% |
| 512B Sequential Write | 9,852.35 | 4.81 | 9,962.60 | 4.86 | 1% |
| 512B Random | 1,030.97 | 0.5 | 888.57 | 0.43 | -16% |
| 4K Sequential Read | 42,360.94 | 165.47 | 43,122.36 | 168.45 | 2% |
| 4K Sequential Write | 9,872.50 | 38.57 | 10,095.13 | 39.44 | 2% |
| 4K Random | 1,005.58 | 3.93 | 867.18 | 3.38 | -16% |
| 8K Sequential Read | 29,887.57 | 233.5 | 30,064.42 | 234.88 | 1% |
| 8K Sequential Write | 9,325.26 | 72.85 | 9,662.30 | 75.49 | 3% |
| 8K Random | 974.62 | 7.61 | 837.35 | 6.55 | -16% |
| 16K Sequential Read | 20,523.30 | 320.68 | 20,533.28 | 320.83 | 0% |
| 16K Sequential Write | 8,594.01 | 134.28 | 8,918.84 | 139.36 | 4% |
| 16K Random | 939.2 | 14.67 | 802.7 | 12.54 | -17% |
| 64K Sequential Read | 6,999.23 | 437.45 | 7,021.71 | 438.85 | 0% |
| 64K Sequential Write | 3,470.35 | 216.9 | 3,333.34 | 208.34 | -4% |
| 64K Random | 747.26 | 46.7 | 630.11 | 39.38 | -19% |
| 128K Sequential Read | 3,608.91 | 451.11 | 3,613.36 | 451.67 | 0% |
| 128K Sequential Write | 1,683.77 | 210.47 | 1,630.54 | 203.82 | -3% |
| 128K Random | 583.66 | 72.96 | 498.09 | 62.27 | -17% |
| 1M Sequential Read | 463.64 | 463.64 | 463.2 | 463.2 | 0% |
| 1M Sequential Write | 213.13 | 213.13 | 194.81 | 194.81 | -9% |
| 1M Random | 182.91 | 182.91 | 163.35 | 163.35 | -12% |

Summary

Overland's industry leadership delivers investment protection to storage buyers through leading-edge, "future proofed" technology. ULTAMUS RAID arms storage managers with the tools they need to implement storage strategies and build IT infrastructures that align with today's and tomorrow's business and data management needs.

| Technology | Benefits | Overland Leadership |
|-------------------------------|---|--|
| Hardware Accelerated RAID 6 | <ul style="list-style-type: none"> Higher availability Uncompromised performance | <ul style="list-style-type: none"> Better protection and availability for storage administrators RAID 6 allows storage to remain online if two drives fail simultaneously Overland's Hardware Accelerated RAID 6 delivers RAID 6 protection without sacrificing performance |
| SAS Drives | <ul style="list-style-type: none"> Higher Performance Lower Costs Investment Protection Higher Availability | <ul style="list-style-type: none"> Better for email, web serving, database and other I/O-intensive applications. SAS, the latest enterprise disk technology provides the performance and reliability of Fibre Channel drives at lower cost SAS and SATA II drives can be combined in the same ULTAMUS RAID configuration to support multiple applications with different performance requirements |
| Enterprise SATA II Drives | <ul style="list-style-type: none"> Higher availability Higher performance Better price/TB Investment protection | <ul style="list-style-type: none"> Better for D2D backup, media streaming and other capacity-intensive applications. Built for 7 x 24 operating environments Fewer drive timeouts and fewer RAID rebuilds 1.4 million hour MTBF at 100% duty cycle – more than double the reliability of desktop SATA II drives Enhances availability through Directed Offline Scan (DOS) that monitors every write and performs diagnostic tests when drives are idle Further enhances availability through Perpendicular Recording that reduces the number of moving parts Protects storage investments against the risk of technological obsolescence |
| SAS Expansion and Scalability | <ul style="list-style-type: none"> Grows with your application requirements | <ul style="list-style-type: none"> Effortlessly scales up to 60 drives in 12 drive, 2U increments No performance penalty |
| 4 Gb/sec Fibre Channel | <ul style="list-style-type: none"> Higher performance Investment protection | <ul style="list-style-type: none"> Better performance for all applications. 4 Gb/sec FC improves the performance of IOPS-intensive applications with reduced latency while improving the performance of streaming applications with higher data transfer rates. |
| Snapshots | <ul style="list-style-type: none"> DR and data protection | <ul style="list-style-type: none"> Better value with included snapshots. Create low-impact copies of data that enable point-in-time recovery. |
| Pricing | <ul style="list-style-type: none"> Low acquisition cost | <ul style="list-style-type: none"> Better price/TB for RAID storage Integrated RAID ASIC technology reduces controller costs and an optimized supply chain reduces platform costs. |

ULTAMUS RAID has incorporated the most advanced technologies in the storage industry into a solution that:

- Is ready to win with differentiated technology at a commodity price point
- Provides investment protection for storage managers
- Creates a new class of end-to-end performance in the entry-space market
- Provides a complete tiered storage strategy (ULTAMUS RAID with both SAS and SATA II drives, REO™, NEO™, ARCvault™)

ULTAMUS RAID Application Benefits

ULTAMUS RAID is the right platform to make IT applications run faster, with higher availability and improved ROI.

| Applications | ULTAMUS Features | Business Benefit |
|---|---|---|
| <ul style="list-style-type: none"> • Database • E-Mail • Web Servers | <ul style="list-style-type: none"> • Multipathing for key operating systems • 4 Gb/sec host connection provides high data transfer rates and low latency • SAS drives provide maximum IOPS (based on 15K RPM) at a lower cost than 4GB FC drives • Hardware accelerated RAID 6 improves data availability | <ul style="list-style-type: none"> • More transactions • Lower cost per IOP • Improved operations and lower management cost with higher availability |
| <ul style="list-style-type: none"> • Video • D2D Backup | <ul style="list-style-type: none"> • Multipathing for key operating systems • 4 Gb/sec host connection provides more consistent data transfers for streaming applications • SATA II drives offer the right performance with very high capacities • Hardware accelerated RAID 6 improves data availability | <ul style="list-style-type: none"> • More data protected online for regulatory requirements • Improved cost per TB • Improved operations and lower management cost with higher availability |
| <ul style="list-style-type: none"> • Tiered Storage | <ul style="list-style-type: none"> • Multipathing for key operating systems • Shared SAS/SATA II architecture allows for optimizing cost, performance, capacity, and access requirements • Single box, many uses • Snapshots provide efficient data protection on the right class of storage | <ul style="list-style-type: none"> • Investment protection, asset management, and flexibility • Lower cost per IOP and TB • Improved operations and lower management cost with higher availability |