Comparison of Hybrid Flash Storage System Performance

Author: Russ Fellows

March 23, 2015
Table of Contents

Executive Summary .................................................................................................................. 4
  Evaluation Result Summary ................................................................................................. 4
Comparison of SSD Performance Options .............................................................................. 6
Evaluation Process .................................................................................................................. 7
  Evaluation Overview ............................................................................................................ 7
  Validation Test Objectives .................................................................................................... 7
  Test Cases ............................................................................................................................ 8
  Measuring Performance ......................................................................................................... 8
Test Findings ........................................................................................................................... 8
  Test 1 - Find Utilization of Flash Media ................................................................................. 10
  Test 2 - Find I/O Rate at Specified Response Times ............................................................. 11
  Test 3 - Test Performance of Frequently Accessed Data - With Optimization ......................... 14
  Test 4 - Test Performance of Infrequently Accessed Data ...................................................... 16
  Test 5 - Test Performance Over Time ................................................................................... 17
  Test 6 - Test Sequential I/O Workload Performance ............................................................ 20
Dell SC Series SC4020 ............................................................................................................ 22
Dell PS Series PS6510XS ....................................................................................................... 23
Evaluation Summary .............................................................................................................. 24
  Issues and Concerns .............................................................................................................. 25
  Final Observations ................................................................................................................ 25
Appendix A – Configuration Overview ................................................................................... 26
  Hardware Environment .......................................................................................................... 26

© 2015 Evaluator Group, Inc. All rights reserved. Reproduction of this publication in any form without prior written permission is prohibited.
Software Testing Applications ................................................................. 27

Appendix B – Storage Workload Details ..................................................28

Appendix C - Test Case Details ...............................................................29
Workload Generation Overview ............................................................... 29

Appendix D - Configuration and Price Comparison .................................32
Comparison of List Price of Configured System ....................................... 32

Appendix E - Test Result Details .............................................................33
Test 3 Data Details .................................................................................. 33
Executive Summary

Storage systems have changed dramatically over the past five years, primarily due to the benefits of utilizing solid-state technologies. While high-end enterprise systems have seen some advances due to solid-state, midrange storage systems have seen the most dramatic enhancements.

Every organization has critical applications, many of which can benefit from the use of solid-state technology to reduce storage latencies and increase application performance. Larger organizations often have multiple applications that need storage performance, along with substantial capacity needs. In these cases, utilizing an all-flash storage system can be the best option. For mid-sized and smaller organizations, a better option is deploying a hybrid storage system that combines large capacities of spinning media with the acceleration potential of solid-state.

For hybrid storage systems, there are two approaches to placing data on the appropriate media: tiering (which moves data) or caching (which makes a copy of data). Both of these approaches can provide enhanced performance and lower cost capacity than all-flash storage systems. However, tiering and caching are optimized for different workload patterns, providing significantly different results for primary and secondary applications.

In this lab evaluation, three systems were tested, with a focus on midrange enterprise storage products. The systems tested included a Dell Storage PS Series system and a Dell Storage SC Series system. These two systems were compared against a competing midrange SAN storage product of a similar configuration and price. The two systems tested included a hybrid SSD+10Krpm array for the PS Series and an all-flash array for the SC Series. The testing did not include additional NL-SAS capacity with the SC4020.

Multiple test case scenarios were constructed in an effort to highlight the performance of each system while running the same workload. These tests utilized synthetic workload generation tools, which are detailed in the Appendices.

We evaluated three systems and characterized their performance as follows:

- Dell SC 4020 (all-flash): Excellent performance for entire capacity range tested
- Dell PS6210XS (hybrid): Good performance for both SSD and spinning media tier
- Competitor (hybrid): Good performance for SSD cached data only

Evaluation Result Summary

We found the Dell PS Series system outperformed the competitor in nearly every test conducted and the Dell SC Series system significantly outperformed the competitor in every test. These results are due to several significant architectural differences between the products, which were contributing factors to the results achieved. The remainder of this report details these findings and test conditions.
The testing was designed to recreate the workloads produced from multiple virtual servers running database and other on-line transaction processing types of workloads. As explained in detail in the remainder of this document, the tests were designed to produce high amounts of random I/O using mixed block sizes and mixed read vs. write ratios. These I/O patterns are common characteristics of virtual server database applications.

Some of the results include:

• For the hybrid systems, there were significant performance differences seen when data resided on SSD media vs. when it resided on rotating media.
• For workloads residing on rotating media, there are significant performance advantages when using enterprise 10k SAS drives, compared to using 7.2k SATA drives.
• Although caching and tiering can mitigate the effects of slow media, they are unable to do so after the effective capacity of the SSD cache / tier is exhausted.
• The movement of data between two tiers took multiple days to complete, while in most cases the movement of data into cache took 24 hours to complete.
• Due to the time to move data, neither caching nor tiering provided immediate benefits to changing workloads.
• For the workloads tested, both the Dell PS6210XS and SC4020 systems outperformed the equivalently configured and priced competitor by significant amounts.

For details on the test cases utilized, the data sets, testing tools and hardware utilized refer to the Appendices A - E.
Comparison of SSD Performance Options

Currently, three primary methods are available to enhance storage system performance by utilizing solid-state NAND Flash storage technologies: 1) Cache reads, writes or both to solid-state; 2) Automatically move data between HDD and solid-state tiers of storage, or 3) deploy an all solid-state system, typically called an all-flash system.

Each method has benefits and shortcomings; each is typically optimal for a specific type of workload or application environment.

<table>
<thead>
<tr>
<th>SSD for Cache or Capacity</th>
<th>Auto Tiering</th>
<th>Caching</th>
<th>All Flash</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSD is primary location of data, SSD adds capacity</td>
<td>SSD is an additional copy of data, SSD does not add capacity</td>
<td>SSD is only location for all data</td>
<td></td>
</tr>
<tr>
<td>Time to optimize Primary Applications</td>
<td>New data placed onto SSD first, hot data migrated over multiple days</td>
<td>New data retained in read cache, hot data fills cache in 18 - 24 hour period</td>
<td>All capacity is high performance, no time required</td>
</tr>
<tr>
<td>Impact of Secondary Applications on System</td>
<td>Primary applications remain in SSD tier for multiple days</td>
<td>New apps quickly displace data, resulting in slow primary applications</td>
<td>Applications all reside on Flash, high performance with limited capacity</td>
</tr>
<tr>
<td>Optimal Workload</td>
<td>Mixed workloads with a portion of data frequently accessed.</td>
<td>Mixed workloads with a portion of data frequently accessed.</td>
<td>Random I/O workloads, with high I/O rates relative to capacity</td>
</tr>
<tr>
<td>Application Examples</td>
<td>Workloads with a mixture of high I/O and large capacity</td>
<td>Workloads with a mixture of high I/O and large capacity</td>
<td>Workloads with a consistently high I/O and smaller capacity</td>
</tr>
<tr>
<td>Cost Effectiveness</td>
<td>Excellent performance for entire SSD Tier, excellent capacity</td>
<td>Excellent performance for recently cached data, good capacity</td>
<td>Excellent performance, poor capacity efficiency</td>
</tr>
</tbody>
</table>

Table 1: Comparing Storage SSD Options for Performance
Evaluation Process

In order to provide a substantive evaluation based on measurements, Evaluator Group together with Dell developed a series of test cases to show system performance for typical workloads encountered in midrange SAN storage environments.

Three different systems were tested, including a Dell PS Series system and a Dell SC system, along with a similarly configured competing midrange storage system. All three systems are hybrid designs, incorporating both spinning HDD media and solid-state NAND Flash media to accelerate workloads.

Evaluation Overview

Testing occurred in December 2014 and January 2015, with Evaluator Group having access to all systems being tested. Testing occurred in Dell labs, using Dell equipment for all tests. Both Dell and Evaluator Group personnel performed the tests, with all testing supervised by Evaluator Group and an additional on-site audit and review of all tests and equipment.

This test validation was performed in order to provide IT administrators and architects with actual data points to help them evaluate and consider the tradeoffs of choices available when choosing storage for virtualized and application environments.

Evaluator Group comments: The tests utilized during this validation were intended to re-create real-world environments and use cases as closely as possible. Although a single test cannot highlight every use case, the scenario tested here is likely to be found in most IT environments.

In instances noted, Evaluator Group established test parameters and the results were reviewed after being run by Dell personnel. This report highlights the results found while performing testing. Evaluator Group commentary provides context and a narrative assessment of the results as experienced by Evaluator Group personnel. Details of the testing, along with the tools are provided in Appendices.

Validation Test Objectives

The following aspects were the elements the tests were designed to highlight:

- Efficient use of solid-state storage media to accelerate workloads
- Maximum capacity available to workloads, operating from high-speed SSD media
- Ability for system to maintain optimized performance for long-running primary applications
- Performance for secondary applications and other shorter duration workloads
- Overall performance of systems running multiple random workloads
- Streaming I/O comparison for analytic workloads
Test Cases

Testing was designed to measure relevant performance information of each system, and to provide information that would allow assessment of the validation objectives previously specified. The following test cases were used for evaluation:

1. Find effective size of SSD cache or tier
2. Find I/O rate while achieving response time threshold
3. Test the performance of a frequently accessed portion of data
4. Test the performance of infrequently accessed data
5. Test performance of a volume after accessing infrequent data
6. Test how quickly data can be moved back into cache or SSD tier (hybrid systems only)
7. Test a sequential I/O workload

In each instance, the test was run on the following systems:

- Dell PS Series 6210XS - hybrid design with SSD and 10k rpm SAS
- Dell SC Series 4020 - hybrid or all SSD, tested all SSD configuration
- Competitor - hybrid design, 7.2k rpm Nearline SAS with read caching on SSD

Measuring Performance

There are multiple methods of measuring storage performance. However, the three most significant attributes are the I/O rate, the response time and the throughput. Each of these affects the other attributes, along with the block size(s) used for performing I/O operations. It is impossible to optimize each of these aspects simultaneously, and typically optimization is made either to achieve a specific response time or to achieve a specific I/O or throughput rate.

For the I/O rate tests (Test #1 - 5), we chose to optimize the response time first, establishing a threshold of maximum average response time rate and then measuring the I/O rate for the given workload. In contrast, for the throughput test (Test #6) the response time was not used and instead an unbounded test was run.

The test setting details are provided in Appendix C.

The performance advantage between the all flash SC4020 and the other systems is substantial across all data points, including when 100% of I/O was going to a portion of the volume being tested.

Test Findings

As previously indicated, the Dell PS series outperformed a competing hybrid system in nearly every test case. Not only did the PS series provide higher I/O rates, it delivered the I/O’s with lower latency. The Dell SC series significantly outperformed the competing system in every test, with response times that
were also substantially lower. Multiple data points throughout the remainder of this section provide the basis for our findings.

_Evaluator Group comments: Our testing found that for sustained workload performance, both Dell systems provided better performance as measured by higher I/O rates, lower latencies and higher throughput than the competing system._

Overall, the evaluation of these three different products showed that there were differences in performance between the systems. Both the Dell PS and Dell SC Series systems utilize a tiering mechanism to accelerate applications. The competing system utilizes a caching mechanism to accelerate applications; copying frequently accessed data onto SSD devices for read workloads while utilizing a solid-state write cache to buffer incoming writes. The effective capacity of the solid-state storage was a significant factor in determining performance, with the spinning media utilized also contributing.

_Evaluator Group comments: The effective capacity of solid-state was a primary factor for performance differences between the systems. Due to the higher effective capacity of solid-state the Dell PS series outperformed the competitor; the SC series utilized all solid-state and provided even higher performance levels. A secondary factor was the higher-performing spinning media in the Dell PS series vs. the competitor’s use of NL-SAS media._

An overview of results is shown below, with full details provided in Appendix E:

- Dell SC4020 outperformed the Competitor’s system in all cases ranging from 3 X - up to 68 X better
  - 68 X _higher_ performance (when 100% of I/O to cold 1.5 TB)
  - 14 X _higher_ performance (when 90% of I/O to hot 1.5 TB)
  - 4X _better_ response time (Max I/O rate at 5 ms. vs. 20 ms. for competitor)
  - 3.1 X _higher_ performance (when 100% of I/O to hot 1 TB)
- Dell PS6210XS outperformed the Competitor’s system in most cases, from 2 X -to 4 X better, with the exception of workloads with less than 1 TB of hot data
  - 2.7 X _higher_ performance (when 100% of I/O to cold 1.5 TB)
  - 4.3 X _higher_ performance (when 90% of I/O to hot 1.5 TB)
  - 2X _better_ response time (Max I/O rate at 10 ms. vs. 20 ms. for competitor)
  - 3.3% _lower_ performance (when 100% of I/O to hot 1 TB)

_Evaluator Group comments: In the majority of test cases using a single 6TB volume, both the Dell SC4020 and the Dell PS6210XS system outperformed the competing system, often by substantial margins. For a majority of workloads and use cases, it is clear that the Dell configured systems provided better performance._
The majority of the tests utilized two workloads, running against different portions of a single volume. These workloads were denoted as “hot” and “cold” based on their potential for caching or tiering, with the hot zone detailed in Appendix B and workloads detailed in Appendix C.

**Test 1 - Find Utilization of Flash Media**

The goal of this test was to find the maximum useable capacity of SSD media on each system and to report results as a percentage of the configured SSD capacity. Finding the maximum useable capacity was accomplished by configuring a volume and running a random I/O workload against it; this allowed the system sufficient time to cache or move the data to the SSD media tier. Once performance increases stopped, the system had optimized the volume. The process was repeated with a larger volume size until performance no longer matched previous volumes’ performance, indicating SSD capacity had been exhausted.

**Test Setup and Operation:**

1. Create a LUN / Volume on each storage system
   a. Run a random read / write workload until all data resides on SSD media
   b. Note cache hit percent as indicated by response times
2. Remove volume created in step 1 and increase size of volume
3. Repeat until the volume no longer resides on SSD media
   a. Response times will be significantly higher, with lower I/O rate once SSDs are consumed
   b. Note largest size of LUN / volume where data resided on SSDs

**Note: This test case was performed by Dell and reviewed during its progression by Evaluator Group**

**Results**

**Competitor**

The maximum amount of usable SSD capacity was found to be less than 1 TB. This resulted in approximately 41% of the reported available 2.4 TB of SSD capacity:

- 75% cache hit maximum for 1.2 TB volume = 0.90 TB cache maximum
- 62% cache hit maximum for 1.6 TB volume = 0.99 TB cache maximum
- Result = Usable cache is less than 50% of available cache (1 TB vs. 2.4 TB)

**Dell PS Series 6210XS**

The physical size of the SSD’s used for the performance tier is 4 TB. Our testing determined that the effective size is approximately 2.5 TB.

- 90% of I/O to 1.5 TB had same I/O rate as 100% of I/O to a 1.5 TB volume
- 100% of I/O to 6 TB had significantly lower I/O rate than 1.5 TB region
- In test 5b (later in report), it was found that 2.5 TB was able to be accelerated
**Dell SC Series 4020**

The amount of usable SSD capacity was at least 6 TB.
- 100% of the 6 TB volume resided on the configured SSD capacity

Shown below in Figure 1 is a chart showing the cache hit rate compared to the primary volume being accessed. Also plotted is the Effective Cache Rate. As seen on the left of the graph at the 0.75 data point, the cache hit approaches 100% when the capacity accessed is 0.75 TB, while the “Effective Cache Size as a % of RAW SSD” value at that point is approximately 30%. At 1 TB, the effective size is 40% with a cache-hit rate of 96%. Several values are plotted for tested capacity points. The lowest cache hit rate was 26%, for a 6 TB volume where the effective use was 65% of the SSD capacity.

![Cache Hit % vs Volume Size](image)

**Figure 1: Competitor - Cache Hit Rate vs. Volume Size**

**Test 2 - Find I/O Rate at Specified Response Times**

The objective of this test was to find the maximum I/O rate for each system while achieving a specific average response time. As with the other tests, the response time was measured on the host rather than on the storage system, thus ensuring complete round-trip measurement for all I/O operations.

As shown below in Figures 2 and 3, the Dell PS series and Dell SC series systems provided higher I/O rates at each response time threshold tested.
Testing these systems utilized a workload as follows:

- Mixed block sizes, avg. 22 KB, with 60% read, 40% write to 6 TB volume
- 90% of I/O to 1.5 TB hot-band, with 95% random workload
- 10% of I/O to remaining 4.5 TB band, with 90% random workload

![Relative I/O Rate vs. Response (90% I/O to 1.5 TB)](image)

**Figure 2: Dell PS vs. Competitor (Relative I/O at Response Time Thresholds)**
As shown above, the Competitor’s system delivered essentially zero (0) I/O’s when a 5ms. read and write response time threshold was enforced. Lower response times are required in order to increase performance of applications requiring I/O operations to be completed in order, as is common with OLTP and database applications. Thus, moving the response time threshold from 20 ms. to 5 ms. can provide a 4X performance improvement in these types of applications.

**Results**

**Competitor**
- Zero (0.0) I/O operations at a 5 ms. threshold

**PS Series 6210XS vs. Competitor**
- \(\frac{2.96}{0.52} = 5.7X \) more I/O at 10 ms. threshold
- \(\frac{2.96}{0.69} = 4.3X \) more I/O at 20 ms. threshold

**SC Series 4020 vs. Competitor**
- \(\frac{10.0}{0.52} = 19.2X \) more I/O at 10 ms. threshold
- \(\frac{10.0}{0.69} = 14.5X \) more I/O at 20 ms. threshold
Test 3 - Test Performance of Frequently Accessed Data - With Optimization

In this test, each system has I/O directed to a portion of the 6 TB volume tested. The workload was run until the system stopped increasing performance or its cache hit percentage. In this way, each system was allowed unlimited time to optimize the workload to achieve optimal performance.

Shown below in Figure 4 is a comparison of the Dell SC4020 to the competing system. The left legend shows the normalized I/O rate, with the right legend indicating the cache hit rates for the competitor. The cache hit rate shown is as reported by the Competitor’s system during testing.

![Relative I/O Rate @ < 20ms Avg. Resp. 6TB LUN](image)

Figure 4: Dell SC4020 vs. Competitor

For this test, the three rightmost data points in Figure 4 and Figure 5 below were created with the following workloads:

- Mixed block size, avg. 22 KB, with 60% read
- 95% random I/O for “hot-band” and 90% random for “cold-band”
- 90% of I/O to 1.5 TB hot-band, 10% of I/O to cold-band
- 100% of I/O to 1.5 TB hot-band, no cold-band
- 100% of I/O to 0.75 TB hot-band, no cold-band
- Compressible data, at 2:1 compressibility setting
- Note: See Appendices B & C for workload details
Shown below in Figure 5 is a similar chart, comparing the competitor to the Dell PS6210XS.

![Relative I/O Rate @ < 20ms Avg. Resp. 6TB LUN](chart)

**Figure 5: PS6210XS vs. Competitor**

### Results

For these tests, a response time threshold of 20 ms. was utilized. As outlined previously, both the Dell SC4020 and PS6210XS outperformed the Competitor for a hot-zone of 1.5 TB. For the case of a 1 TB hot-zone, the Competitor outperformed the PS6210XS but still did not match the performance of the SC4020.

#### Competitor

After allowing time for performance to stabilize at maximum rate:

- Achieved a normalized I/O rate of 0.7 for 90% of I/O to 1.5 TB with 71% cache hit
- Achieved a normalized I/O rate of 1.6 for 100% of I/O to 1.5 TB with 88% cache hit
- Achieved a normalized I/O rate of 3.8 for 100% of I/O to 1.0 TB with 96% cache hit

#### Dell PS Series 6210XS

After allowing time for performance to stabilize at maximum rate:

- Achieved a normalized I/O rate of 3.0 for 90% of I/O to 1.5 TB
- Achieved a normalized I/O rate of 3.0 for 100% of I/O to 1.5 TB
- Achieved a normalized I/O rate of 3.0 for 100% of I/O to 1.0 TB
**Dell SC Series 4020**

After allowing time for performance to stabilize at maximum rate:

- Achieved a normalized I/O rate of 9.8 for 90% of I/O to 1.5 TB
- Achieved a normalized I/O rate of 9.8 for 100% of I/O to 1.5 TB
- Achieved a normalized I/O rate of 9.8 for 100% of I/O to 1.0 TB

**Test 4 - Test Performance of Infrequently Accessed Data**

This test was a continuation of the previous testing, and examined the performance of each system as a greater percentage of I/O was directed to a region that previously had little access. The order of testing began with 90% of I/O to a small region until a maximum I/O rate was achieved at a steady state. This indicated that optimization had completed and no further performance gains could be achieved.

_Evaluator Group comments: This test highlights the performance capabilities of systems when a small percentage of data resides in an SSD accelerated tier or cache. This may occur if an additional workload displaces the primary application from the accelerated media. As shown, the results show a 2 - 3X advantage for the PS series vs. the competitor and a 40 - 80X advantage for the SC series vs. the competitor. These are significant differences._

For this portion of testing, the tests began with 100% of I/O to the entire region, thereby leveraging the portion of data that had already been cached, tiered or otherwise accelerated. Next, 100% of I/O was run to the previous cold-band region of 4.5 TB, then 100% of I/O to a 1.5 TB portion of the cold-band.

For this test, the three leftmost data points were created with the following workloads:

- Mixed block size, avg. 22 KB, with 60% read, 95% random
- 100% of I/O to entire 6 TB volume, no warm-up period
- 100% of I/O to 4.5 TB cold-band, no warm-up period
- 100% of I/O to 1.5 TB cold-band, no warm-up period
- Note: See Appendices B & C for workload details

**Results**

As with the previous test, the response time threshold utilized was 20 ms. Both the Dell SC4020 and PS6210XS outperformed the competitor for workloads to a cold zone.

**Competitor**

_Note: The Competitor’s results were achieved with average response times < 20 ms._

Without allowing time for performance to stabilize at maximum rate:
• Achieved a normalized I/O rate of 0.2 for 100% of I/O to 6 TB with 26% cache hit
• Achieved a normalized I/O rate of 0.1 for 100% of I/O to 4.5 TB with 8% cache hit
• Achieved a normalized I/O rate of 0.1 for 100% of I/O to 1.5 TB with 8% cache hit

**Dell PS Series 6210XS**

*Note: The PS Series results were achieved with average response times < 20 ms.*

*Without* allowing time for performance to stabilize at maximum rate:

• Achieved a normalized I/O rate of 0.6 for 100% of I/O to 6 TB
• Achieved a normalized I/O rate of 0.4 for 100% of I/O to 4.5 TB
• Achieved a normalized I/O rate of 0.3 for 100% of I/O to 1.5 TB

**Dell SC Series 4020**

*Note: The SC Series results were achieved with average response times < 5 ms.*

*Without* allowing time for performance to stabilize at maximum rate:

• Achieved a normalized I/O rate of 8.3 for 100% of I/O to 6 TB
• Achieved a normalized I/O rate of 8.3 for 100% of I/O to 4.5 TB
• Achieved a normalized I/O rate of 8.3 for 100% of I/O to 1.5 TB

**Test 5 - Test Performance Over Time**

This test is similar to Test #3. However, the primary objective of this test was to test the performance of the Dell PS series compared to the Competitor’s system over time. Since this test is designed to find how caching and tiering changes performance over time, the SC series system was not tested, as it was already previously demonstrated in earlier tests that the all flash SC4020 system performed well across multiple workloads, with both new or cold data, as well as highly accessed or hot data sets.

**Test 5a - New Data Set Results**

For Test 5a, the workload consisted of the following data patterns:

• Mixed block size, avg. 22 KB, with 60% read, 95% random
• 100% of I/O to 1.5 TB cold-band, no warm-up period
• Note: See Appendices B & C for workload details

**Competitor**

The system was allowed to run past the time when it achieved steady results, in order to correspond to the Dell PS system at 48 hours.

• Achieved a peak normalized I/O rate of 2.88 at 22 hours into testing
• Subsequently leveled out with a steady-state I/O rate of 1.7 from 36 hours on

**Dell PS Series 6210XS**

The system was allowed to run until it achieved steady results, at 48 hours.
• Achieved a peak normalized I/O rate of 3.1 at 47 hours into testing
• Delivered steady-state normalized I/O rate of 3.1 from 47 hours on

![Normalized I/O Rate @ < 20 ms, 48 Hour Run](image)

**Figure 6: Normalized I/O rate vs. Hours - New Data Set**

It should be noted that for this test, the maximum allowed average response time was 20 milliseconds. The I/O rates remained unchanged for both systems past 48 hours.

**Test 5b - Alternating Data Set Results**

For Test 5b, there were two, alternating workloads. There were a total of 4 regions tested in total. I/O to Hot2 and Cold2 were never zero but were greatly reduced while running workloads Hot1 and Cold1. The sizes of these regions differed from that of other tests described in this report.

Testing began after both systems had run both workload 1 and workload 2 for 48 hours.

All Test Data:
• Mixed block size, avg. 22 KB, with 60% read, 2:1 compressible data
• All I/O to “hot” zones was 95% random, all I/O to “cold” zones was 90% random
• Ratio of I/O for hot and cold was 90% to the hot zone, and 10% to the cold zone

**Workload 1:**
• No warm-up period after storage had reached a steady-state on workload 2
• Hot1 and Cold1 had 95% of total I/O, with remaining 5% of I/O going to Hot2 and Cold2
  ○ 85.5% of I/O to Hot1 (1.25 TB)
  ○ 9.5% of I/O to Cold1 (1.75 TB)
  ○ 4.5% of I/O to Hot2 (1.25 TB)
  ○ 0.5% of I/O to Cold2 (1.75 TB)
Workload 2:
- No warm-up period after storage had reached a steady-state on workload 1
- Hot2 and Cold2 had 100% of total I/O, with 0% of I/O going to Hot1 and Cold1
  - 0% of I/O to Hot1 (1.25 TB)
  - 0% of I/O to Cold1 (1.75 TB)
  - 90% of I/O to Hot2 (1.25 TB)
  - 10% of I/O to Cold2 (1.75 TB)

**Competitor**

Testing began with workload 1 for 48 hours, and then changed to workload 2 with the following results:
- Achieved a peak normalized I/O rate of 1.06 at 19 hours into testing for workload 1
- Subsequently leveled out with a steady-state I/O rate of 0.8 from 36 hours on
- Achieved a peak normalized I/O rate of 1.03 at 21 hours into testing for workload 2
- Subsequently leveled out with a steady-state I/O rate of 0.9 from 28 hours on

**Dell PS Series 6210XS**

Testing began with workload 1 for 48 hours, and then changed to workload 2 with the following results:
- Delivered steady-state normalized I/O rate of 3.1

---

_Evaluator Group comments:_ This test produced interesting results. First, the caching approach of the competitor did not warm the cache as quickly as literature would leave a user to believe. Additionally, the spike in performance for the competitor at approximately 24 hours into the testing, with a subsequent 50% fall off in amount of performance was unexpected. These results were re-created several times, to ensure that it the results were not due to other factors. In summary, up to 17 hours of testing, both the PS6210XS and competitor produced similar results. From 17 hours to 29 hours, the competitor had better results, and then from 30 hours onward, the PS6210XS again provided increasingly better results up to 48 hours, at which time both systems achieved steady-state.
Shown below in Figure 7 are the normalized I/O Rates for the alternating workloads described above.

![Normalized I/O Rate - Alternating Workloads](image)

**Figure 7: Normalized I/O rate vs. Hours - Alternating Data Sets**

**Test 6 - Test Sequential I/O Workload Performance**

The final test was designed to measure the throughput rate of the storage system while using large, sequential block transfers. This test examined the Competitor’s system against the Dell PS6210XS. The test details are provided in Appendix C. In both instances, the workload was tested against 4 volumes, each approximately 50 GB in size.

**Note: This test case was performed by Dell and reviewed during its progression by Evaluator Group**

**PS Series vs. Competitor Results**

No attempt was made to direct what media was used for these tests. In both cases, the storage system determined how data was placed. The aggregate throughput results are charted below in Figure 5.
As shown above, in all cases the Dell PS Series system outperformed the Competitor, both for reads as well as writes.

**SC4020 Results**

The sequential workload test was not performed on the Dell SC4020, due to the fact that the connectivity bandwidth was 4 @ 8 Gb FC vs. 2 @ 10 Gb iSCSI for the other systems. For bandwidth testing the SAN connectivity would provide a significant advantage for the SC4020, and as a result a head to head comparison was not run.
Dell SC Series SC4020

Since their inception, Dell SC Series systems have been designed to virtualize storage capacity into multiple tiers. The built-in support for tiering has been enhanced with support for multiple SSD types, along with multiple types of spinning media with auto-tiering capabilities to move data between media types based on access patterns.

- **Architecture**
  - SAN storage with iSCSI and FC connectivity
  - Devices may reside in a single virtualized pool, or in separate pools
  - Multiple device types / tiers supported (up to 12)
    - 7.2K NL-SAS, 10K SAS, 15K SAS, SSD, RAID levels, etc.
  - Data progression (auto-tiering) supported on all volumes
  - Includes all Advanced Storage Features
    - Thin provisioning, snapshots, clones, replication, etc.

- **Data Progression (Auto Tiering)**
  - Default is 3 days to promote data, 12 days for demotion and real-time for SSD
  - Reports built-in to system, including amount of data to be moved up or down next cycle

---

![Diagram of Dell SC Series - Auto Tiering Architecture](image_url)

**Figure 9: Dell SC Series - Auto Tiering Architecture**

© 2015 Evaluator Group, Inc. All rights reserved. Reproduction of this publication in any form without prior written permission is prohibited.
Dell PS Series PS6510XS

The Dell PS Series line is designed as a scale-out system with the following features:

- **PS Series Scale Out SAN systems**
  - Scale out design, multiple nodes form a distributed “Peer Group”
  - Load balancing across members within a group of up to 16 nodes
  - Each node is a 2 controller, HA system with drive media
  - Uses iSCSI connectivity
  - Drives are HDD and SSD using SAS interface
  - Includes all advanced Storage Features
    - Thin provisioning, snapshots, clones, replication, etc.

- **XS system sub-LUN auto tiering**
  - Storage pool is defined as a tier
  - Storage media type comprises a pool, with volumes assigned to a pool / tier
  - Two tiers supported – SSD and 10K SAS drives
  - Classification of each page for I/O rate: high, medium or low
  - Data moved between tiers based on workload class
  - Movement done in the background automatically
    - Hours to multiple days for movement, depending on workload
  - A portion of SSD is reserved as incoming write cache
Evaluation Summary

There are debates over the best design for providing performance and capacity cost effectively. Alternatives include cache-centric designs and tiering designs. Each approach can deliver performance and capacity, but does so differently. By utilizing three similarly priced configurations all competing in the same market segment, it was possible to directly observe the performance advantages of these specific systems while running the same workloads.

Furthermore, a critical part of the test was choosing representative workloads and data set sizes. Although a synthetic workload was created, it was carefully chosen to be similar to the workloads created by virtualized database applications generated by IOmark-VM and VMmark. Another factor was the use of a data set size larger than the SSD Tier or Cache of either the Competitor or the PS Series system, thereby forcing the systems to utilize some portion of their rotating media. This also exposed the efficiency of the methods for moving data and choosing data to place on SSD.

The comparisons to the SC4020 appear less obvious, since the tested configuration was an all SSD system. However, the fact that it supports tiering and was priced similarly to the other systems while using only 50% of its SSD capacity makes this system a valid comparison.

Although it is possible to design corner case tests that favor nearly any system, the Competitor’s system was still nearly 3X slower than the SC4020 even in its best case of 100% of data going to less than 5% of capacity. This indicates that the SC4020 provides consistently higher performance in all test cases. In this same test, the Competitor did outperform the PS6210XS, although not by a substantial amount. In nearly all other tests, both Dell solutions outperformed the Competitor significantly.

Additionally, the response times, or latencies, had a bi-modal distribution for both PS Series and the Competitor. This is due to the fact that two media types were being accessed with dramatically different response time characteristics. However, one difference seemed to be attributable to the faster enterprise 10K SAS drives rather than nearline SAS. Due to the relatively high response times of the Competitor’s spinning media, it was necessary to lower the I/O rate for the competitor in order to achieve an overall average response time of less than 20 ms. This was not as big a factor for the PS Series, which required lowering the I/O rate from maximum, but not nearly to the same degree.

None of these factors played a role in the response times for the SC4020, which utilized only SSD for testing. However, SSD wear leveling does impact performance, and there was a drop-off in performance during over-write operations. The lower rate was reported, rather than the higher initial I/O rate. Thus, for I/O bursts of several seconds the SC4020 could achieve higher I/O rates, up to 25% higher than those reported.

Another observation of the results is that when the Competitor’s system is operating near its maximum I/O rate, with a high cache hit rate the spinning media are handling a significant amount of
I/O, since all writes are being flushed to these devices. As a result, any data that does not reside in cache will experience significant I/O latencies, since the spinning media are already near their maximum queue depth limit. This is different than the tiering design with the Dell PS Series, in which case the majority of I/O will be performed exclusively to the SSD tier, leaving the spinning tier with a relatively light load and able to perform read and write operations to the un-accelerated data with acceptable latencies.

Issues and Concerns

Although the test cases utilized a range of configurations, they did not test all use cases and workloads that might occur in customer environments. There may be workloads that utilize a small portion of active data, resulting in instances where the Competitor caching storage system may outperform the tiering design of the Dell PS6210XS. However, it is not expected that there are any realistic workloads or scenarios where the competing system would outperform an all SSD SC4020 system.

Final Observations

Rotating media performs substantially slower than SSD devices do, despite the effect of read or write caching or tiering. Although both caching and tiering can mitigate the effects of slow media in many cases, they cannot do so in all cases. Therefore, for random workloads, the more SSD capacity and devices a system has, the better its performance.

When I/O is performed to rotating media, there are measurable differences based on the type of media. Nearline, 7,200 rpm media do not perform small, random block I/O as fast as 10,000 rpm media does. This is one reason why when I/O occurred to rotating media the PS Series solution outperformed the competing system.

For high throughput transfers, the use of SSD and the type of rotating media used remain important factors. The use of fewer large Nearline media cannot match the throughput rate of a greater amount of faster media, even with sequential large block transfers. The advantages of the SC4020 included all SSD and significantly higher connectivity, 32 Gb FC versus the 20 Gb iSCSI.

The majority of virtual applications and database applications are highly random and latency sensitive. As a result, it is critical to choose a storage platform that can cost-effectively deliver the necessary performance. We found that for the workloads tested, both the Dell PS6210XS and SC4020 systems outperformed an equivalently configured and priced competitor by significant amounts.
Appendix A - Configuration Overview

Hardware Environment

Hypervisors

- VMware ESXi version 5.5 was used on the two physical workload servers
- VMware vCenter version 5.5 for management of hosts and VMs

Guest OS

- Guest VMs utilized Ubuntu 12.04 LTS OS

Physical Servers for PS Series & Competitor testing

- A total of 3 servers were available to generate application workloads
- System 1: Dell R710, with 2 CPU sockets, PCIe bus
  - 2 x 4 Intel CPU’s with Hyper-threading (8 @ 2.27 GHz E5520 CPU’s)
  - 128 GB RAM
  - 2 x 10 Gb Ethernet connections to SAN network
  - 2 x 1 Gb Ethernet connections to LAN management networks
  - OS: VMware vSphere 5.5
- System 2: Dell R710, with 2 CPU sockets, PCIe bus
  - 2 x 4 Intel CPU’s with Hyper-threading (8 @ 2.27 GHz E5520 CPU’s)
  - 48 GB RAM
  - 2 x 10 Gb Ethernet connections to SAN network
  - 2 x 1 Gb Ethernet connections to LAN management networks
  - OS: VMware vSphere 5.5
- System 3: Dell R710, with 2 CPU sockets, PCIe bus
  - 2 x 4 Intel CPU’s with Hyper-threading (8 @ 2.27 GHz E5520 CPU’s)
  - 96 GB RAM
  - 2 x 10 Gb Ethernet connections to SAN network
  - 2 x 1 Gb Ethernet connections to LAN management networks
  - OS: Windows Server 2012R2

Physical Server for SC Series testing

- System 1: Dell R320, with 1 CPU sockets, PCIe bus
  - 1 x 4 Intel CPU’s with Hyper-threading (4 @ 1.8 GHz E5-2403 CPU’s)
  - 32 GB RAM
  - 4 x 8 Gb FCP connections to FC SAN network (Qlogic HBA)
  - 2 x 1 Gb Ethernet connections to LAN management networks
  - OS: Windows Server 2008 R2
Networking

- SAN: Dell, 10 Gb Ethernet switch
- LAN: Dell, 1 Gb network as used to connect all hosts for management

Storage

Dell PS Series

- Dell PS Series 6210XS
  - 2 controllers, each with 2 @ 10 Gb iSCSI connectivity
  - 1 shelf with 7 @ 800 GB SSD disks + 17 @ 10K 1.2TB, total of 24 devices
  - 5.6 TB SSD raw + 20.4 TB HDD capacity = 26 TB raw
  - OS version 7

Dell SC Series

- Dell SC Series SC4020
  - 2 controllers, each with 4 @ 8 Gb FC connectivity
  - 1 shelf with 24 drive slots
  - 12 @ 1.6 TB Read Intensive SSD’s (11 in RAID group, 1 hot spare)
  - 11 SSD’s configured as RAID 10 (7.3 TB usable)
  - 12 @ 3 TB NL-SAS (not tested)

Competitor

- Competing model (Top of line controller as of June 2014)
  - 2 controllers, each with 2 @ 10 Gb iSCSI connectivity
  - 1 shelf with 16 drives, 24 TB raw capacity
  - 4 @ 600 GB SSD read cache + 12 @ 2 TB NL-SAS backing storage
  - Firmware was most recent at time of purchase (June 2014)

Software Testing Applications

The workloads used to test the configurations were generated using two common synthetic workload generation tools: iometer and vdbench.

iometer

The iometer tool was used to generate and measure a portion of the workloads, including Test #5 for sequential data throughput rate testing.

vdbench

The vdbench tool was used to generate a majority of tests, primarily those that utilized a hot and cold band with different workloads running against each portion. Version 5.03 of vdbench was used.
Appendix B - Storage Workload Details

As shown below, the primary volume configuration used for the I/O rate and latency sensitive testing was a single, 6 TB volume. The tests utilized two different workload configurations, each directed to utilize a different range of the same volume. A so-called “Hot” band or region used the first 25%, or 1.5 TB. The “Cold” band utilized the remaining capacity, or 4.5 TB. This is depicted below.

![Diagram of Hot and Cold Regions](image)

Figure 10: Volume configuration used for testing two workload sets

This configuration is meant as an approximation of observed application access patterns across many different application and workload types. It is also similar to the well-known “Pareto” distribution\(^1\), which indicates that 80% of I/O is directed to 20% of capacity.

The total volume and hot-band sizes were chosen such that the hot-band could theoretically reside entirely in the SSD Tier or SSD Cache of all three systems tested, as 1.5 TB is less than the stated usable SSD capacity of each system.

**Note:** For the large block, sequential throughput test (Test #6), a single 50 TB volume was used rather than the configuration depicted above.

**Note 2:** For Test 5b, the size of the hot and cold regions was different, as denoted in Test 5b.

---

Appendix C - Test Case Details

Workload Generation Overview
As detailed previously, both iometer and vdbench workload generation tools were utilized.

Block Size Distributions Utilized for Random Workloads
Shown below is the distribution of I/O sizes, and the overall weighted average of the distribution utilized for testing both the hot and cold zones during the random I/O test cases, used with vdbench.

<table>
<thead>
<tr>
<th>IO Size</th>
<th>% of Workload</th>
<th>Weighted Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>512</td>
<td>5</td>
<td>0.026</td>
</tr>
<tr>
<td>1KB</td>
<td>3</td>
<td>0.03</td>
</tr>
<tr>
<td>4KB</td>
<td>4</td>
<td>0.16</td>
</tr>
<tr>
<td>8KB</td>
<td>10</td>
<td>0.8</td>
</tr>
<tr>
<td>16KB</td>
<td>30</td>
<td>4.8</td>
</tr>
<tr>
<td>32KB</td>
<td>45</td>
<td>14.4</td>
</tr>
<tr>
<td>64KB</td>
<td>3</td>
<td>1.92</td>
</tr>
<tr>
<td>Average</td>
<td>100%</td>
<td>22.1</td>
</tr>
</tbody>
</table>

Test 1 Workload - Effective SSD Size
For this test, find the effective size of the SSD tier or cache, the iometer tools was utilized. Multiple settings were used in order to fill the cache, or SSD tier. This was done as follows:

1. Initialize the system with random write at 256 KB, random data sets (non compressible or deduplicatable)
2. Write workload (99% write, 100% random) in order to observe the read cache hit percentage
3. Once competitor system reaches plateau then the system has finished caching data
4. Change to 8 KB, 40% write percentage to finalize stabilizing of cache hit rate
Test 2 Workload - I/O Rate at Response Time

Test the performance of a frequently accessed portion of data.

- The primary workload tool utilized was vdbench
- A sample configuration file is provided on the following page
  1. The “hot” band workload was set to receive 90% of the I/O, by setting “skew=90”
  2. The “cold” band workload was set to receive 10% of I/O by setting skew=10
  3. Configured to generate 2:1 compressible data
  4. All other parameters were remained as detailed

Test 3 Workload - Performance of Frequently Accessed Data

Test the performance of a frequently accessed portion of data.

- The primary workload tool utilized was vdbench
- A sample configuration file is provided on the following page
- The differences between sample vdbench configuration and this test are as follows:
  5. The “hot” band workload was set to receive 100% of the I/O, by setting “skew=100”
  6. The “cold” band workload was set to receive 0% of I/O
  7. Configured to generate 2:1 compressible data
  8. All other parameters were remained as detailed

Test 4 Workload - Performance of Infrequently Accessed Data

Test the performance of in-frequently accessed portion of data.

- The primary workload tool utilized was vdbench with sample configuration on the next page
- The differences between sample vdbench configuration and this test are as follows:
  1. The “cold” band workload was set to receive 100% of the I/O, by setting “skew=100”

Test 5 Workload - Performance of Frequently Accessed Data - No Optimization Time

Test the performance of a volume after running test #3.

- The primary workload tool utilized was vdbench with sample configuration on the next page
- There were no differences between the settings use for this test, and test #2.

Test 6 Workload - Large Block Sequential Performance

Test the performance large block, sequential workload

- The primary workload tool utilized was iometer, using 256 KB blocks
Example vdbench parameter file

The specific settings for each workload changed somewhat, primarily the “skew” factor denoted below. This parameter directs a percentage of total I/O to a specific workload. Thus, both the “hot” and “cold” workload skew parameters were changes, such that their sum was 100%.

The transfer block-size (xfersize) distribution remained unchanged, as did the read-write percentage. The random percentage (seekpct) was set at 95% of the hot band and 90% for the cold band. When only a single workload was run, the higher, 95% seek percentage was utilized.

In some instances, a specific I/O rate was specified, rather than “curve” or “max”. When using a specified I/O rate, the rate was determined based upon initially running a maximum I/O rate workload.

Additionally, a compression setting of 2:1 was used during a pre-write phase, to create 2 to 1 compressible data used to pre-fill the 6 TB volume during some tests as indicated. The testing then used this setting to ensure overwrites also created 2:1 compressible data.

---

* VDbench parameter file

* hd=default,jvms=2

* Note: Compression setting used during some tests
* compratio=2

sd=default,openflags=directio,threads=128
sd=sd6,lun=\..\PhysicalDrive5

* Workload definition values

* Hot Workload 23.5 KB (weighted avg.) with 95% random, 60% read
  wd=hot,sd=sd*,xfersize=(512,5,1k,3,4k,4,8k,10,16k,30,32k,45,64k,3),seekpct=95,rdpct=60,range=(0,24),skew=90

* Cold Workload 23.5 KB (weighted avg.) with 90% random, 60% read
  wd=cold,sd=sd*,xfersize=(512,5,1k,3,4k,4,8k,10,16k,30,32k,45,64k,3),seekpct=90,rdpct=60,range=(25,98),skew=10

* Run a curve with 60 second warmup, report every 1 second
  rd=rd1,wd=*,iorate=curve,elapsed=300,warmup=60,interval=1

---

Table 2: vdbench sample configuration file
Appendix D - Configuration and Price Comparison

Comparison of List Price of Configured System

The detailed listing of each system’s configuration was provided previously in Appendix A. Provided below are details of pricing for each system, and the sources used to determine those prices.

Competitor’s System

- GSA Price = $94,710
- Dual, active / passive controller, 3U configuration, Single shelf w/ controllers
- 4 x 600 GB SSD (Read Cache) + 12 x 2 TB 7.2K NL-SAS (Backig Store)

Dell PS 6210XS System

- GSA Price = $77,133
- Dual, active / passive controller, 2U configuration, Single shelf w/ controllers
- 7 x 800 GB SSD + 17 x NL-SAS

Dell SC 4020 System

- GSA Price = $85,908 (12 SSD’s only - tested configuration)
- Dual, active 6035, 512U
- 12 x 1.6 TB SSD
- [ Note: Typically systems include 12 x 1.6 TB SSD + 12 x 3 TB NL-SAS ] = GSA Price $106,950

As shown above, the tested systems all had similar configurations, consisting of dual controllers and a single tray of storage media with a somewhat similar usable capacity. Additionally, the configured prices were similar and these systems would be quoted against each other.
# Appendix E - Test Result Details

## Test 3 Data Details

Data for Figure 5.

<table>
<thead>
<tr>
<th></th>
<th>100% IO to 1.5TB cold</th>
<th>100% IO to 4.5TB cold</th>
<th>100% IO to 6TB</th>
<th>90% IO to 1.5TB hot</th>
<th>100% IO to 1.5TB hot</th>
<th>100% IO to 1 TB hot</th>
</tr>
</thead>
<tbody>
<tr>
<td>PS6210XS</td>
<td>0.3</td>
<td>0.4</td>
<td>0.6</td>
<td>3.0</td>
<td>3.0</td>
<td>3.0</td>
</tr>
<tr>
<td>Competitor</td>
<td>0.1</td>
<td>0.1</td>
<td>0.2</td>
<td>0.7</td>
<td>1.6</td>
<td>3.1</td>
</tr>
<tr>
<td>Cache Hit</td>
<td>8%</td>
<td>8%</td>
<td>26%</td>
<td>71%</td>
<td>88%</td>
<td>97%</td>
</tr>
</tbody>
</table>

Data for Figure 4

<table>
<thead>
<tr>
<th></th>
<th>100% IO to 1.5TB cold</th>
<th>100% IO to 4.5TB cold</th>
<th>100% IO to 6TB</th>
<th>90% IO to 1.5TB hot</th>
<th>100% IO to 1.5TB hot</th>
<th>100% IO to 1 TB hot</th>
</tr>
</thead>
<tbody>
<tr>
<td>SC4020</td>
<td>8.3</td>
<td>8.3</td>
<td>8.3</td>
<td>10.0</td>
<td>9.8</td>
<td>9.8</td>
</tr>
<tr>
<td>Competitor</td>
<td>0.1</td>
<td>0.1</td>
<td>0.2</td>
<td>0.7</td>
<td>1.6</td>
<td>3.1</td>
</tr>
<tr>
<td>Cache Hit</td>
<td>8%</td>
<td>8%</td>
<td>26%</td>
<td>71%</td>
<td>88%</td>
<td>97%</td>
</tr>
</tbody>
</table>

Not shown in figure: 100% of IO to 0.75TB hot: SC4020: 9.8, Competitor, 3.8, Cache hit rate 99%
About Evaluator Group
Evaluator Group Inc. is dedicated to helping IT professionals and vendors create and implement strategies that make the most of the value of their storage and digital information. Evaluator Group services deliver in-depth, unbiased analysis on storage architectures, infrastructures and management for IT professionals. Since 1997 Evaluator Group has provided services for thousands of end users and vendor professionals through product and market evaluations, competitive analysis and education. www.evaluatorgroup.com Follow us on Twitter @evaluator_group

Copyright 2015 Evaluator Group, Inc. All rights reserved.
No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying and recording, or stored in a database or retrieval system for any purpose without the express written consent of Evaluator Group Inc. The information contained in this document is subject to change without notice. Evaluator Group assumes no responsibility for errors or omissions. Evaluator Group makes no expressed or implied warranties in this document relating to the use or operation of the products described herein. In no event shall Evaluator Group be liable for any indirect, special, consequential or incidental damages arising out of or associated with any aspect of this publication, even if advised of the possibility of such damages. The Evaluator Series is a trademark of Evaluator Group, Inc. All other trademarks are the property of their respective companies.

This document was developed with Dell funding. Although the document may utilize publicly available material from various vendors, including Dell and others, it does not necessarily reflect the positions of such vendors on the issues addressed in this document.