Lab Evaluation of NetApp® Hybrid Array with Flash Pool™ Technology

Evaluation report prepared under contract with NetApp

Introduction

As flash storage options proliferate and become accepted in the enterprise, IT professionals must analyze various types of flash storage solutions in order to determine the best fit for many of their applications. Continuous access to data is a key factor to consider when selecting a storage system for important applications.

NetApp addresses these challenges with their Hybrid Array – Flash Pool technology along with clustered Data ONTAP®. Flash Pool technology accelerates storage performance for volumes and resources for host applications and clients.

NetApp commissioned Demartek to evaluate the effectiveness of Flash Pool with different types and numbers of hard disk drives using an online transaction processing (OLTP) database workload, and to evaluate the performance of Flash Pool in a clustered Data ONTAP environment during a cluster storage node failover scenario. The configurations tested included SAS HDD + SSD for performance and SATA HDD + SSD for balance of performance and capacity – storage efficiency.

Executive Summary and Key Findings

The NetApp Flash Pool technology provides accelerated and more consistent storage performance than the performance available with hard disk drive only storage volumes.

- When compared to SAS HDD configurations, the addition of the Flash Pool (SAS+SSD) resulted in 283% gain in IOPS and 46% reduction in $/IOPS.
- Flash Pool (SATA+SSD) provides better $/IOPS and $/GB compared to the SAS HDD configuration with a 50% reduction in the number of spinning drives.
- In the long-running OLTP workload scenario, Flash Pool was able to reduce latency by up to a factor of 66x.
- In the cluster failover scenario, the Flash Pool storage performance in degraded mode was able to match the performance of the normal operation in terms of IOPS and bandwidth, and was only slightly higher in terms of latency.

Architecture

NetApp Flash Pool is a storage cache option within the NetApp Virtual Storage Tier product family, available for NetApp FAS storage systems. A Flash Pool configures solid state drives (SSDs) and hard disk drives (HDDs) into a single storage pool, known as an "aggregate" in NetApp parlance, with the SSDs providing a fast response time cache for volumes that are provisioned on the Flash Pool aggregate.

Other NetApp flash-based caching solutions include Flash Cache™, which consists of one or more flash PCIe cards installed in a storage controller acting as a system-wide cache, and the E-Series SSD cache.

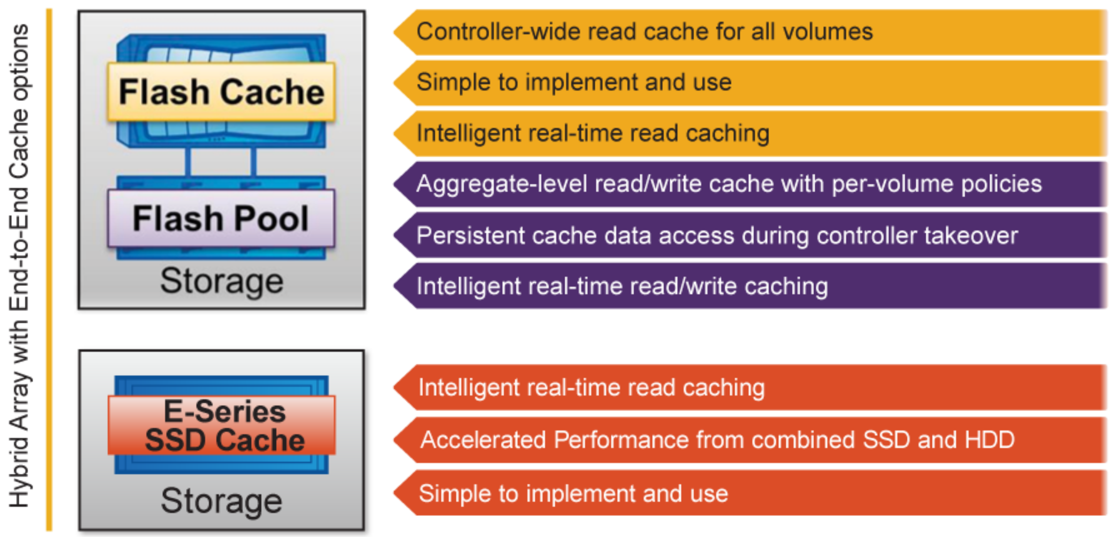


Figure 1 – NetApp Hybrid Array Options

Flash Pool Caching

Flash Pool is a hybrid-caching model wherein SSDs are combined with one type of HDD (performance or capacity drives) to form a Flash Pool Aggregate (an aggregate is a collection of RAID groups that stores one or more application data).

Many workloads such as OLTP, databases, file services etc. are characterized by a subset of the dataset that are accessed repeatedly – common known as the hot spot. The SSDs within the Flash Pool aggregate are used to dynamically cache hot data.

From the perspective of an HDD, the most "expensive" activities are random I/Os. The role of the SSDs within a Flash Pool is to offload random read and repetitive random write operations from the HDDs; thereby improving latency and increasing IOPS for workloads that are disk I/O bound. Because these random operations are handled by the SSDs, the

HDDs can better handle other operations. Flash Pool automatically manages the data insertion and eviction of the cache based on the hot or cold characteristics of the dataset.

In addition the cached data on the SSD is fully RAID protected and High Availability compatible. This assures that Flash Pool provides consistent performance over planned or un-planned failover scenarios. Flash Pool is fully integrated with Data ONTAP capabilities such as compression, deduplication, etc.

In a NetApp Flash Pool aggregate, the SSDs are used to cache data for all volumes that are provisioned on that aggregate. Flash Pool is specifically targeted at accelerating repeat random read operations and off-loading small-block random overwrite operations (which are a specific class of writes) from HDDs. Although a Flash Pool SSD cache is a single entity within the aggregate, for the purposes of configuration and cache data management the read cache and write cache are two separate entities.

Flash Pool Read Caching

Flash Pool caches random read requests of all sizes. Caching of random reads significantly improves read I/O response times for the volumes provisioned in a Flash Pool that have read caching enabled. It is possible to exclude volumes from using the Flash Pool read cache.

Flash Pool Write Caching

Flash Pool write caching is targeted at caching overwrites of random data where the operation size is 16KB or smaller. Caching small-block random overwrites off-loads write operations that can consume many HDD I/O cycles on data that will soon be invalidated by the next overwrite. In addition, read-after-writes are also served from the SSDs.

In Database OLTP environments, there can be a high frequency of random (repetitive) writes that have a relatively small I/O size. Unfortunately for database applications, this high rate of random writes can get bogged down by HDDs. By absorbing these writes on the SSD, both the writes and read after writes (read associated to writes on the SSD) get served from the SSD.

The Flash Pool does not accelerate normal writes, as the Data ONTAP® operating system is already write-optimized through the use of write cache and non-volatile memory (NVRAM or NVMEM).

Clustered Data ONTAP

NetApp storage systems can be configured as clustered Data ONTAP, which support two or more nodes in a single storage cluster. The cluster forms a shared pool of storage resources that are available to applications, SAN hosts and NAS clients. The shared pool appears as a single system image for management purposes, providing a single common point of management, through GUI or CLI tools, for the entire cluster.

This clustered storage architecture provides these basic benefits:

- Fault tolerance in the event of a disruption on one node, maintaining access to data for the host applications and clients
- Scale-out growth, allowing seamless addition of controllers to the storage cluster
- Non-disruptive operations such as software and firmware upgrades

Testing Flash Pool Effectiveness

To determine the effectiveness of NetApp Flash Pool, two test scenarios were performed. Because Flash Pool operates as an SSD cache, we would expect to see significant improvement in latency for workloads that access data storage on the back-end hard disk drives (HDD). For a storage system configured as a cluster, we would also expect to see improvements in node failure conditions such as failover, failback and operation during degraded mode.

The first test scenario compares the effectiveness of Flash Pool with different types and numbers of HDD using an online transaction processing (OLTP) workload. The goal of this test is to show the improvements in latency while maintaining a steady I/O rate in terms of IOPS and bandwidth.

The second test scenario observes the performance during a storage cluster failover event. We found that the data in the cache was persistent and remained "warm" during the failover process, as the performance data shows.

The same type of application workload was run for both test scenarios, but with different parameters. These workloads consisted of a database OLTP workload with repeated reads and some overwrites. A series of larger (higher IOPS) workloads were run for the long-running OLTP workloads with various HDD types to show the full benefit of Flash Pool at higher workload levels. A smaller (lower IOPS) workload was run for the failover/failback tests that was enough to place a reasonable load on the storage system and then conduct the failover/failback tests.

This particular OLTP workload has a working set that moves across the entire database over time. It is a cache-friendly workload, in that there are repeated accesses to the working set as it makes its way through the entire data set.

These workloads were initialized before the test measurements were taken. For the storage cluster HA failover/failback test, the workload was running in a steady state when the storage cluster failover steps began.

Long-Running OLTP Workload Results

The purpose of the long-running OLTP workloads (12 hour runs) was to show the effects of the Flash Pool acceleration with different HDD configurations, including SAS HDDs, a small number of SATA HDDs and a larger number of SATA HDDs. Specifically, the configurations included:

- **Baseline** – Large configuration of 450GB SAS HDDs and no flash acceleration
- **Test 1** – Large configuration of 450GB SAS HDDs and 200GB SSDs (“SAS Perf.”)
- **Test 2** – Large configuration of 1TB SATA HDDs and 200GB SSDs (“SATA Perf.”)
- **Test 3** – Small configuration of 1TB SATA HDDs and 200GB SSDs (“SATA Efficiency”)
- **Note**: the goal of these tests is not to show the maximum IOPS and throughput that the system can attain, but to provide a steady workload so that the effects of the Flash Pool cache can be seen.

Use Case	Goals	Configuration
SAS - Baseline	This is the baseline for comparison across different test cases	8 shelves 450GB SAS 10K rpm
SAS+SSD Perf Acceleration	Adding SSDs to the baseline for higher IOPS and better $/IOPS as compared to the baseline	8 shelves 450GB SAS, 2 shelves 200GB SSD
SATA+SSD Perf Acceleration	From SAS to SATA+SSD for higher IOPS combined with better $/IOPS and $/GB as compared to the baseline	12 shelves 1TB SATA, 2 shelves 200GB SSD
SATA+SSD Storage Efficiency	Balance of improved $/IOPS and $/GB as compared to the baseline.	4 shelves 1TB SATA, 1 shelf 200GB SSD
SATA+SSD Storage Efficiency #2	The 3TB SATA drives have similar performance as the 1TB SATA drives– with higher capacity resulting in lower $/GB.	4 shelves 3TB SATA, 1 shelf 200GB SSD

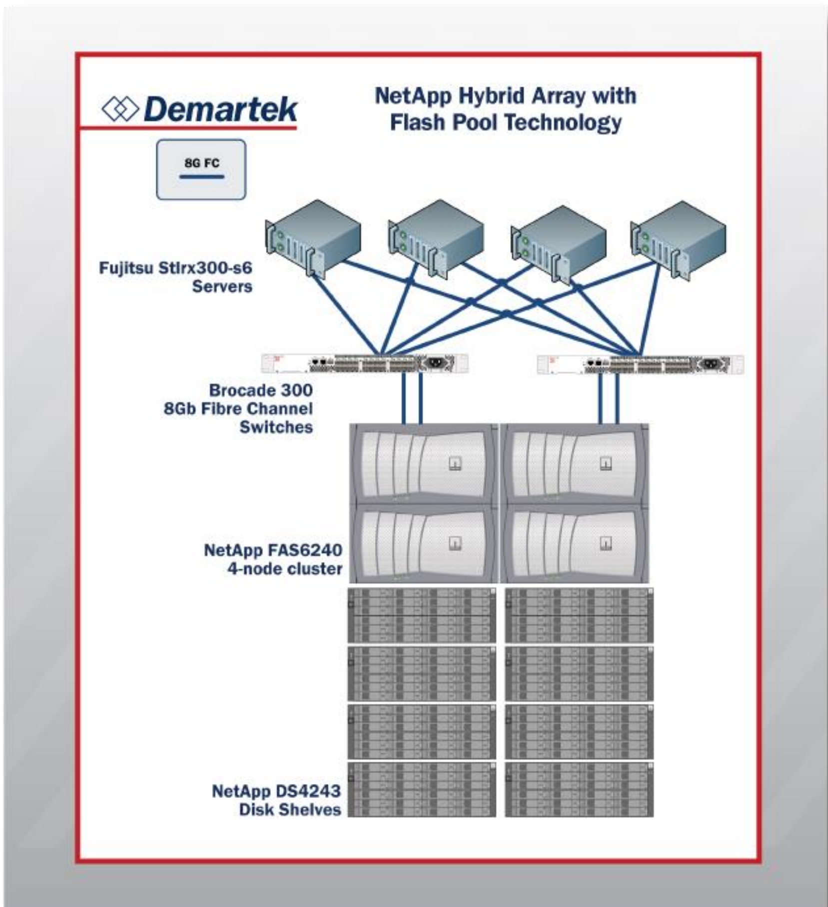
The large configurations of SAS and SATA HDDs used a total of eight disk shelves that included 192 HDDs and 48 SSDs. 160 of the 192 HDDs were used for data and 44 of the 48 SSDs were used for the Flash Pool cache for the large configuration tests. The small configuration was half the size of the large configuration, with four disk shelves and half the number of drives available and used.

As we stated previously, the NetApp Flash Pool aggregate uses SSDs to cache data for all volumes that are provisioned on that aggregate. Flash Pool is specifically targeted at accelerating repeat random read operations and off-loading small-block random overwrite operations. Given this type of caching approach, we observed two behaviors with Flash Pool:

- The SSD cache warms-up over time and latencies decrease during the warm-up period. Once fully warmed, overall latency is significantly reduced compared to an equivalent HDD-only configuration.
- Once the SSD cache is fully warmed, I/O rates stabilize.

The three HDD configurations were able to sustain different I/O rates over the long term. The particular I/O rates for each configuration are somewhat interesting, but it is the latency that is the most interesting aspect of these tests. Flash Pool dramatically reduced latency, fairly quickly given the length of the test runs.

The database consisted of approximately 22TB of data and 3.2TB of logs. The databases and logs were spread evenly over the storage volumes. Each node of the NetApp FAS6240 cluster had its own root volume for ONTAP and twelve data volumes with one LUN per data volume. The application workloads performed their I/O with an average block size of approximately 8KB and the overall read/write ratio was approximately 39% reads and 61% writes.

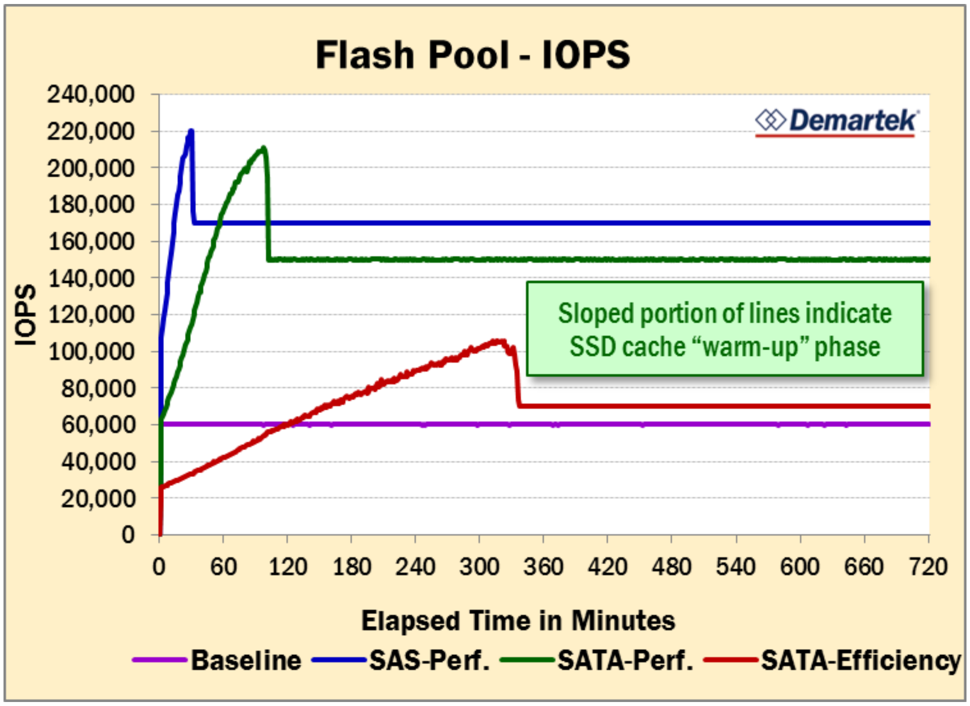


The workloads were designed to achieve a steady state performance in terms of IOPS and bandwidth, with the parameters adjust for each HDD type.

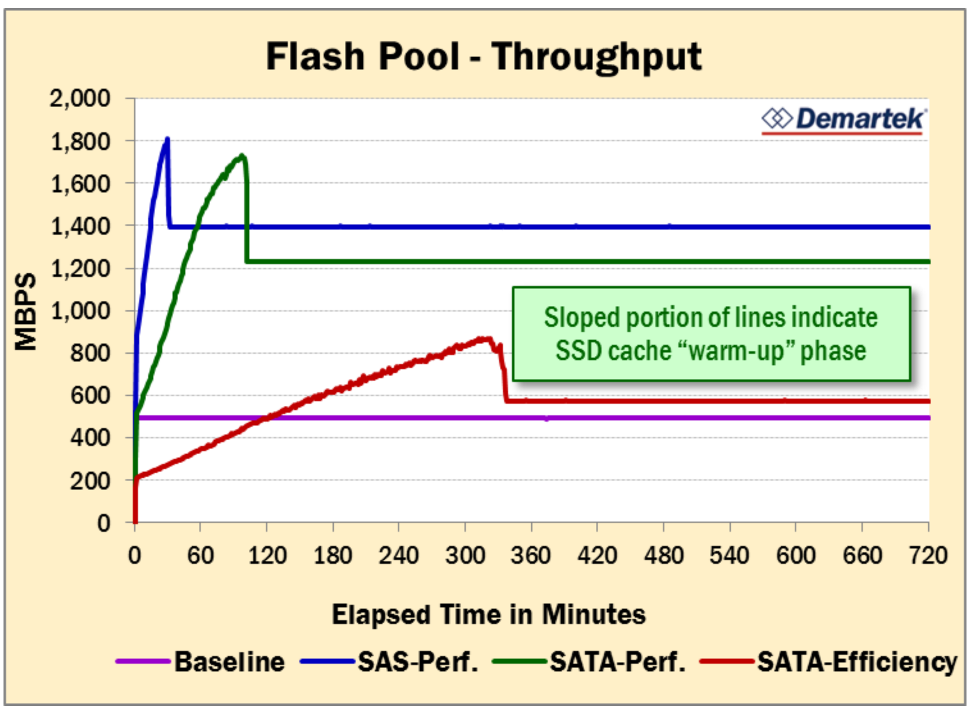
Performance	IOPS		Throughput MB	
	IOPS	Compare to Baseline	Per Node	Compare to Baseline
SAS - Baseline	60,000		123	
SAS + SSD Perf Acceleration	170,000	283%	360	293%
SATA + SSD Perf Acceleration	150,000	250%	300	244%
SATA +SSD Storage Efficiency	70,000	117%	150	122%
SATA + SSD Storage Efficiency #2	70,000	117%	150	122%

Capacity and Cost	Raw Capacity		$/IOPS	$/GB
	GB	Compare to Baseline	Compare to Baseline	Compare to Baseline
SAS - Baseline	86,400			
SAS + SSD Perf Acceleration	86,400		-46%	
SATA + SSD Perf Acceleration	288,000	333%	-33%	-49%
SATA +SSD Storage Efficiency	96,000	111%	-18%	-14%
SATA + SSD Storage Efficiency #2	288,000	333%	-1%	-65%

OLTP Results – IOPS

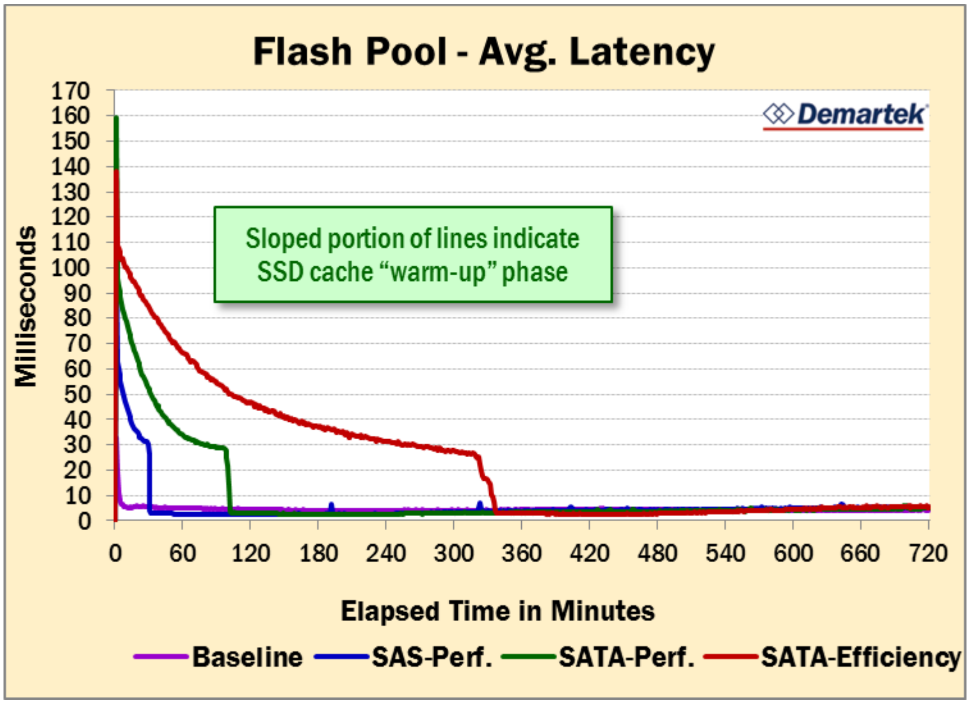


OLTP Results – MBPS



OLTP Results – Latency

Notice that the latencies decrease dramatically for each of the tests, regardless of the specific IOPS or MBPS rates shown above.



The average latency data was collected over one minute intervals. The initial average latencies were quite high, as shown by the graph above. However, because the volumes were created with Flash Pool SSD caching, the latencies immediately began to drop and continued to drop until the cache was fully warmed.

The table below shows run time and latency statistics.

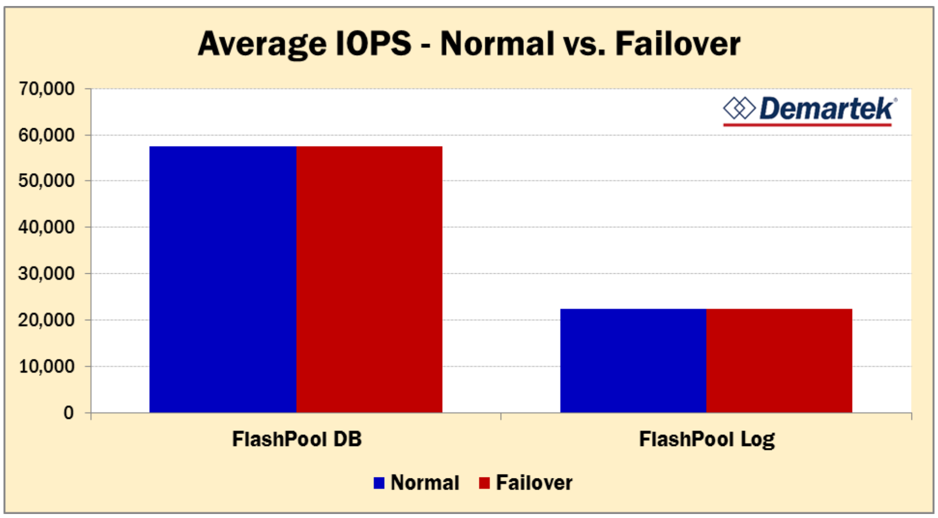
	Total run time (minutes)	Run time at which latency first dropped to approx. 3 ms (minutes)	Avg. Latency after cache warm-up (Milliseconds)
SAS-Perf.	720	32	3.775
SATA-Efficiency	720	337	3.799
SATA-Perf.	720	102	3.536

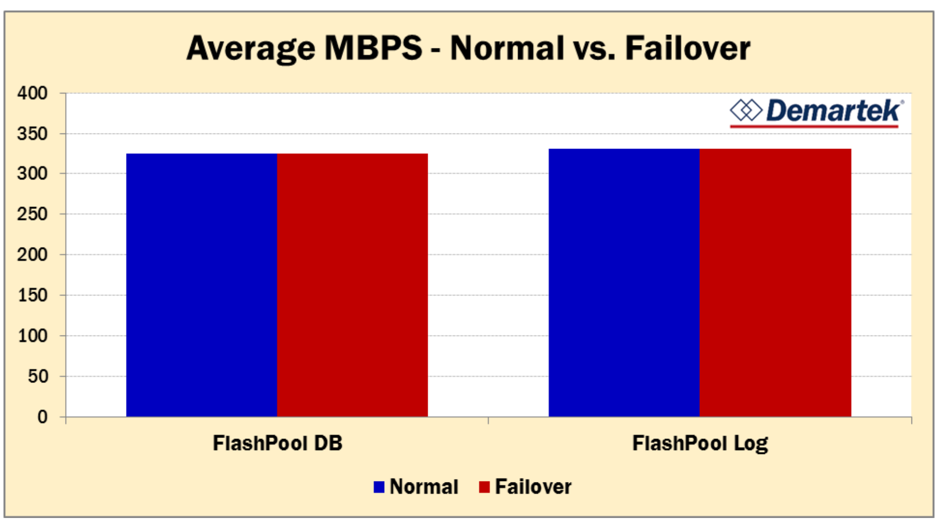
Storage Cluster HA Failover/Failback Results

For the cluster failover/failback tests, we used a two-node cluster running a slightly lower-performing workload than for the previous tests. This was due to system availability at the time of running the tests. However, for the purpose of demonstrating the cluster failover and related activities, we needed a reasonable but not necessarily identical workload.

Storage Cluster HA Failover/Failback Test Results

To evaluate the effectiveness of the Flash Pool solution in the clustered Data ONTAP configuration, we compared the performance in the normal mode (both nodes running normally) and in a failover mode ("degraded" mode). We found that for IOPS and MBPS, the results were nearly identical.





Summary and Conclusion

Flash Pool provides significant performance advantages and consistent latency for database applications, much more so than storage volumes consisting only of hard disk drives.

- In the long-running OLTP workload scenario, Flash Pool was able to reduce latency by up to a factor of 66x.
- In the cluster failover scenario, the Flash Pool storage performance in degraded mode was able to match the performance of the normal operation in terms of IOPS and bandwidth, and was only slightly higher in terms of latency.
- When compared to SAS HDD configurations, the addition of the Flash Pool (SAS+SSD) resulted in 283% gain in IOPS and 46% reduction in $/IOPS.
- Flash Pool (SATA+SSD) provides better $/IOPS and $/GB compared to the SAS HDD configuration with a 50% reduction in the number of spinning drives.
- Adding Flash Pool Hybrid Array will provide higher performance at lower cost compared to the HDD only solution. There is flexibility in either performance only (SAS+SSD) or balance of performance and capacity (SATA+SSD).

The most current version of this report is available at http://www.demartek.com/Demartek_NetApp_Hybrid_Array_with_Flash_Pool_Technology_Evaluation_2014-05.html on the Demartek website.