Technical FAQ

NetApp Flash Pool

Data ONTAP 8.1.4P1, 8.2.1 and Later Releases

Skip Shapiro, NetApp
Version 4.7.1
June 2015

Abstract

This technical FAQ addresses the common questions about NetApp® Flash Pool™ technology with an emphasis on information that complements the content in TR-4070: Flash Pool Design and Implementation Guide. The questions addressed cover a range of Flash Pool topics including hardware and software requirements, architecture, configuration, performance and interoperability for both clustered Data ONTAP® and Data ONTAP operating in 7-Mode deployments.

Data Classification

NetApp Confidential—Limited Use

Version History

Version	Date	Document Version History
4.7.1	July 2015	MetroCluster with FibreBridge stack depths added to table 8
4.7	June 2015	Data ONTAP 8.3.1 updates
4.6.5	May 2015	Cumulative minor clarifications and corrections in 4.6.1 through 4.6.5. Minimum supported releases updated in accordance with SSD support.
4.6	February 2015	Data ONTAP® 8.3 updates and reorganization for revised template.
4.5	May 2014	FAS2500 and FAS8080 EX data added. Data ONTAP 8.2.2 table added. Several new questions added and few other answers were revised.
4.4	February 2014	FAS8020, 8040 and 8060 data added. Some tables replaced with links to the Storage Subsystem Tech FAQ.
4.3	January 2014	400GB and 1.6TB SSD data added; SnapMirror® interoperability updated and SnapVault® interoperability added to the topic.
4.2	November 2013	Data ONTAP 8.2.1 updates for AWA and FAS2200 Flash Pool™ cache maximums, plus other revisions and additions for enhanced clarity.
4.1	June 2013	FAS2200 mixed internal drive configuration and format/branding updates.
4.0	May 2013	Updated based Data ONTAP 8.2 enhancements and additional SSD support. Replaced content that appears in TR-4070 with references.
3.2	December 2012	Corrected SyncMirror® information concerning the mirrored plex counting toward the cache limit.
3.1	December 2012	Updated FAS2200 series SSD limits.
3.0	December 2012	Major update. Added 200GB SSD drive, DS2246 SSD offerings, updated cache limits for 8.1.1 and 8.1.2, proof of concept recommendations, MetroCluster™ updates, performance graphs for OLTP and SFS, updated shelf mixing information, fixed some typos, eviction scanner information, pure SSD support information, and additional command information (cache used, disable cache, and more). Corrected the X441A-R5 usable capacity.
2.1	August 2012	Fixed references link to the Storage Subsystem Technical FAQ.
2.0	August 2012	Restructured the content and added several new items based on a total product review (hardware and software).
1.1	July 2012	Added an item to address combined Flash Pool and Flash Cache™ limits.
1.0	May 2012	Paul Update: Initial document created and published.

TABLE OF CONTENTS

Version History 1
TABLE OF CONTENTS 3
List of Tables 6
1 Overview 7
What is NetApp Flash Pool technology? 7
What is a Flash Pool aggregate? 7
How does Flash Pool caching work? 7
What are the benefits of using Flash Pool caching? 7
What application workloads benefit from using Flash Pool cache? 7
What are the new capabilities are available for Flash Pool aggregates with Data ONTAP 8.3? 8
What are the new capabilities are available for Flash Pool aggregates with Data ONTAP 8.3.1? 8
2 Requirements 8
What are the system-level software and hardware requirements for using Flash Pool caching? 8
Are DS14mk2 AT and DS14mk2 FC shelves supported in a Flash Pool aggregate? 9
What are the aggregate-level requirements for using Flash Pool caching? 9
3 Architecture 10
How do read caching and write caching work in Flash Pool aggregate? 10
Do all volumes provisioned on a Flash Pool aggregate use the cache? 10
Is Flash Pool cache divided into separate segments for read-inserted and write-inserted blocks? 11
Does using a Flash Pool aggregate improve write latency or write throughput performance? 11
Does a Flash Pool aggregate increase the maximum throughput capability of a storage controller? 11
Can blocks, LUNs, files or entire volumes be pinned in Flash Pool cache? 11
What is the eviction scanner? 11
When and how often does the eviction scanner run? 11
How full can a Flash Pool cache get? 12
How is the retention of blocks in a Flash Pool cache managed? 12
Is there a performance impact to the controller when the eviction scanner is running? 12
Is the capacity of the SSDs in a Flash Pool aggregate counted towards the aggregate capacity? 12
Do the SSDs in Flash Pool aggregates count toward the maximum drive count for a system? 12
Do the SSDs used in Flash Pool aggregates affect the maximum storage capacity of a system? 12
Is 10% of Flash Pool cache capacity allocated for WAFL reserve? 12
Is Flash Pool SSD cache capacity decremented by aggregate-level Snapshot® reserve? 12
Does Data ONTAP use the TRIM command when managing the SSDs? 13
How does wear leveling work with SSDs in a Flash Pool aggregate? 13
How is media usage monitored for SSDs used in a Flash Pool aggregate? 13
4 Configuration 14
How is a Flash Pool aggregate created? 14
Are the SSDs and HDDs in a Flash Pool aggregate configured in the same RAID group? 14
Can more than one type of HDD be configured in a Flash Pool aggregate? 14
Can SSDs, or HDDs, with different capacities be used in the same Flash Pool aggregate? 14

Can an HDD-only aggregate be converted into a Flash Pool aggregate? ... 15
Can a Flash Pool aggregate be converted into an HDD-only Flash Pool aggregate? ... 15
Can an all-SSD aggregate be converted into a Flash Pool aggregate by adding HDDs? ... 15
Can SSDs be added to a Flash Pool aggregate to increase the cache size? ... 15
Can SSDs be removed from a Flash Pool aggregate to decrease the cache size? ... 15
Can a root aggregate that is provisioned with dedicated (non-partitioned) HDDs be a Flash Pool aggregate? ... 15
Can a root aggregate that is provisioned using advanced data partitioning (ADP) be a Flash Pool aggregate? ... 15
Can Flash Pool cache be provisioned using shared (partitioned) SSDs? ... 15
What is a storage pool? ... 16
How is Flash Pool cache provisioned using a storage pool of shared (partitioned) SSDs? ... 16
Can SSDs be added to a storage pool to increase its size? ... 16
Can SSDs be removed from a storage pool to decrease its size? ... 16
What are the benefits of using a storage pool to provision Flash Pool cache? ... 16
When should Flash Pool cache be provisioned from a storage pool versus using dedicated (non-partitioned) SSDs? ... 16
What is the recommended minimum number of data SSDs per Flash Pool aggregate when dedicated (non-partitioned) SSDs are used? ... 17
What is the recommended minimum number of data SSDs in a storage pool? ... 17
What is the recommended minimum number of dedicated (non-partitioned) SSDs to add to of a Flash Pool aggregate when expanding the cache? ... 18
What is the recommended minimum number of SSDs to add when expanding a storage pool? ... 19
How many Flash Pool aggregates are supported on a single system? ... 19
How many storage pools are supported on a single system? ... 19
How is Flash Pool cache size calculated? ... 19
What is the maximum number of dedicated (non-partitioned) SSDs per Flash Pool aggregate and per system (HA pair or single controller) when used in Flash Pool aggregates? ... 20
What type of RAID protection is recommended for a SSD RAID group in a Flash Pool aggregate on a system that is running Data ONTAP 8.1.1 or a later 8.1.x release? ... 21
What type of RAID protection is recommended for a SSD RAID group in a Flash Pool aggregate on a system that is running a Data ONTAP 8.2.x release? ... 22
What type of RAID protection is recommended for a RAID group of dedicated (non-partitioned) SSDs in a Flash Pool aggregate on a system that is running Data ONTAP 8.3 or a later release? ... 22
What is the best practice for RAID group size when using dedicated (non-partitioned) SSDs for Flash Pool cache? ... 22
Can a RAID-DP group of dedicated (non-partitioned) SSDs in a Flash Pool aggregate be converted to RAID-4 while the aggregate is online? ... 22
Can a RAID 4 group of dedicated (non-partitioned) SSDs in a Flash Pool aggregate be converted to RAID-DP while the aggregate is online? ... 22
What are the RAID recommendations for storage pool allocation units used for Flash Pool cache? ... 23
Can the RAID type used for a storage pool allocation unit be changed after the allocation unit is assigned to a Flash Pool aggregate? ... 23
What is the hot spare recommendation for SSDs in a Flash Pool aggregate? ... 23
What is the performance impact when a Flash Pool SSD RAID group is in degraded mode? ... 23
What is the system impact when a SSD RAID group in a Flash Pool aggregate fails? ... 23
What is the maximum Flash Pool cache size supported on a single controller or HA pair system? ... 23
Can Flash Cache cards and Flash Pool aggregates be configured on the same system? ... 24
Does Flash Cache reduce the amount of Flash Pool cache capacity allowed on a system? ... 24
What methods are available to size cache for a Flash Pool aggregate? ... 24

What is Automated Workload Analyzer (AWA)? ... 25
How long should AWA run to get a good representation of an aggregate's workload? ... 25
Can AWA run on a Flash Pool aggregate? ... 25
Can AWA run on a node that has Flash Cache enabled? ... 25
What does the default AWA print command output look like? ... 25
Does AWA provide volume-level granularity for workload analysis? ... 27
How much CPU and memory are used when AWA is running? ... 28
Can AWA be run on more than one aggregate at a time? ... 28
Can AWA be used to size cache for Flash Cache? ... 28
How can Flash Pool cache size be estimated when a FAS system is replacing a system that doesn't run Data ONTAP? ... 28
Can Predictive Cache Statistics (PCS) be used to estimate how much Flash Pool cache would be beneficial to add to a system? ... 28
What happens if the maximum supported cache capacity is exceeded on a system? ... 29
How many drives can be configured in a single SAS stack when some of the drives are SSDs? ... 29
Can different shelf types and drive types be configured in the same storage stack? ... 30
Can a shelf configured with both SSDs and HDDs be used to provision a Flash Pool aggregate? ... 30
Can a Flash Pool aggregate be configured on only one controller in an HA pair? ... 30
Are FAS2200 series configurations that have a mix of internal SSDs and HDDs supported? ... 30
Are FAS2500 series configurations that have a mix of internal SSDs and HDDs supported? ... 30
Is Flash Pool supported on FAS FlexArray and V-Series system configurations? ... 30
What are the default Flash Pool caching policies? ... 30
Can the caching policies of a volume be changed from the default Flash Pool caching policies? ... 30
What Flash Pool caching policy options are available? ... 31
What are the use cases for using a read caching policy other than "random_read"? ... 32
What is the use cases for using the write caching policy "none"? ... 33
Can the caching of blocks from a volume in a Flash Pool aggregate be disabled? ... 33

5 Performance ... 33

How does implementing Flash Pool aggregates affect system performance? ... 33
How soon does latency improve after a HDD-only aggregate is converted to a Flash Pool aggregate? ... 33
Does system performance degrade as the layout of data on the SSDs in Flash Pool cache ages? ... 33
Are clustered Data ONTAP quality of service (QoS) objectives used to manage the insertion and retention of blocks in Flash Pool cache? ... 33
Can System Performance Modeler (SPM) be used to size system configurations with Flash Pool? ... 33
... 34
How can Flash Pool aggregate performance be monitored? ... 34
What does the output of "stats show -p hybrid_aggr" preset command look like? ... 34
What does the output of "statistics cache flash-pool show" command look like? ... 35
How can a workload be analyzed to determine if configuring Flash Pool cache will be beneficial? ... 36
How can a workload be analyzed to determine if implementing Flash Pool would be beneficial, on a FAS system running a release earlier than Data ONTAP 8.2.1? ... 36

6 Interoperability ... 37

Can NetApp storage efficiency features be used with volumes provisioned on a Flash Pool aggregate? ... 37
How do Snapshot copies work with Flash Pool? ... 37

How does Flash Pool work with SnapMirror and SnapVault? 37
Are Flash Pool aggregates supported on NetApp Storage Encrypted (NSE) systems? 38
Can a Flash Pool aggregate be configured with SyncMirror RAID protection? 38
Are Flash Pool aggregates supported on MetroCluster systems? 38
In a 7-Mode MetroCluster, do the SSDs in both plexes of a Flash Pool aggregate count toward the maximum Flash Pool cache size for the system? 38
What is the maximum Flash Pool cache size in a 4-node clustered Data ONTAP MetroCluster? 38
Can FlexShare be used to prioritize how volumes are cached in a Flash Pool aggregate? 39
Is SnapLock software supported with Flash Pool aggregates? 39
Can the constituents of an Infinite Volume be provisioned on Flash Pool aggregates? 39

7 Flash Cache and Flash Pool Maximum Cache Sizes 40

8 Additional Resources 43

9 Contact us 43

List of Tables

Table 1) SSD part numbers and capacities, storage shelf and Data ONTAP support for Flash Pool use. 9
Table 2) Steps to create a Flash Pool aggregate 14
Table 3) Recommended minimum number of SSD data drives per Flash Pool aggregate. 17
Table 4) Recommended minimum number of SSD data drives per storage pool. 18
Table 5) Recommended minimum number of dedicated SSDs to add when expanding Flash Pool cache. 18
Table 6) Recommended minimum number of SSD to add when expanding a storage pool. 19
Table 7) Maximum number of SSD data drives in Flash Pool aggregate by maximum cache size. 20
Table 8) Maximum number of HDDs and SSDs in a mixed SAS stack. 29
Table 9) Flash Pool caching policy options 31
Table 10) Flash Pool “stats show -p hybrid_aggr” statistics descriptions. 35
Table 11) Data ONTAP 8.1.4P1 and later[1] Flash Cache, Flash Pool, and combined maximum cache sizes. 40
Table 12) Data ONTAP 8.2.2 and later[1] Flash Cache, Flash Pool, and combined maximum cache sizes. 41
Table 13) Data ONTAP 8.3 and later[1] Flash Cache, Flash Pool, and combined maximum cache sizes. 42

List of Figures

Figure 1) A sample “storage disk show –a” output 13
Figure 2) Sample default AWA print command output 26
Figure 3) Sample AWA print output with volume detail 27
Figure 4) Cache size estimation using write data rate and retention time. 28
Figure 5) SPM v1.6.1 hardware configuration detailed inputs screen shot. 34
Figure 6) Sample output of the “stats show -p hybrid_aggr” command. 34
Figure 7) Sample output of the “statistics cache flash-pool” command. 35

1 Overview

What is NetApp Flash Pool technology?

Answer: NetApp Flash Pool™ technology is least recently used (LRU), write-back caching solution, for FAS systems running the Data ONTAP® operating system, that uses solid state drives (SSDs) to cache randomly-accessed data from volumes provisioned on hard disk drives (HDDs). Data cached on the SSDs is persistently available through storage controller takeover events, and it is also RAID-protected. Flash Pool caching is best-suited for random read and random read-write workloads.

What is a Flash Pool aggregate?

Answer: A Flash Pool (or hybrid) aggregate consists of one or more RAID groups of hard disk drives (HDDs) and one or more RAID groups of solid state drives (SSDs). LUN and volume data are stored on the HDDs, and the SSDs are used to cache frequently-accessed blocks. The HDDs can be either performance disk drives (10,000 rpm or 15,000 rpm) or capacity disk drives (7200 rpm or slower). Only one type of HDD—performance or capacity—can be provisioned in an aggregate. Provisioning LUNs and volumes on a Flash Pool aggregate can improve application and storage system performance, as compared to using HDD-only systems aggregates, at a comparable or lower deployment cost.

How does Flash Pool caching work?

Answer: The default Flash Pool policies insert 4KB blocks of user data and metadata into cache that have been read randomly from HDD, and 4KB blocks that are random updates (overwrites) of logical blocks that were previously written as random updates. A block that is inserted is retained in cache for as long as it is ‘hot’, meaning it is requested more often than the least recently used blocks in the cache.

A block inserted into cache via the read caching policy is a copy of a block that is stored on HDD.

A block that is inserted into cache via the write caching policy is the only block with that valid data; there is no other block on HDD with the same data. Flash Pool cache is provisioned with one or more SSD RAID groups to provide protection against losing data inserted via the write caching policy. When a block inserted via the write caching policy becomes the least recently used block, it is written to HDD (destaged) before it is evicted from cache.

What are the benefits of using Flash Pool caching?

Answer: The primary benefit of using Flash Pool caching is better (lower) latency for random read-write workloads when compared to the latency from a HDD-only aggregate. Random read operations served from cache (read cache hits), and frequent repeat random write updates (write cache hits), offload those operations from HDDs, which enables the HDDs to handle other workloads such as sequential reads and writes, and non-cached random reads. In many cases, provisioning volumes on a new Flash Pool aggregate results in a lower-cost storage solution than if the volumes were provisioned on a non-cache assisted aggregate of performance (high rpm) HDDs, while providing equal or better performance.

What application workloads benefit from using Flash Pool cache?

Answer: Any application that consists of a high mix of random reads, or random reads and writes, will typically benefit from using Flash Pool cache. And because Flash Pool cached data is persistently available through storage controller takeovers and reboots, application performance remains more consistent during these events. The benefits make Flash Pool caching a good fit for many database applications, including transactional applications such as enterprise resource planning (ERP), customer relationship management (CRM), and data warehouses. Virtualized server and desktop workloads, and file serving applications, often benefit as well.

What are the new capabilities are available for Flash Pool aggregates with Data ONTAP 8.3?

Answer: Here's a summary list of the new Flash Pool capabilities with Data ONTAP 8.3, all of which are covered in more detailed in other sections of this FAQ:

- **Larger maximum cache sizes** – sizes are 4 times larger than with 8.2.x for most FAS models
- **Storage pool SSD partitioning** – caches for up to 4 Flash Pool aggregates can be provisioned from a shared storage pool.
- **Cluster-shell command line interface** - cluster-shell commands are available to manage Flash Pool aggregates, including commands to modify caching policies.
- **Overwrites of any size are cacheable** – previous releases only allowed overwrite transfers of 16KB or less to be inserted into cache via the write caching policy; in 8.3 there is no size limit.

  Note: All random read I/O sizes are Flash Pool cacheable starting with Data ONTAP 8.1.1.
- **Read caching from read-only volumes** – random reads from read-only volumes, such as SnapMirror® and SnapVault® destination volumes, are inserted into cache.
- **Read caching of compressed blocks** – random reads from HDDs of compression groups (compressed blocks) are inserted into Flash Pool cache.
- **Higher eviction scanner start threshold** – eviction scanner starts evicting cached blocks when the cache is 85% full; the threshold in earlier releases is 75% full.

What are the new capabilities are available for Flash Pool aggregates with Data ONTAP 8.3.1?

Answer: The enhancements to Flash Pool capabilities in Data ONTAP 8.3.1 are:

- **New caching polices** – 3 read cache insertion policies provide options that are similar to what is available with Flash Cache cards
- **AWA volume-level reporting** – summary workload and cache hit rate reporting in Automated Workload Analyzer (AWA) for volumes with the highest combination of throughput and hit rates
- **Inline compression support** – adaptive inline compression with 8KB group size is an optional storage efficiency volume policy for Flash Pool aggregates

  Faster reads from cache – Read fast path to storage enhancements enable lower latency random reads from Flash Pool SSD cache

2 Requirements

What are the system-level software and hardware requirements for using Flash Pool caching?

The general system requirements for Flash Pool caching are:

- Data ONTAP 8.1.4P1, 8.2.1 or later for clustered Data ONTAP or 7-Mode.
- Both controllers in the HA pair must be running the same version of Data ONTAP.
- Supported platforms:
  - FAS22x0: 2220, 2240-2, and 2240-4
  - FAS25xx: 2520, 2552, and 2554
  - FAS and V3xx0: 3220, 3240, 3250, and 3270; 3160 and 3170
  - FAS and V6xx0: 6210, 6220, 6240, 6250, 6280, and 6290; 6030, 6040, 6070, and 6080
  - FAS8000: 8020, 8040, 8060, and 8080 EX

Note: FAS systems using non-FAS storage shelves via FlexArray, and V-Series systems, must use only NetApp SSDs and HDDs in a Flash Pool aggregate.

- Supported storage shelves: DS4243, DS4246, DS2246, DS4486, and DS14mk4 FC

  Note: The random read performance of the DS4486 HDDs is not as good as the performance of the NL-SAS drives in DS4246 shelves.

  Note: Flash Pool aggregates are not supported using DS14mk2 FC shelves in MetroCluster™ systems.

- Supported hard disk drives (HDDs): performance (10,000 or 15,000 rpm) or capacity (7200 rpm)
- Supported solid state drives (SSDs): Table 1 below contains the list of supported SSD models, the part numbers, the right-sized usable capacities, the supported shelves and the Data ONTAP support.

Table 1) SSD part numbers and capacities, storage shelf and Data ONTAP support for Flash Pool use.

SSD Model	Right-Sized Usable Capacity	SSD Part Numbers	Supported Shelves	Data ONTAP Flash Pool Support
100GB	95,146 MiB	X441A-R5	DS4243 DS4246	8.1.4P1, 8.2.1 and later
200GB	190,532 MiB	X446A-R6 X446B-R6 X448A-R6	DS2246 DS2246 DS4246	8.1.4P1, 8.2.1 and later
400GB	381,304 MiB	X438A-R6 X575A-R6	DS2246 DS4246	8.2.1 and later
800GB	762,847 MiB	X447A-R6 X449A-R6	DS2246 DS4246	8.1.4P1, 8.2.1 and later
800GB NSE	762,847 MiB	X440A-R6 X577A-R6	DS2246 DS4246	8.2.2 and later
1.6TB	1,525,935 MiB	X439A-R6 X576A-R6	DS2246 DS4246	8.2.1 and later

Note: Some SSD models are not supported with certain configurations of Data ONTAP and FAS/V controller. Check Hardware Universe for SSD support details for a specific configuration.

Are DS14mk2 AT and DS14mk2 FC shelves supported in a Flash Pool aggregate?

Answer: No, neither DS14mk2 shelf is supported for use in a Flash Pool aggregate

What are the aggregate-level requirements for using Flash Pool caching?

The aggregate requirements for Flash Pool caching are:

- The aggregate must be a 64-bit aggregate (32-bit aggregates, traditional volumes and striped volumes are not supported).

  Note: With Data ONTAP 8.1.x and later releases, a 32-bit aggregate can be converted into a 64-bit aggregate, and the 64-bit aggregate can then be converted into a Flash Pool aggregate. However, there are situations in which a 64-bit aggregate converted from a 32-bit aggregate cannot become a Flash Pool aggregate. Refer to the *Storage Management Guide* for the version of Data ONTAP in use to confirm whether an aggregate can be converted into a Flash Pool aggregate.

- The aggregate must be an HDD aggregate (It cannot be an all-SSD aggregate).
- The aggregate cannot be in a FAILED, LIMBO, offline, or foreign state.
- Supported RAID configurations: RAID 4, RAID-DP® technology, and SyncMirror® software, including local SyncMirror® and MetroCluster™.

3 Architecture

How do read caching and write caching work in Flash Pool aggregate?

Answer: Caching of user data and metadata blocks data significantly improves latencies for random read operations. The default Flash Pool read caching policy inserts all randomly requested 4KB blocks that are read from HDDs into the cache. A block that is stored on a HDD in uncompressed format is inserted into cache before it is evicted from a read buffer in the storage controller memory. A block that is stored on HDD in a compression group is inserted into cache, as part of the entire compression group—in compressed format—simultaneously to the blocks in the compression group being inserted into read buffers in uncompressed format. There is no option to cache data from sequential read operations.

Blocks are inserted into Flash Pool cache using a Data ONTAP® consistency point (CP) write operation, which assures consistency between data in memory and data on storage. Flash Pool cache is SSD storage acting as a cache. A block or compression group inserted into Flash Pool read cache is a copy of what is already stored on HDD, and thus the cache and stored blocks are identical and consistent.

The default Flash Pool write caching policy inserts overwrite blocks into Flash Pool cache, which are updates to previously written blocks . There are three requirements for a write to be classified as an overwrite and thus be eligible for insertion into Flash Pool cache:

- The previously-written block that the incoming block is replacing must have been classified as a random write. A write block is considered to be random based the relationship of its inode to the inodes of other blocks that are written during the same consistency point.
- The incoming block must be random when compared to the blocks in its pending consistency point.
- The new write must occur before four eviction scanner passes have completed since the block was last written.

Unlike with blocks inserted into cache via the read caching policy, a write-inserted block is not a copy of a block that is stored on disk. The block in cache is the only valid version of that data. This is why Flash Pool SSD cache is configured with RAID protection.

With Data ONTAP 8.1.x and 8.2.x releases, 16KB and smaller random overwrite I/Os are inserted into cache. With Data ONTAP 8.3 and later releases, random overwrites of all sizes are inserted into cache. Inserting frequently updated blocks into Flash Pool cache, instead of writing them to HDDs, is more efficient because a SSD can handle much higher random I/O than a HDD can. In addition, blocks inserted via the write caching policy can be read from cache. This makes Flash Pool cache very effective for random workloads that include reads of a recently updated, randomly written blocks.

Flash Pool metadata tracks all cached blocks—whether they are inserted via the read or writing caching policy—to determine how frequently each 4KB block is read (in other words, how “hot” the data is). When the cache is nearly full, a block that is read less frequently than other blocks will become less important to retain in the cache, and the least frequently used blocks will eventually be evicted from the cache. This process assures that most frequently read blocks—the hottest blocks—are retained in cache, and there is room in the cache to insert new hot blocks.

A block that was inserted into cache via a read caching policy is a copy of a block on disk, so the cached block is simply marked for eviction and will be erased when the next erasure operation occurs. A block inserted into cache via the write caching policy is written to disk via a CP write operation before it is marked for eviction from the cache.

Do all volumes provisioned on a Flash Pool aggregate use the cache?

Answer: Yes. When a Flash Pool aggregate is created, the default read and write caching policies are applied to all existing, and subsequently-created, volumes. The caching policies of a volume can be modified by using the node-level “priority hybrid-cache set” command, or starting with Data ONTAP 8.3,

the cluster-shell “volume modify” command. The complete syntax for each command, and the policy options, are covered in the Configuration section of this FAQ.

Is Flash Pool cache divided into separate segments for read-inserted and write-inserted blocks?

Answer: No. Flash Pool cache is a single cache space; there are separate caches for reads and writes. Flash Pool cache is best characterized as a random read cache that has a default write-back caching policy for random overwrites.

Does using a Flash Pool aggregate improve write latency or write throughput performance?

Answer: No, in most cases. Data ONTAP delivers low write latency by buffering write data in memory, and logging a copy in NVRAM; writes are acknowledged as soon as these operations are completed which is long before any random write data is inserted into Flash Pool cache.

Situations such as back-to-back consistency points (CPs) are typically caused by the inability of HDDs to keep up with the amount of write data being flushed from write memory buffers. For random write workloads that are disk bound (or close to it), random writes to Flash Pool cache may help by offloading those operations from HDDs to SSD cache. Offloading random write I/O to SSDs may enable the HDDs to keep up better with sequential write and write append operations.

Does a Flash Pool aggregate increase the maximum throughput capability of a storage controller?

Answer: No, configuring one or more Flash Pool aggregates on a controller does not increase the maximum throughput capability of the controller. The maximum throughput capability of a controller, for a given mix of workloads, is determined by the controller’s memory, CPU cores, and the version of Data ONTAP in use. Converting an HDD-only aggregate to a Flash Pool aggregate where the throughput performance of volumes on the aggregate, or of the controller, is constrained by the number and type of HDDs configured in the aggregate, will often increase throughput performance. However, the maximum throughput performance is still bounded by controller memory and CPU resources.

> **Note:** If the throughput of a controller is constrained by memory or CPU utilization, converting a HDD-only aggregate into a Flash Pool aggregate will not improve performance. Moving workloads to another controller, or upgrading to a higher performing controller, are recommended to address performance constraints due to high memory or CPU utilization.

Can blocks, LUNs, files or entire volumes be pinned in Flash Pool cache?

Answer: No. Blocks are inserted into, and evicted from, Flash Pool cache based on the cache insertion policies and cache retention algorithms. There is no policy or mechanism to pre-warm the cache with data that doesn’t match a policy, nor a mechanism to circumvent the data retention algorithm.

What is the eviction scanner?

Answer: The eviction scanner runs the cached data retention process assuring that hotter blocks are retained in cache longer than cooler blocks, and that cold blocks are marked for eviction so that cache space becomes available to insert new hot data into the cache.

When and how often does the eviction scanner run?

Answer: The eviction scanner checks every 60 seconds to determine if it needs to start running. If the eviction scanner is already running, it continues to run; otherwise, it starts running if one of two conditions is met:

- The cache is 75% or more full (85% full with Data ONTAP 8.3 and later releases)
- System heuristics project that the cache will exceed 75% (or 85%) usage within the next 60 minutes.

If the scanner is not already running and neither of the preceding two conditions is met, the process remains inactive (goes back to sleep).

How full can a Flash Pool cache get?

Answer: With Data ONTAP 8.1.x and 8.2. releases, the eviction scanner starts to run when a Flash Pool cache is 75% full, or if Data ONTAP projects it will become 75% full within one hour. With Data ONTAP 8.3 and later releases, the eviction scanner threshold is 85%. With all Data ONTAP releases, if the cache becomes 90% full, no new data will be inserted into the cache until the level drops below 90%.

Note: When sizing the capacity of a cache, the eviction scanner start threshold should be factored in which enables the cache to hold the estimated working set without forcing data eviction.

How is the retention of blocks in a Flash Pool cache managed?

Answer: Data ONTAP uses a “heat map” to assign a value to each block in Flash Pool cache. Each new block inserted into a cache is assigned a neutral priority or “warmth” level. A block inserted via the read caching policy is elevated to the next higher warmth level when it is read between eviction scanner runs. If a block is not read between eviction scanner runs, it is demoted a level. When a block is demoted to the lowest level, it will be evicted from the cache in the next run of the eviction scanner.

A block inserted via the write caching policy is also assigned neutral warmth. With releases prior to Data ONTAP 8.3, a write-inserted block is demoted with each eviction scanner run until it is replaced by write update. When a block reaches the lowest warmth level, it is destaged (copied) to HDD before it is evicted from cache in the next scanner run. Starting with Data ONTAP 8.3, a write-inserted block can be promoted if it is read between eviction scanner runs, however the highest warmth it can reach is neutral.

Is there a performance impact to the controller when the eviction scanner is running?

Answer: The eviction scanner is a very low-impact process. Extensive testing shows that the eviction scanner has no effect on end-user or application performance . What might be visible when monitoring controller performance is very active destaging of blocks from cache to HDDs, or cache insertion and eviction churn that can occur when the cache much smaller than the cacheable working set; this is often called cache overrun.

Is the capacity of the SSDs in a Flash Pool aggregate counted towards the aggregate capacity?

Answer: No. The capacity of the SSDs is treated as cache, and it does not count as part of the aggregate capacity. Only the data HDDs in a Flash Pool aggregate count toward the maximum aggregate storage capacity.

Do the SSDs in Flash Pool aggregates count toward the maximum drive count for a system?

Answer: Yes. All drives, including SSDs that are used for cache capacity in Flash Pool aggregates, count toward the maximum supported drive count of a system.

Do the SSDs used in Flash Pool aggregates affect the maximum storage capacity of a system?

Answer: Yes. The SSDs in Flash Pool aggregates, and spares, are included in the total drive count of system. Consequently, the total quantity of SSDs reduce the maximum number of HDDs that can be configured on a system, thereby decreasing the maximum configurable storage capacity.

Is 10% of Flash Pool cache capacity allocated for WAFL reserve?

Answer: No. The SSDs in a Flash Pool aggregate are cache, and not storage. The cache capacity is excluded from space that is allocated for aggregate-level WAFL reserve.

Is Flash Pool SSD cache capacity decremented by aggregate-level Snapshot® reserve?

Answer: No. The SSDs in a Flash Pool aggregate are considered to be cache, and not storage; the cache capacity is excluded from space that is reserved for aggregate Snapshot™ copies.

Note: Except for Flash Pool aggregates that are configured with local SyncMirror or as part of a MetroCluster system, aggregate Snapshot copies should not be used.

Does Data ONTAP use the TRIM command when managing the SSDs?

Answer: No. The process of freeing up stale data blocks is performed by a background process within Data ONTAP. This process eliminates the need to use the TRIM feature.

How does wear leveling work with SSDs in a Flash Pool aggregate?

Answer: A characteristic of all flash-based SSDs is a specified finite number of cycles that data can be written to any cell within the media. Wear leveling is a standard feature of SSDs that assures data is written evenly across all the cells such that all cells wear at roughly the same rate. Wear leveling is done with SSDs used for Flash Pool cache just as it is for SSDs used as storage. For additional information on wear leveling and NetApp SSD, see TR-3938: NetApp and Solid State Drive Technology.

How is media usage monitored for SSDs used in a Flash Pool aggregate?

Answer: SSD media is monitored using the node-level command “storage disk show –a”, and specifying a SSD ID. The output includes a percentage of the rated media life that has been used, and a percentage of the spare blocks that have been consumed. The spare blocks consumed percentage is the better indicator of how much of the SSD media has been used. A sample output of the “storage disk show –a” command is shown in figure 1, followed by descriptions of some of the fields.

Figure 1) A sample “storage disk show –a” output

```
stl6280DLEcmode1::> node run -node stl6280DLEcmode1-01
stl6280DLEcmode1-01> storage show disk -a 6a.00.0
Disk: 6a.00.0
Shelf: 0
Bay: 0
Serial: XXVE0P4A
Vendor: NETAPP
Model: X446_RALS200MCHT
Rev: NA01
RPM: N/A
WWN: 5:000:cca013:17ad88
UID: 5000CCA0:1317AD88:00000000:00000000:00000000:00000000:00000000:00000000:00000000:00000000
Downrev: no
Pri Port: A
Sec Name: 6b.00.0
Sec Port: B
Power-on Hours: 8156
Blocks read:      8827073939
Blocks written:   11862011056
Time interval: 196:10:38
Glist count: 0
Scrub last done: 00:15:37
Scrub count: 161
LIP count: 0
Dynamically qualified: No
Power cycle count: 0
Power cycle on error: 0
Current owner: 1873768414
Home owner: 1873768414
Reservation owner: 1873768414
Rated life used: 0 %
Spare blocks consumed: 0 %
Spare blocks consumed limit: N/A
```

- **Power-on Hours**: This is the number of hours the SSD has been powered on.
- **Rated life used:** This is a vendor-specific estimate of the percentage of the life used based on actual device usage and the manufacturer's projection of device life. A value of 100 (or greater) indicates that the estimated endurance has been consumed, but it might not indicate a device failure.

- **Spare blocks consumed:** Reported by the SSD as a percentage, it is the current worst-case state of the internal resource pool of the SSD. This parameter is monitored by Data ONTAP to trigger Event Management System (EMS) and AutoSupport™ events when the spare-blocks-consumed limit has been reached.
- **Spare blocks consumed limit:** This is the threshold value used internally by the SSD to trigger a predictive failure event based on spare blocks consumed. The limit for currently shipping SSDs is 90% even though Data ONTAP currently shows the spare-blocks-consumed limit as N/A.

4 Configuration

How is a Flash Pool aggregate created?

- The three basic steps to creating a Flash Pool aggregate and the process can be performed using System Manager 2.2, 3.0 or a later release (including System Manager 8.3). If the command line interface is used, the steps shown in table 2 are used. If an existing HDD-only aggregate is being converted into a Flash Pool aggregate, the first step is skipped.

Table 2) Steps to create a Flash Pool aggregate

Step	7-Mode Commands	Clustered ONTAP Commands
Create an HDD-only aggregate	`aggr create <aggr name>` with the appropriate options to set disk type, RAID type, RAID size, etc.	`storage aggregate create -aggregate <aggr name>` with the appropriate options to set disk type, RAID type, RAID size, etc.
Convert the aggregate into a Flash Pool aggregate	`aggr options <aggr name> hybrid_enabled on`	`storage aggregate modify -aggregate <aggr name> -hybrid-enabled true`
Add a SSD RAID group	`aggr add <aggr name>` with the appropriate options to set disk type, RAID type, RAID size, etc.	`storage aggregate add-disks -aggregate <aggr name>` with the appropriate options to set disk type, RAID type, RAID size, etc.

The full syntax with options for the “aggr create and “aggr add” commands are in the7-Mode Storage Management Guide. The syntax with options for the “storage aggregate create” and “storage aggregate add-disks” commands are in the Clustered Data ONTAP Physical Storage Management Guide.

Are the SSDs and HDDs in a Flash Pool aggregate configured in the same RAID group?

Answer: No. SSDs and HDDs are different drive types and must be configured in separate RAID groups.

Can more than one type of HDD be configured in a Flash Pool aggregate?

Answer: No. A Flash Pool aggregate consists of SSDs and only one type of HDD: 7200 rpm, 10,000 rpm, or 15,000 rpm disk drives. HDDs that have different rotational speeds cannot be used in a Flash Pool aggregate. The SSDs and HDDs are configured in separate RAID groups.

Note: Provisioning SAS and FC-AL disk drives that have the same capacity and rotational speed in the same aggregate is not recommended because the SAS stack architecture is more resilient than the FC-AL architecture.

Can SSDs, or HDDs, with different capacities be used in the same Flash Pool aggregate?

Answer: Yes. A Flash Pool aggregate has one or more RAID groups of SSDs, and one or more RAID groups of one type of HDD. All the drives in a RAID group should have the same usable capacity; if a

drive with a larger right-sized capacity is added to a RAID group of smaller drives, the only the smaller capacity will be usable on the larger drive. Best practice is to provision different capacity drives in separate RAID groups so that the full capacity of each drive can be used. If larger capacity SSDs or HDDs are added to a Flash Pool aggregate, a new RAID group should be created for the added drives.

Can an HDD-only aggregate be converted into a Flash Pool aggregate?

Answer: Yes. In fact, a Flash Pool aggregate is provisioned by first creating a HDD-only aggregate, and then converting the aggregate into a Flash Pool aggregate and adding a SSD RAID group. An HDD-only aggregate that is already provisioned with active volumes can be converted into a Flash Pool aggregate while those volumes are being accessed for read and write I/O.

Can a Flash Pool aggregate be converted into an HDD-only Flash Pool aggregate?

Answer: No. The SSDs, or SSD storage pool partitions, that are used as Flash Pool cache are provisioned in a RAID group, and a RAID group cannot be removed or destroyed within an active aggregate. Even though the SSDs are principally used as a cache, they are still part of the aggregate. If the objective is to repurpose the SSDs that are currently used for Flash Pool cache elsewhere on the same or another controller, all data must first be migrated from the Flash Pool aggregate to a another aggregate. Then the source Flash Pool aggregate and all RAID groups are destroyed, and the SSDs and HDDs are returned to the spare drives pool.

Can an all-SSD aggregate be converted into a Flash Pool aggregate by adding HDDs?

Answer: No. In an all-SSD aggregate, the SSDs are used to store data. These SSDs cannot be used as Flash Pool cache without first migrating all the data from the SSDs to another aggregate, destroying the all-SSD aggregate and releasing the drives to the spares pool. Only then can the SSDs be repurposed for cache use in a Flash Pool aggregate.

Can SSDs be added to a Flash Pool aggregate to increase the cache size?

Answer: Yes. SSDs can be added to an existing SSD RAID group (up to the maximum supported size of the RAID group) or by creating a new RAID group.

Can SSDs be removed from a Flash Pool aggregate to decrease the cache size?

Answer: No. SSDs cannot be removed from a RAID group, and a RAID group of SSDs cannot be removed from a Flash Pool aggregate. These restrictions also apply to HDDs in all types of aggregates.

Can a root aggregate that is provisioned with dedicated (non-partitioned) HDDs be a Flash Pool aggregate?

Answer: Yes, the root volume aggregate can be a Flash Pool aggregate if it is provisioned with dedicated (non-partitioned) HDDs and SSDs. Support for this requires Data ONTAP 8.1.2 or later release. Configuring a root aggregate that uses dedicated HDDs as a Flash Pool aggregate, with cache from shared (partitioned) SSD storage pool is not recommended.

Can a root aggregate that is provisioned using advanced data partitioning (ADP) be a Flash Pool aggregate?

Answer: No. Only a root aggregate that is provisioned on dedicated (non-partitioned) HDDs can be converted into a Flash Pool aggregate. A root aggregate created with advanced data partitioning (ADP) cannot become a Flash Pool aggregate.

Can Flash Pool cache be provisioned using shared (partitioned) SSDs?

Answer: Yes. Starting with Data ONTAP 8.3, the cache in a Flash Pool aggregate can be provisioned using either dedicated (non-shared) or shared (partitioned) SSDs that are configured as a *storage pool*.

A Flash Pool aggregate can use only one of the other type of cache; that is, dedicated and shared SSDs cannot be used in the same Flash Pool aggregate.

What is a storage pool?

Answer: A *storage pool* is a group of like-capacity SSDs, and each drive is divided into four equal *partitions*. Each row of partitions (called an *allocation unit*) can be assigned to a Flash Pool aggregate, and up to four Flash Pool aggregates within a HA pair can have cache provisioned from the same storage pool. Or, more than one allocation unit can be assigned to the same Flash Pool aggregate. A RAID policy is selected when an allocation unit is assigned to an aggregate.

How is Flash Pool cache provisioned using a storage pool of shared (partitioned) SSDs?

Answer: Provisioning Flash Pool cache from a storage pool (a shared and partitioned group of physical SSDs) with Data ONTAP 8.3 can be done using System Manager, which runs within ONTAP 8.3). Alternatively, the cluster–shell command line interface can be used to do this; the commands and process are documented in the *Physical Storage Management Guide*.

Can SSDs be added to a storage pool to increase its size?

Answer: Yes. SSDs of the same capacity as those already in a storage pool can be added as data drives, up to the maximum RAID group size supported for the RAID type used for the allocation units.

Can SSDs be removed from a storage pool to decrease its size?

Answer: No. SSDs cannot be removed from a storage pool. To liberate SSDs from a storage pool requires destroying the storage pool, and that can only be done after all the Flash Pool aggregates that use an allocation unit from the storage pool are destroyed.

What are the benefits of using a storage pool to provision Flash Pool cache?

Answer: A storage pool enables two to four Flash Pool aggregates within a HA pair to share one pool of SSDs for the purpose of provisioning cache. The primary benefit is better cache capacity efficiency with only one or two physical drives for parity as compared to four or as many as eight when non-partitioned SSDs are used. In addition, performance can be better than a Flash Pool cache that has three or fewer non-partitioned data SSDs if the storage pool allocation units are spread across a larger number of SSDs.

When should Flash Pool cache be provisioned from a storage pool versus using dedicated (non-partitioned) SSDs?

Answer: The prime use case for a storage pool is when provisioning cache for one or two Flash Pool aggregates on each node of a FAS2200 or FAS8020 HA pair. By using a storage pool instead of dedicated SSDs, two to four fewer SSDs are needed (for parity and spares) and better cache performance may be realized. For larger Flash Pool caches, and deployments with multiple Flash Pool aggregates across both nodes in a HA pair, the constraints associated with storage pool use make cache provisioned with dedicated (non-partitioned) SSDs more attractive. Key things to consider when deciding whether to use a storage pools are:

- All four allocation units are the same size and grow at the same rate when SSDs are added to a storage pool.
- The RAID type – RAID 4 or RAID-DP – applied to an allocation unit cannot be changed.
- The RAID reconstruction rate of a failed SSD in a storage pool, when allocation units are assigned to both nodes, is up to 10 times slower than when SSDs are not shared between nodes.
- It is important to provision enough data SSDs in a storage pool to meet the cache I/O demands of all the Flash Pool aggregates that share the SSDs.

What is the recommended minimum number of data SSDs per Flash Pool aggregate when dedicated (non-partitioned) SSDs are used?

Answer: The recommended minimum number of dedicated (non-partitioned) data SSDs per Flash Pool aggregate is a function of the controller model. Although SSDs are capable of handling high IOPS, a small number of them are required to provide the desired cache throughput performance. Table 3 shows the recommended minimum number of data SSDs per Flash Pool aggregate.

Table 3) Recommended minimum number of SSD data drives per Flash Pool aggregate.

System Family	100GB SSD	200GB SSD	400GB SSD	800GB SSD	1.6TB SSD
FAS2200 series	1	1	1	Not supported	Not supported
FAS2500 series	1	1	1	1	1
FAS/V3100 series[1]	3	2	2	Not supported	Not supported
FAS/V3200 series[2]	3	2	2	2[3]	2[4]
FAS8020 / FAS8040	3	2	2	2	2
FAS/V6000 series	9	5	5	Not supported	Not supported
FAS/V6200 series	9	5	5	5	5[5]
FAS8060 / FAS8080 EX	9	5	5	5	5

Note: (1) Use in Flash Pool aggregates is supported with the FAS/V3160 and FAS/V3170 only.

Note: (2) Use in Flash Pool aggregates is supported on all models except FAS/V3210.

Note: (3) Use in Flash Pool aggregates is supported on all models, starting with Data ONTAP 8.2.1, except FAS/V3210.

Note: (4) Use in Flash Pool aggregates is supported on FAS/V3250 only, starting with Data ONTAP 8.2.1. Use is supported on all models, except FAS/V3210, with Data ONTAP 8.3.

Note: (5) Use in Flash Pool aggregates is supported on all models, starting with Data ONTAP 8.2.1, except FAS/V6210. Use is supported on all models with Data ONTAP 8.3.

What is the recommended minimum number of data SSDs in a storage pool?

Answer: The recommended minimum number of data SSDs in a storage pool is a function of the controller model. Although SSDs are capable of handling high IOPS, a small number of them are required to provide the desired cache throughput performance. Table 4 shows the recommended minimum number of data SSDs per storage pool.

Table 4) Recommended minimum number of SSD data drives per storage pool.

System Family	100GB SSD	200GB SSD	400GB SSD	800GB SSD	1.6TB SSD
FAS2200 series	Not supported	2	2	Not supported	Not supported
FAS2500 series	Not supported	2	2	2	2
FAS/V3200 series[1]	Not supported	2	2	2	2
FAS8020 / FAS8040	Not supported	3	3	3	3
FAS/V6200 series	Not supported	5	5	5	5
FAS8060 / FAS8080 EX	Not supported	5	5	5	5

Note: Because a storage pool provides cache for up to 4 Flash Pool aggregates, provisioning the minimum number of data SSDs in a storage pool might throttle Flash Pool cache performance. It is recommended that SPM be used to size the number of SSDs needed in a storage pool to meet Flash Pool aggregate workloads.

Note: (1) FAS/V3210 is not supported with Data ONTAP 8.3.

What is the recommended minimum number of dedicated (non-partitioned) SSDs to add to of a Flash Pool aggregate when expanding the cache?

Answer: The recommended minimum number of SSDs to add when expanding the cache capacity of a Flash Pool aggregate is a function of controller model, with the intent being good cache performance after expansion. Table 5 lists the recommended minimum number of SSDs to add per Flash Pool aggregate.

Table 5) Recommended minimum number of dedicated SSDs to add when expanding Flash Pool cache.

System Family	100GB SSD	200GB SSD	400GB SSD	800GB SSD	1.6TB SSD
FAS2200 series	1	1	1	Not supported	Not supported
FAS2500 series	1	1	1	1	1
FAS/V3100 series[1]	3	2	2	Not supported	Not supported
FAS/V3200 series[2]	3	2	2	2[3]	2[4]
FAS8020 / FAS8040	3	2	2	2	2
FAS/V6000 series	6	3	3	Not supported	Not supported
FAS/V6200 series	6	3	3	3	3[5]
FAS8060 / FAS8080 EX	6	3	3	3	3

Note: (1) Use in Flash Pool aggregates is supported with the FAS/V3160 and FAS/V3170 only.

Note: (2) Use in Flash Pool aggregates is supported on all models except FAS/V3210.

Note: (3) Use in Flash Pool aggregates is supported on all models, starting with Data ONTAP 8.2, except FAS/V3210.

Note: (4) Use in Flash Pool aggregates is supported on FAS/V3250 only, starting with Data ONTAP 8.2.1. Use is supported on all models, except FAS/V3210, with Data ONTAP 8.3.

Note: (5) Use in Flash Pool aggregates is supported on all models, starting with Data ONTAP 8.2.1, except FAS/V6210. Use is supported on all models with Data ONTAP 8.3.

What is the recommended minimum number of SSDs to add when expanding a storage pool?

Answer: The recommended minimum number of SSDs to add to storage pool is a function of the controller model, with the intent being good cache performance after expansion. Table 6 lists the recommended minimum number of SSDs to when expanding a storage pool.

Table 6) Recommended minimum number of SSD to add when expanding a storage pool.

System Family	100GB SSD	200GB SSD	400GB SSD	800GB SSD	1.6TB SSD
FAS2200 series	Not supported	1	1	Not supported	Not supported
FAS2500 series	Not supported	1	2	1	1
FAS/V3200 series[1]	Not supported	2	2	2	2
FAS8020 / FAS8040	Not supported	2	2	2	2
FAS/V6200 series	Not supported	3	3	3	3
FAS8060 / FAS8080 EX	Not supported	3	3	3	3

Note: Because a storage pool provides cache for up to 4 Flash Pool aggregates, provisioning the minimum number of data SSDs in a storage pool might throttle Flash Pool cache performance. It is recommended that SPM be used to size the number of SSDs needed in a storage pool to meet Flash Pool aggregate workloads.

Note: (1) FAS/V3210 is not supported with Data ONTAP 8.3.

How many Flash Pool aggregates are supported on a single system?

Answer: There is no specified maximum number of Flash Pool aggregates that can be configured on a system. A maximum number can be calculated based on the recommended minimum number of data SSDs per Flash Pool aggregate (as shown in table 4) and the maximum number of data SSDs that shown in table 6.

For best utilization of, and storage efficiency from, the SSDs in Flash Pool aggregates, it is recommended to configure a small number of Flash Pool aggregates, each of which has relatively large HDD storage capacity. This results in more of the SSDs being used as data drives and thus larger Flash Pool caches.

How many storage pools are supported on a single system?

Answer: There is no specification for the maximum number of storage pools per controller or HA pair, however the maximum Flash Pool cache size or maximum number of SSDs per system do limit how many storage pools can be configured. The best practice recommendation is to use storage pools for Flash Pool cache provisioning only when a small number of Flash Pool aggregates with similar cache sizes are planned for the system.

How is Flash Pool cache size calculated?

Answer: For a Flash Pool cache provisioned with dedicated (non-partitioned) SSDs, the cache size is calculated using the formula:

Number of data SSDs in the aggregate × *right-sized used capacity* per SSD

For example, a Flash Pool cache that is provisioned with six 200GB data drives has a cache size equal to 6 x 190,532 MiB, which totals 1,143,192 MiB (or 1,116.4 GiB, or 1.09 TiB).

The command `df -AH -hybrid` shows the available and used SSD space in a Flash Pool aggregate.

The same calculation methodology is used to determine the total of all Flash Pool caches provisioned on a controller or a HA pair. Table 1 lists the right-sized used capacities of the SSDs supported for Flash Pool use.

For a Flash Pool cache provisioned from one or more storage pool allocations units, the size of each allocation unit can be calculated using this formula:

[number of SSDs in the storage pool – parity SSDs] × *right-sized used capacity* per SSD × 25%

The number of parity drives will be 1 or 2, based on whether RAID 4 or RAID-DP protection will be used for the allocation units. The product of the calculation is a close approximation of the actual allocation unit usable capacity. After allocation units are assigned to a Flash Pool aggregate the actual cache capacity can be viewed using the command "storage pool show-aggregate".

Consult the Clustered Data ONTAP *Physical Storage Management Guide* and *Manual Page Reference* for more information about storage pool commands.

What is the maximum number of dedicated (non-partitioned) SSDs per Flash Pool aggregate and per system (HA pair or single controller) when used in Flash Pool aggregates?

Answer: The maximum number of SSDs per Flash Pool, and per system when SSDs are used as Flash Pool cache, varies by FAS model, Data ONTAP release, and the capacity of the SSD used. Only data SSDs count against the Flash Pool cache maximum. In table 7, column 1 lists the maximum Flash Pool cache sizes that apply to different FAS models and across multiple Data ONTAP releases. Columns 2 through 6 show the maximum number of data SSDs configurable within each maximum cache size. The maximum number of data SSDs is calculated by dividing the right-sized usable capacity of the drive (see table 1) into the maximum Flash Pool cache size.

Table 7) Maximum number of SSD data drives in Flash Pool aggregate by maximum cache size.

Flash Pool Max Cache Size[1]	100GB[2] Data SSDs	200GB[2] Data SSDs	400GB[2] Data SSDs	800GB[2] Data SSDs	1.6TB[2] Data SSDs
300 GiB	3	1	Not supported [3]	Not supported	Not supported
400 GiB	4	2	1	Not supported	Not supported
0.5 TiB	5	2	1	Not supported	Not supported
0.6 TiB	6	3	1	Not supported	Not supported
800 GiB	8	4	2	Not supported	Not supported
1.0 TiB	11	5	2	Not supported	Not supported
1.2 TiB	13	6	3	Not supported	Not supported
1.6 TiB	17	8	4	2	Not supported
2.0 TiB	22	11	5	2	Not supported
3.2 TiB	35	17	8	4	2
4.0 TiB	44	22	10	5	2
6.0 TiB	66	33	16	8	4
6.4 TiB	70	5	17	8	4
8.0 TiB	88	44	21	10	5

Flash Pool Max Cache Size[1]	100GB[2] Data SSDs	200GB[2] Data SSDs	400GB[2] Data SSDs	800GB[2] Data SSDs	1.6TB[2] Data SSDs
12.0 TiB	132	66	32	16	8
16.0 TiB	176	88	44	21	10
18.0 TiB	198	99	49	24	12
24.0 TiB	240[4]	132	65	32	16
36.0 TiB	240[4]	198	98	49	24
48.0 TiB	240[4]	240[4]	131	65	32
72.0 TiB	240[4]	240[4]	197	98	49
96.0 TiB	240[4]	240[4]	240[4]	131	65
144.0 TiB	240[4]	240[4]	240[4]	197	98

Note: (1) Each value in this column matches the maximum Flash Pool cache size for one or more FAS/V system model running one or more Data ONTAP release.

Note: (2) SSD part numbers, right-sized usable capacities, and Data ONTAP release requirements are listed in Table 1.

Note: (3) Not supported, wherever shown, means the maximum number of data SSDs is less than the minimum number recommended in T\table 3.

Note: (4) The maximum number of SSDs supported with FAS/V3xx0, FAS/V6xx0, and FAS80x0 HA pairs is 240. The maximum number of data SSDs in Flash Pool aggregates or storage pools is less than 240 because parity and spare drives count towards the limit.

What type of RAID protection is recommended for a SSD RAID group in a Flash Pool aggregate on a system that is running Data ONTAP 8.1.1 or a later 8.1.x release?

Answer: For systems that run Data ONTAP 8.1.1 and later 8.1.x releases, the SSD and HDD RAID groups in a Flash Pool aggregate must have the same RAID policy, which means RAID protection type and RAID group size. Since RAID-DP is the recommended RAID type for HDDs, the SSD RAID groups in Flash Pool aggregates will also be RAID-DP protected. The SSD RAID group will be the same size, in terms of the number of data and parity drives, as was applied to the HDD RAID groups.

For information on default and maximum RAID group sizes by drive type and RAID policy, refer to the Storage Subsystem Technical FAQ.

Note: The option “raid.min_spare_count” does not apply to SSD RAID groups in a Flash Pool aggregate. This enables a FAS2200 series that is running a Data ONTAP 8.1.x release to have a Flash Pool aggregate with four (4) SSDs and no SSD spare, without generating a low spares error message. However, the absence of a hot SSD spare means Flash Pool performance will be affected when a SSD fails, until a replacement SSD is installed and the SDD RAID-DP group rebuild is completed.

Note: The RAID group size for a HDD-only aggregate that will be converted into a Flash Pool aggregate can be increased before configuring the SSD RAID group. However, the maximum size of the RAID group is governed by the maximum size supported for the type of HDD in the aggregate (for example, 20 drives for a SATA drive smaller than 6TB in a RAID-DP group). In addition, if a HDD RAID group is added after the larger SSD RAID group is configured, the updated larger group size will apply to the new HDD RAID group. Following the best practice of configuring all HDD RAID groups to the same size (plus or minus one drive) in the aggregate requires manual attention when adding new HDD RAID groups.

What type of RAID protection is recommended for a SSD RAID group in a Flash Pool aggregate on a system that is running a Data ONTAP 8.2.x release?

Answer: For systems that run Data ONTAP 8.2 or a later release, the SSD and HDD RAID policies (RAID protection type and group size) can be different. This means a Flash Pool aggregate can be configured with the HDDs in RAID-DP groups—per best practice—while RAID 4 is used for the SSD RAID group. Or, if RAID-DP is used for both the SSD and HDD RAID groups, the group sizes for each drive type can be different; in particular, the SSD RAID group can be larger.

RAID 4 protection for the SSDs in a Flash Pool aggregate is recommended when only one SSD RAID group per Flash Pool aggregate meets the cache capacity needs and at least one hot spare SSD per controller is configured to assure immediate RAID rebuild should an SSD fail. For example, on a FAS2200 system that has one Flash Pool aggregate, a two–data drive RAID 4 SSD RAID group plus one hot spare is a better configuration that a two–data drive RAID-DP SSD RAID group and no hot spare.

Nevertheless, for best resiliency and performance in systems where the cost impact of additional parity SSDs is relatively minor, the best practice recommendation for a SSD RAID group in a Flash Pool aggregate is RAID-DP protection with at least one hot SSD spare per node.

For information on default and maximum RAID group sizes by drive type and RAID policy, refer to section 4 of the Storage Subsystem Technical FAQ.

What type of RAID protection is recommended for a RAID group of dedicated (non-partitioned) SSDs in a Flash Pool aggregate on a system that is running Data ONTAP 8.3 or a later release?

Answer: For systems that run Data ONTAP 8.3 or a later release, the SSD and HDD RAID policies (RAID protection type and group size) can be different. This means a Flash Pool aggregate can be configured with the HDDs in RAID-DP groups—per best practice—while RAID 4 is used for the SSD RAID group. Or, if RAID-DP is used for both the SSD and HDD RAID groups, the group sizes for each drive type can be different; in particular, the SSD RAID group can be larger.

RAID 4 protection for the SSDs in a Flash Pool aggregate is recommended when only one SSD RAID group per Flash Pool aggregate meets the cache capacity needs and at least one hot spare SSD per controller is configured to assure immediate RAID rebuild should an SSD fail. For example, on a FAS2200 system that has one Flash Pool aggregate, a two–data drive RAID 4 SSD RAID group plus one hot spare is a better configuration that a two–data drive RAID-DP SSD RAID group and no hot spare.

What is the best practice for RAID group size when using dedicated (non-partitioned) SSDs for Flash Pool cache?

Answer: The smallest RAID group size is based on the recommended minimum number of data SSDs per aggregate (refer to table 5). When RAID 4 protection is used, the RAID group will have one drive more than the recommended minimum number of data SSDs. When RAID-DP is used, the group size will be two drives larger than the recommended minimum number of data SSDs.

RAID 4 groups up to 14 SSDs are supported for Flash Pool cache, however RAID-DP is recommended for groups with 10 or more data SSDs.

Can a RAID-DP group of dedicated (non-partitioned) SSDs in a Flash Pool aggregate be converted to RAID-4 while the aggregate is online?

Answer: Yes. The RAID policy of a SSD RAID group can be changed online just as is the case with HDD RAID groups. The system must be running Data ONTAP 8.2 or a later release to use RAID 4 for a SSD RAID group in a Flash Pool aggregate that has HDDs protected in RAID-DP groups.

Can a RAID 4 group of dedicated (non-partitioned) SSDs in a Flash Pool aggregate be converted to RAID-DP while the aggregate is online?

Answer: Yes.

What are the RAID recommendations for storage pool allocation units used for Flash Pool cache?

Answer: Best practice is to use the same RAID type—RAID-DP or RAID 4—for all four allocation units within a storage pool. RAID-DP is recommended for storage pools that have or will expand to 8 or more SSDs. This is because up to four aggregates depend on the resiliency of the storage pool. RAID 4 can be used for storage pools of 7 or fewer SSDs.

Note: After the RAID type is selected for an allocation unit it cannot be changed. This is different from RAID groups of dedicated (non-partitioned) SSDs where RAID type is changeable.

Note: Using the same RAID type for all allocation units in a storage pool avoids resiliency risks if the storage pool is expanded beyond 7 SSDs, and unusable capacity for RAID 4 protected allocation units if the storage pool is expanded beyond 14 SSDs.

Can the RAID type used for a storage pool allocation unit be changed after the allocation unit is assigned to a Flash Pool aggregate?

Answer: No. The RAID selected for an allocation unit as it is assigned to a Flash Pool aggregate cannot be changed.

Note: When using SSDs with RAID 4 protection, at least one hot SSD spare per controller must always be configured to assure adequate data protection should an SSD fail.

What is the hot spare recommendation for SSDs in a Flash Pool aggregate?

Answer: A hot spare SSD is always recommended. For every node that is configured with a Flash Pool aggregate, there should be at least one SSD hot spare that has a capacity equal to the largest SSD. A hot spare is strongly recommended even when Flash Pool SSD cache is protected with RAID-DP; the hot spare enables a RAID rebuild to start immediately if a SSD fails. SSD RAID rebuilds are very fast as compared to HDD rebuilds (typically less than 1 hour), and this minimizes the potential performance impact due to a degraded Flash Pool SSD cache.

Starting with Data ONTAP 8.2, RAID 4 protection can be used for the Flash Pool SSDs in the same aggregate as HDDs that are configured in RAID-DP groups. When RAID 4 is used, a SSD hot spare is required.

What is the performance impact when a Flash Pool SSD RAID group is in degraded mode?

Answer: A degraded SSD RAID group affects the performance of the Flash Pool cache; for example, every cache read will require parity reconstruction. The rebuild of an SSD RAID group is very fast (typically less than 30 minutes) compared to HDD RAID rebuilds, which can take many hours. However, if no SSD hot spare is available on the controller, Flash Pool cache performance will be adversely affected until the failed SSD is replaced and RAID rebuild is completed. Consequently, the best practice recommendation is to provision at least one SSD hot spare per controller.

What is the system impact when a SSD RAID group in a Flash Pool aggregate fails?

Answer: If the SSD RAID group in a Flash Pool aggregate fails, the entire Flash Pool aggregate fails and is taken offline. A RAID group failure consists of two (or more) drives in a RAID 4 protected group, or three or more drives in a RAID-DP protected group, simultaneously either having uncorrectable errors or being in a failed state. However, if the SSD RAID group fails (exceeding the RAID type's ability to protect the data in the RAID group), the entire aggregate will be in a failed state and unable to serve data. Using RAID-DP is the best protection against the possibility of this failure event.

What is the maximum Flash Pool cache size supported on a single controller or HA pair system?

Answer: The maximum supported Flash Pool cache capacity varies by platform controller model and Data ONTAP release. Maximum cache sizes for Flash Cache, Flash Pool, and the combined use of Flash Cache and Flash Pool, by platform and for several Data ONTAP releases, are listed in the tables in section 7 of this FAQ. For other releases, refer to the Hardware Universe for maximum cache sizes.

The total Flash Pool cache capacity configured on a single-controller or HA pair system is the sum of the right-sized usable capacities of all data SSDs in all Flash Pool aggregates assigned to the controller, or present on the HA pair system. When multiple Flash Pool aggregates are configured on the same controller or HA pair system, each one can have a different cache size and different HDD storage capacity.

When both Flash Cache and Flash Pool are configured on a single controller or HA pair system, the combined total configured cache capacity must not exceed the combined per controller, or the combined per HA pair, maximum cache values as shown in the tables.

Note: The amount of Flash Pool cache configured on each controller of a HA pair can dissimilar. However, the cache size on a controller must not exceed the per controller maximum values that apply to systems running Data ONTAP 8.1.1 and later 8.1.x releases. For HA pairs running Data ONTAP 8.2 and later releases, the maximum Flash Pool cache size is enforced for the HA pair and not per controller; that means up to the maximum supported cache can be configured on one controller within the HA pair.

Can Flash Cache cards and Flash Pool aggregates be configured on the same system?

Answer: Yes. Flash Cache™ cards and Flash Pool aggregates can be used on the same controller, or separately each node of an HA pair system. Configuration requirements when using both caching products on the same system are covered in the Configuration section of this FAQ. Blocks from volumes that are provisioned in a Flash Pool aggregate are only cacheable in that aggregate's Flash Pool cache; blocks from volumes in a Flash Pool are automatically excluded from using Flash Cache, when Flash Cache cards are configured on the same system, even when Flash Pool caching for a volume has been disabled.

Note: Configuring Flash Cache cards and a Flash Pool aggregate on the same controller only makes sense if there are at least two data aggregates on the controller.

Note: Flash Cache and Flash Pool caching each use controller memory and CPU; configuring both products on the same controller uses more of these resources than when only one caching product is used. Using both caching products on the same controller should considered only for controller models that support at a maximum of 2.0TiB or more of combined cache.

Does Flash Cache reduce the amount of Flash Pool cache capacity allowed on a system?

Answer: Yes. When Flash Cache and Flash Pool are combined on the same controller, the combined capacity of both solutions must fall within the maximum combined cache capacity that is supported for the platform. (The maximum combined cache sizes are listed by platform model for a few Data ONTAP releases section 7.) When both Flash Cache and Flash Pool are used in a HA pair, the maximum amount of Flash Pool cache per HA pair is calculated as follows:

Maximum Flash Pool cache = Maximum combined cache – the larger* Flash Cache capacity configured on a node of the HA pair

* Each node in an HA pair may have a different amount of Flash Cache capacity.

What methods are available to size cache for a Flash Pool aggregate?

Answer: There are four methods available to size Flash Pool cache; which one to use depends on the use case.

1. For a FAS or V-Series system running Data ONTAP 8.2.1 or later releases, the Automated Workload Analyzer (AWA) is the best tool with which to size Flash Pool cache. AWA is covered extensively in later in this section.
2. For a FAS or V-Series system running a Data ONTAP release earlier than 8.2.1, and that does not have Flash Cache installed in the node, Predictive Cache Statistics (PCS) can be used to

estimate Flash Pool cache size. The next question in this section covers how to adjust PCS results for Flash Pool cache sizing.

3. For a FAS or V-Series system running a Data ONTAP release earlier than 8.2.1, and that has Flash Cache installed, the amount of Flash Cache installed is a reasonable starting point for sizing total Flash Pool cache. Note: If there are multiple data aggregates on the system and only one will be converted to a Flash Pool aggregate, sizing the Flash Pool cache to be equal to the Flash Cache size for all aggregates on the node could overstate the Flash Pool cache size, depending on the workloads of the other aggregates.
4. For sizing Flash Pool cache when a FAS or V-Series system will replace a system that is not running Data ONTAP, the grid in Table 7, which calculates cache size by multiplying the random write data rate (per day or per month) by the number of days newly written data needs to be cached, can be used.

What is Automated Workload Analyzer (AWA)?

Answer: AWA is a feature of Data ONTAP 8.2.1 (and later releases) that analyzes the total workload of all volumes and LUNs on a specified aggregate, recommends a Flash Pool cache size, and estimates the cache hit rates for that size cache as well as smaller increments. AWA only models offload

AWA is available with both clustered Data ONTAP and Data ONTAP operating in 7-Mode. AWA runs from the node shell command line interface only and requires advanced privilege access. OnCommand® management software does not currently support AWA.

AWA consists of three node-shell commands:

```
• wafl awa start <aggrname>
• wafl awa stop  <aggrname>  [-q] (quiet)
• wafl awa print <aggrname>  [-l] (intervals)
                             [-t] <n> (summaries for top n volume are printed)
```

How long should AWA run to get a good representation of an aggregate's workload?

Answer: It is recommended to run AWA for at 3 to 7 days to capture workload variations that occur over the course of several days to a week. By default AWA samples the workload in 10-minute intervals, which means 400 to 1000 intervals are needed for good results. AWA stores up to the last 1024 intervals.

More granular data about workload variation can be acquired by running a scripting the AWA print command to run every 6 hours to summarize all the interval data from AWA start plus the last 36 intervals of data.

Can AWA run on a Flash Pool aggregate?

Answer: Yes. AWA can be run on an existing Flash Pool aggregate to analyze the workload and to see if the configured cache is the optimal size for the workload.

Can AWA run on a node that has Flash Cache enabled?

Answer: Yes. AWA can run on an aggregate that is provisioned on a node that has Flash Cache cards installed and enabled. The AWA analysis ignores any benefit from Flash Cache for the volumes on the aggregate as if the volumes were already in a Flash Pool aggregate.

What does the default AWA print command output look like?

Answer: The output of the "wafl awa print" command is designed to be easy to understand. Figure 2 is a sample output for Data ONTAP 8.3.1; there are minor differences in the output format in earlier Data ONTAP releases.

Figure 2) Sample default AWA print command output

```
### FP AWA Stats ###

                     Host lada66a                     Memory 93788 MB
            ONTAP Version NetApp Release R8_3_1x_awa_2809198_1504220853: Tue Apr 21
18:35:12 PDT 2015

Basic Information
                Aggregate lada66a_aggr1
             Current-time Thu Apr 23 16:42:17 PDT 2015
               Start-time Thu Apr 23 08:03:51 PDT 2015
      Total runtime (sec) 31103
    Interval length (sec) 600
          Total intervals 54
        In-core Intervals 1024

Summary of the past 20 intervals
                                max
                                ------------
        Read Throughput (MB/s): 181.772
       Write Throughput (MB/s): 550.611
            Cacheable Read (%): 12
           Cacheable Write (%): 30
Max Projected Cache Size (GiB): 787.077

Summary Cache Hit Rate vs. Cache Size
Referenced Cache Size (GiB): 787.077
Referenced Interval: ID 53 starting at Thu Apr 23 16:33:07 PDT 2015
         Size        20%        40%        60%        80%       100%
 Read Hit (%)          9         20         28         32         35
Write Hit (%)         17         21         23         25         30

The entire results and output of Automated Workload Analyzer (AWA) are estimates.
The format, syntax, CLI, results and output of AWA may change in future Data ONTAP
releases. AWA reports the projected cache size in capacity. It does not make
recommendations regarding the number of data SSDs required. Please follow the
guidelines for configuring and deploying Flash Pool that are provided in tools and
collateral documents. These include verifying the platform cache size maximums and
minimum number and maximum number of data SSDs.

### FP AWA Stats End ###
```

Here is a brief explanation of the AWA summary print output:

- **Read throughput and write throughput**: These are the values from the interval that had the highest total throughput. Random and sequential transfers are included in each value.
- **Cacheable read and cacheable write**: These values are from the interval that had the highest total cacheable rates. The values may be from a different interval from the one reported for throughputs. All random reads are cacheable by Flash Pool, so the cacheable read value is the percentage of the read workload that is random. The cacheable write value reflects the percentage of the write workload that would be inserted into Flash Pool cache per the default “random-write” write caching policy.
- **Max projected cache size**: This is the recommended cache size for the aggregate based on the set of intervals in the report. The cache size typically increases as more intervals are included in the output.
- **Summary cache hit rate vs. cache size**: This table shows what the projected cache hit rates would be for cache sizes that are smaller than the max projected cache size (100%). A small number of outlier intervals can disproportionately influence the calculation of the max projected cache size; when that occurs a smaller cache may deliver close to or the same cache hit rates as the max projected cache size.

Does AWA provide volume-level granularity for workload analysis?

Answer: Yes. Starting with Data ONTAP 8.3.1, the “wafl awa print” command option “-t n” enables reporting of workload and projected cache hit rates and size for a user specified “n” number of volumes within the aggregate. The volumes are reported in descending order based on the highest combination of throughput and projected cache hit rates in the aggregate. If the value for “n” is not specified, the default is 8.

Figure 3 is a sample print command output with volume-level details.

Figure 3) Sample AWA print output with volume detail

```
### FP AWA Stats ###

                     Host lada66a                    Memory 93788 MB
            ONTAP Version NetApp Release Rfullsteam_awarc_2662016_1501071654: Wed
Jan  7 17:43:42 PST 2015

Basic Information
               Aggregate lada66a_aggr1
            Current-time Fri Jan  9 16:14:29 PST 2015
              Start-time Fri Jan  9 12:30:16 PST 2015
      Total runtime (sec) 13452
    Interval length (sec) 600
          Total intervals 24
        In-core Intervals 1024

Summary of the past 20 intervals
                                max
                                ------------
        Read Throughput (MB/s): 134.059
       Write Throughput (MB/s): 1333.279
            Cacheable Read (%): 27
           Cacheable Write (%): 22
Max Projected Cache Size (GiB): 216.755

Summary Cache Hit Rate vs. Cache Size
Referenced Cache Size (GiB): 216.755
Referenced Interval: ID 22 starting at Fri Jan  9 16:04:38 PST 2015
         Size        20%        40%        60%        80%       100%
 Read Hit (%)          1          3          6         11         21
Write Hit (%)          5         11         13         14         23

AWA Summary for top 8 volumes

        Vol interval len (sec) 19200
     In-core volume intervals 8

     Volume #1 lada66a_vol8
     Summary of the past 32 intervals
                                     max
                                     ------------
             Read Throughput (MB/s): 1.751
            Write Throughput (MB/s): 18.802
                 Cacheable Read (%): 11
                Cacheable Write (%): 16
     Max Projected Cache Size (GiB): 29.963
             Projected Read Hit (%): 31
            Projected Write Hit (%): 16

       Volume #2 lada66a_vol7
     Summary of the past 32 intervals
                                     max
                                     ------------
             Read Throughput (MB/s): 1.640
            Write Throughput (MB/s): 17.691
                 Cacheable Read (%): 14
                Cacheable Write (%): 13
     Max Projected Cache Size (GiB): 28.687
             Projected Read Hit (%): 29
            Projected Write Hit (%): 13
```

How much CPU and memory are used when AWA is running?

Answer: AWA uses 2% or less CPU, regardless of the number of aggregates that are simultaneously under analysis, and it uses less than 0.2% of system memory per aggregate.

Can AWA be run on more than one aggregate at a time?

Answer: Yes. AWA can be run simultaneously on multiple aggregates at the same time, up to the total number of aggregates supported per node, without concern about an impact to performance.

Can AWA be used to size cache for Flash Cache?

Answer: Yes. However, AWA is designed for Flash Pool workload analysis and cache sizing. To use AWA for Flash Cache sizing all aggregates on a node must be analyzed and the resulting recommended cache sizes then added manually. In addition, because AWA sizes cache based on random reads and random overwrites, and Flash Cache only caches reads, AWA overstates the cache size needed for Flash Cache.

How can Flash Pool cache size be estimated when a FAS system is replacing a system that doesn't run Data ONTAP?

Answer: Table 7 can be used to get a range of cache sizes based on estimates of the random write data rate and the target retention time for data in cache.To calculate cache sizes for random write data rates and cache retention periods that aren't in the table, use the following formula:

Random write data rate (GiB/day) × cache retention time (days) ÷ 1024 = cache size (TiB)

Figure 4) Cache size estimation using write data rate and retention time.

Usable Cache Size / Working Set Size

New Random Data		Target Cached Data Retention Period						
Per Day	Per Mo	7 days	15 days	30 days	45 days	60 days	90 days	120 days
20 GiB	0.6 TiB	0.1 TiB	0.3 TiB	0.6 TiB	0.9 TiB	1.2 TiB	1.8 TiB	2.3 TiB
40 GiB	1.2 TiB	0.3 TiB	0.6 TiB	1.2 TiB	1.8 TiB	2.3 TiB	3.5 TiB	4.7 TiB
60 GiB	1.8 TiB	0.4 TiB	0.9 TiB	1.8 TiB	2.6 TiB	3.5 TiB	5.3 TiB	7.0 TiB
80 GiB	2.3 TiB	0.5 TiB	1.2 TiB	2.3 TiB	3.5 TiB	4.7 TiB	7.0 TiB	9.4 TiB
100 GiB	2.9 TiB	0.7 TiB	1.5 TiB	2.9 TiB	4.4 TiB	5.9 TiB	8.8 TiB	11.7 TiB
120 GiB	3.5 TiB	0.8 TiB	1.8 TiB	3.5 TiB	5.3 TiB	7.0 TiB	10.5 TiB	14.1 TiB
140 GiB	4.1 TiB	1.0 TiB	2.1 TiB	4.1 TiB	6.2 TiB	8.2 TiB	12.3 TiB	16.4 TiB
160 GiB	4.7 TiB	1.1 TiB	2.3 TiB	4.7 TiB	7.0 TiB	9.4 TiB	14.1 TiB	18.8 TiB
180 GiB	5.3 TiB	1.2 TiB	2.6 TiB	5.3 TiB	7.9 TiB	10.5 TiB	15.8 TiB	21.1 TiB
200 GiB	5.9 TiB	1.4 TiB	2.9 TiB	5.9 TiB	8.8 TiB	11.7 TiB	17.6 TiB	23.4 TiB

Note: This cache size calculation assumes that all new random write data will be actively read for a period of time afterward and thus should be cached to provide best performance.

Can Predictive Cache Statistics (PCS) be used to estimate how much Flash Pool cache would be beneficial to add to a system?

Answer: PCS projects the performance benefit of adding Flash Cache cards to a controller that currently has no configured extended cache, based on the workloads the controller is servicing. PCS only considers the benefit of caching data for read requests, and it was designed specifically for Flash Cache prediction. Therefore, the effectiveness of PCS in estimating the benefit of converting an HDD-only aggregate to a Flash Pool aggregate is limited. PCS can be used to get a rough estimate of the read

cache benefit of Flash Pool for a system that isn't configured with Flash Pool or Flash Cache if all the aggregates will be converted to Flash Pool aggregates or if the aggregates that are under consideration for conversion are known to contain volumes that have only sequential read/write workloads. Follow these guidelines when using PCS to estimate Flash Pool benefits:

- PCS can only be used on a system that has neither Flash Cache nor Flash Pool configured.
- Use a PCS cache size that is 75% (or 85% for Data ONTAP 8.3 and later) of the total SSD data drive capacity that is being considered to avoid premature eviction of hot data.
- If the workloads being analyzed by PCS are likely to have a high rate of random overwrites (for example, transaction processing), adjust the resultant PCS hit rate down by 20% because PCS only accounts for read activity. This adjustment adds conservatism in sizing the HDDs needed to support the write workload.

What happens if the maximum supported cache capacity is exceeded on a system?

Answer: Data ONTAP enforces supported cache maximums. However, Flash Pool cache expansion that exceeds the maximum cache capacity for a HA pair may not be detected immediately because the local node must poll its partner to determine how much Flash Cache and Flash Pool cache is installed. A delay in receiving the polling response means cache maximums may be exceeded if data SSDs are added without regard to the documented maximum cache sizes.

If the maximum cache capacity for a controller is exceeded, EMS messages similar to the ones below will appear periodically on the system:

```
 [localhost:raid.hybrid.SSDTotExceed:error]: The sum of sizes of SSD disks of
hybrid aggregates <actual capacity> exceeds the <HA pair limit> maximum.

[localhost: raid.hybrid.SSDHAExceed:error]: The sum of all sizes of SSD disks
of hybrid aggregates <actual capacity>, exceeds <node limit> (half of the
maximum in an HA pair). If the sum of all sizes of SSD disks of hybrid
aggregates on this node and its partner exceeds the <HA pair limit> maximum,
takeover might fail.
```

How many drives can be configured in a single SAS stack when some of the drives are SSDs?

Answer: The maximum number drives (HDDs and SSDs) per stack, when both SSDs and HDDs are present in the stack, depends on the system model, as shown in the table below.

Table 8) Maximum number of HDDs and SSDs in a mixed SAS stack.

System Model	Maximum Number Drives per Stack
All FAS/V-Series systems, except those listed below	96 [2]
FAS2220 and FAS2520[1]	60
FAS2240, FAS2252, FAS2254, and FAS8020 [1]	144 [3]
MetroClusters with FibreBridge 6500: − 1−24 SSDs in the stack − 25−48 SSDs in the stack	 168 96

Note: (1) Drives internal to the system enclosure are included in the maximum number.

Note: (2) The limit enables desired stack bandwidth and latency for high workload deployments.

Note: (3) More drives are supported because only one stack per system is supported, or because a second stack requires using a PCIe slot on a controller has only 2 slots. The maximum number of SSDs per stack is 96. (As of July 2015, QuoteEdge only allows configuration of up to 3 expansion shelves; up to 2 more shelves can be added later.)

Can different shelf types and drive types be configured in the same storage stack?

Answer: Yes. The requirements and recommendations that apply to storage stacks configured with different shelf and drive types are covered in Section 3.2 of the Storage Subsystem Technical FAQ.

Can a shelf configured with both SSDs and HDDs be used to provision a Flash Pool aggregate?

Answer: Yes. Mixed SSD and HDD shelves are supported and offered specifically for Flash Pool aggregate use. The intent is to use one mixed shelf to provision a Flash Pool aggregate with a small number SSDs (1 or 2 data drives).

For deployments that require more data SSDs per Flash Pool aggregate, or that have more than one Flash Pool aggregate per HA pair, the recommendation is to configure the cache using one or more DS2246 shelves that have SSDs only. DS2246 shelves with SSDs are offered in 12- and 24-drive configurations. It is also recommended to configure SSD-only shelves on a dedicated SAS stack, whenever possible, to provide the best possible performance from the SSDs.

Can a Flash Pool aggregate be configured on only one controller in an HA pair?

Answer: Yes. There is no requirement or best practice to configure Flash Pool aggregates on both nodes of an HA pair

Are FAS2200 series configurations that have a mix of internal SSDs and HDDs supported?

Answer: Yes. All internal SSDs and HDDs that are supported with FAS2200 series controllers and in Flash Pool aggregates are also supported in mixed SSD-HDD internal drive configurations. Note that only a small number of all the possible mixed SSD-HDD combinations are orderable, but all other combinations are supported.

Are FAS2500 series configurations that have a mix of internal SSDs and HDDs supported?

Answer: Yes. All internal SSDs and HDDs that are supported with FAS2500 series controllers and in Flash Pool aggregates are also supported in mixed SSD-HDD internal drive configurations. Note that only a small number of all the possible mixed SSD-HDD combinations are orderable, but all other combinations are supported.

Is Flash Pool supported on FAS FlexArray and V-Series system configurations?

Answer: Yes. Flash Pool is supported on FAS FlexArray and V-Series system configurations. However, only SSDs and HDDs in FAS storage shelves can be configured in Flash Pool aggregates. SSDs and HDDs from E-Series systems, and from non-NetApp storage arrays are not supported in Flash Pool aggregates.

What are the default Flash Pool caching policies?

Answer: The default Flash Pool read caching policy is “random_read” and the default write caching policy is “random_write”. Starting with Data ONTAP 8.3, these two policy settings are paired together as the policy “auto” when using the cluster-shell command line interface.

Can the caching policies of a volume be changed from the default Flash Pool caching policies?

Answer: Yes. The default caching policies apply to all volumes when a Flash Pool aggregate is created—including when an existing HDD aggregates is converted into a Flash Pool aggregate—and when new volumes are provisioned on the aggregate. The default read caching policy inserts all blocks that are read randomly into cache, and the default write caching policy inserts random updates to previously written randomly written blocks into cache. In most deployments, the default caching policies are recommended for all volumes. Other policies, such as disabling read or write block insertion, or caching all random writes, are available. To change the caching policies of a volume, use one of the following commands:

- **Node-level**:
```
priority hybrid-cache set <volume_name>
<read-cache>=<caching_policy> <write-cache>=<caching_policy>
```
- **Cluster-shell** (Data ONTAP 8.3 and later releases)
```
volume modify -vserver <vserver name> -volume <volume name>
-caching-policy <caching_policy>
```

Note: The "priority hybrid-cache set" command requires advanced privilege access to the controller node. The command can be used to modify the just the read caching policy, just the write caching policy, or both policies for a volume.

What Flash Pool caching policy options are available?

Answer: The default Flash Pool caching policies are recommended for nearly all deployments. Occasionally, there are circumstances that warrant changing the policy for one or more volumes. Unless you have strong evidence that that a non-default policy will provide better results, use the default policies.

With releases prior to Data ONTAP 8.3, the read and write caching policies of a volume are modified using the node-shell "priority hybrid-cache set" command. This applies to both 7-Mode and clustered Data ONTAP releases. The read cache policy options are: none, random_read, random_read_write, and meta; the default policy is random_read. The write cache policy options are: none and random_write; the default policy is random_write.

Starting with Data ONTAP 8.3, the read and write caching policy options listed above are paired together into combined policies when the cluster-shell "volume modify" command is used.

Starting with Data ONTAP 8.3.1, six additional paired read and write caching options were added based on three new read cache insertion policies: all_read, all_read_random_write, and all.

Table 9 shows the complete set of paired-policy options available in Data ONTAP 8.3.1, the underlying read and write caching policies, and the types of read and write data the each option inserts into cache.

Table 9) Flash Pool caching policy options

			Data Cached per Read Policy[1]				Data Cached per Write Policy[2]
Cluster Shell Policy	Node Shell Read Caching Policy	Node Shell Write Caching Policy	Random Reads	Sequential Reads	Random Writes	Sequential Writes	Random Overwrites
auto (default)[3]	**random_read**	**random_write**	**Yes**	**No**	**No**	**No**	**Yes**
none[3]	none	none	No	No	No	No	No
random_read	random_read	none	Yes	No	Yes	No	No
noread-random_write	none	random_write	No	No	No	No	Yes
meta	meta	none	Meta data only	No	No	No	No
meta-random_write	meta	random_write	Meta data only	No	No	No	Yes
random_read_ write	random_read_ write	none	Yes	No	Yes	No	No
random_read_ write-random_write	random_read_ write	random_write	Yes	No	Yes	No	Yes

			Data Cached per Read Policy[1]				Data Cached per Write Policy[2]
Cluster Shell Policy	Node Shell Read Caching Policy	Node Shell Write Caching Policy	Random Reads	Sequential Reads	Random Writes	Sequential Writes	Random Overwrites
all_read[4]	all_read	none	Yes	Yes	No	No	No
all_read-random_write[4]	all_read	random_write	Yes	Yes	No	No	Yes
all_read_ random_ write[4]	all_read_random _write	none	Yes	Yes	Yes	No	No
all_read_random_ write-random_write[4]	all_read_random _write	random_write	Yes	Yes	Yes	No	Yes
all[4]	all	none	Yes	Yes	Yes	Yes	No
all-random_write[4]	all	random_write	Yes	Yes	Yes	Yes	Yes

Note: (1) A block inserted into cache via any read caching policy is a copy of a block that is stored on HDD.

Note: (2) A block inserted via the random_write write caching policy is the only copy of the data; there is no second copy on HDD.

Note: (3) The cluster shell policies “auto” and “none”, and their constituent node shell policies, can be set with admin-level privilege access. All other policy options require advanced privilege access

Note: (4) These policy options are available starting with Data ONTAP 8.3.1

What are the use cases for using a read caching policy other than “random_read”?

Answer: The default “random_read” read caching policy is recommended for nearly all workloads.

Disabling read caching on one or more volumes with the policy “none” can mitigate performance problems resulting from a severely undersized cache (the working set is much bigger than the cache).

The “meta” policy might deliver better performance for workloads that read mostly meta data and read user data blocks only once.

The “random_read_write” policy may be appropriate for workloads that have a high mix of random reads and a small percentage of random writes that are read, but not cached by the write caching “random_write” policy. With the “random_read_write” policy, blocks are written to HDD before they are inserted into Flash Pool cache via the read policy.

The “all_read” policy inserts all random read and sequential read data into cache. It may be useful when both types of read data are requested repeatedly, however inserting a large amount of sequential read data can displace randomly read data which typically is more beneficial to have in cache.

The “all_read_random_write” policy inserts random write data into read cache, in addition to inserting random and sequential read data. This policy can reduce the latency for reads of just-written random blocks. However, the random write blocks are inserted into cache after they are written to HDDs; this increases workload on the system because two writes to storage – one to HDDs, followed by a second one to SSD cache – are required for every incoming random write block from a host or client.

The “all” policy inserts all types of read and write data into read cache – random and sequential. This means all incoming write blocks are written twice – once to HDDs and a second time to SSD cache –

which significantly increases the workload on the system. The policy can be used to pre-populate read cache with incoming write data, after which the read policy should be changed to stop inserting write data.

What is the use cases for using the write caching policy “none”?

Answer: Disabling write caching on one or more volumes with the policy “none” can mitigate performance problems resulting from a severely undersized cache (the working set is much bigger than the cache).

Can the caching of blocks from a volume in a Flash Pool aggregate be disabled?

Answer: Yes. The per the caching policies and management commands covered above, the “none” policies are used to disable caching for a volume. Disabling caching for a volume stops new blocks from being inserted into the cache, and all reads are served from blocks on HDDs. The disable command does not immediately evict or destage blocks from the cache, however both actions will eventually occur.

5 Performance

How does implementing Flash Pool aggregates affect system performance?

Answer: A Flash Pool aggregate that has a properly-sized cache will lower the average latency of operations for volumes that have workloads with repeat random read, when compared to a HDD-only aggregate. In addition, read cache hits and frequently-updated random writes offload I/O from HDDs to SSDs, which enables the HDDs to handle a higher volume of other types of workload. Converting a HDD-only aggregate into a Flash Pool aggregate can increase volume throughput performance if the HDDs in the aggregate overloaded.

Creating the first Flash Pool aggregate does increase the CPU utilization on multiple cores, however on a controller that has at least 20% headroom (no CPU core is consistently above 80% utilization) controller performance should continue to be satisfactory. Adding more Flash Pool aggregates increases CPU utilization slightly.

How soon does latency improve after a HDD-only aggregate is converted to a Flash Pool aggregate?

Answer: Flash Pool cache needs to be populated with data (the cache needs to be “warmed”) before the benefits of reading blocks from the cache can be realized. The amount of time needed until there is a noticeable reduction in average read latency, or a reduction in HDD utilization, depends on the workload and the cache size. It can take as little as 30 minutes, or as long as several hours, to see significant improvement, and the range of time will be longer to achieve maximum benefits.

Does system performance degrade as the layout of data on the SSDs in Flash Pool cache ages?

Answer: No. After a Flash Pool cache is full to the level where the eviction scanner starts to run, average controller read latency and throughput should reach stable levels, to the extent that the overall workload on the controller is stable.

Are clustered Data ONTAP quality of service (QoS) objectives used to manage the insertion and retention of blocks in Flash Pool cache?

Answer: No. Insertion policies are determined by volume- or aggregate-level cache policies, and data retention (or eviction) is controlled by a modified version of a least recently used (LRU) algorithm.

Can System Performance Modeler (SPM) be used to size system configurations with Flash Pool?

Answer: Yes. An option to size with Flash Pool cache is available in SPM. Sizings can be done that compare system configured with Flash Pool, or Flash Cache, SSDs-only, or HDDs-only. Figure 3 is shows a screen shot of the Detailed Inputs section of the Pick Target Platforms window, where you can

select the desired Flash Pool cache size and SSD capacity to use for the sizing, or let SPM automatically

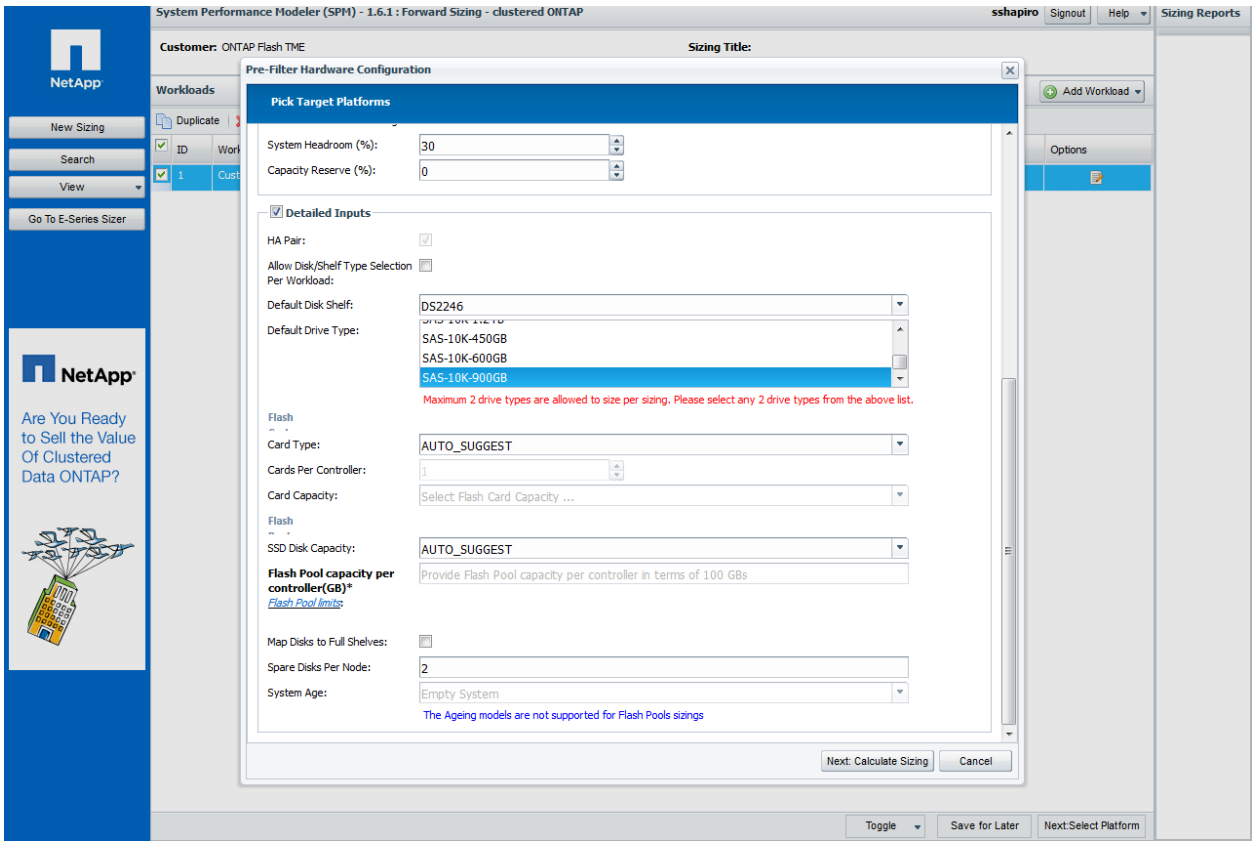

Figure 5) SPM v1.6.1 hardware configuration detailed inputs screen shot.

How can Flash Pool aggregate performance be monitored?

Answer: Graphs of the read and write requests served by a Flash Pool aggregate, and the portion of those requests that are handled by the SSD cache, can be displayed with System Manager 2.2 and later releases. Read and write cache hit rates can be viewed using command line interface commands. The node-level preset command is “stats show -p hybrid_aggr”, and in Data ONTAP 8.3 the cluster-shell command is “statistics cache flash-pool show”. Sample outputs of each command are shown below.

What does the output of “stats show -p hybrid_aggr” preset command look like?

Answer: Figure 4 is a sample output of the `stats show -p hybrid_aggr` command and table 10 describes what the values in each column mean.

Figure 6) Sample output of the “stats show -p hybrid_aggr” command.

Instance	ssd blks used	blks rd cached	blks wrt cached	read ops replaced rate /s	%	write blks replaced rate /s	%	rd cache evict /s	wr cache destage /s	rd cache ins rate /s	wr cache ins rate /s	read hit latency	read miss latency
flash_aggr	52326120	25374159	17624092	9438	88	0	0	5582	5915	2844	3320	5.61	14.68
flash_aggr	52331900	25368605	17616100	8836	88	0	0	4710	6197	3181	3392	5.85	14.84
flash_aggr	52334712	25365170	17608177	8277	88	0	0	3159	5289	3132	3080	6.42	14.05
flash_aggr	52339765	25364531	17608212	9478	89	0	0	4192	6015	3748	3526	5.56	14.36
flash_aggr	52347675	25357944	17605207	10864	89	0	0	4931	6175	1467	1513	4.45	13.14
flash_aggr	52347417	25352786	17595598	10211	89	2187	16	258	31	2762	2219	5.22	14.98
flash_aggr	52350696	25354544	17594364	12657	90	4967	43	0	0	2758	4967	3.47	13.56
flash_aggr	52357194	25357295	17599044	12619	89	6657	51	0	0	2905	6657	3.81	14.28
flash_aggr	52366556	25354910	17600111	11514	89	2134	30	0	0	3791	2134	4.13	13.85
flash_aggr	52370836	25355974	17597018	11704	90	4409	44	0	0	4731	4409	3.85	14.04

Table 10) Flash Pool "stats show -p hybrid_aggr" statistics descriptions.

Statistic	Description
instance	The Flash Pool aggregate that this row of output describes.
ssd blks used	The amount of data currently in the SSD cache. This includes read cache data, write cache data, and Flash Pool metadata.
blks rd cached	The amount of data, in 4KB blocks, currently in the read cache.
blks wrt cached	The amount of data, in 4KB blocks, currently in the write cache.
read ops replaced rate /s	The number of disk read operations replaced per second. This is actual disk I/Os, not 4KB blocks. A read operation can contain more than one 4KB block.
read ops replace rate %	The percentage of disk read operations replaced, on a per-second interval, averaged from one line of output to the next.
write blks replaced rate /s	The number of 4KB blocks written to SSD instead of to HDD.
write blks replaced rate %	The percentage of write blocks written to SSD instead of to HDD, on a per-second interval, averaged from one line of output to the next.
rd cache evict /s	The rate data was evicted from read cache in 4KB blocks per second.
wr cache destage /s	The rate data was destaged from write cache to HDD in 4KB blocks per second.
rd cache ins rate /s	The rate data was inserted into read cache in 4KB blocks per second.
wr cache ins rate /s	The rate data was inserted into write cache in 4KB blocks per second.
read hit latency	The latency, in milliseconds, of read operations that are successfully served from the Flash Pool SSD cache.
read miss latency	The latency, in milliseconds, of read operations that are not served from Flash Pool SSD cache and must read from HDD instead.

What does the output of "statistics cache flash-pool show" command look like?

Answer: Figure 5 is a sample output of the `statistics cache flash-pool show` command. Information about the options available with this command can be found in the *Clustered Data ONTAP 8.3 Command: Manual Page Reference.*

Figure 7) Sample output of the "statistics cache flash-pool" command.

```
cluster1::> statistics cache flash-pool show
cluster1 : 12/31/2013 16:00:04
                              Read Write       Cache
                               Hit   Hit        Used    Read   Write
  Aggregate  Vserver  Volume   (%)   (%)         (%)  Blocks  Blocks Rejects
----------- -------- ------- ---- ----- ----------- ------- ------- -------
      aggr1        - -total-   55    11          84    5342    1675     678
      aggr2      vs1    vol1   81    26          85   11621    4092    1289
```

How can a workload be analyzed to determine if configuring Flash Pool cache will be beneficial?

Answer: The Automated Workload Analyzer (AWA) feature, which is available starting with Data ONTAP 8.2.1, analyzes the read-write workload mix and the percentage of reads and writes that are cacheable, on a per aggregate basis. This makes AWA the best method available to analyze workloads for Flash Pool consideration. More information about AWA is available earlier in this section of the FAQ.

How can a workload be analyzed to determine if implementing Flash Pool would be beneficial, on a FAS system running a release earlier than Data ONTAP 8.2.1?

Answer: Read workload counters are available using the diagnostic mode of Data ONTAP. Diagnostic mode should only be used by, or with the guidance of, NetApp personnel.

First aspect of workload to understand for the purpose of assessing fitness for Flash Pool is the mix of random versus sequential reads. The following two counters do that:

- `stats show readahead:readahead:seq_read_reqs`
- `stats show readahead:readahead:rand_read_reqs`

If a substantial fraction of the read workload is random, the next step is to understand how much of the read workload is being serviced from each type of system resource. That is done by collecting the following statistics:

- `stats show wafl:wafl:read_io_type.cache` (system memory)
- `stats show wafl:wafl:read_io_type.ext_cache` (Flash Cache module)
- `stats show wafl:wafl:read_io_type.disk` (HDD aggregates)
- `stats show wafl:wafl:read_io_type.bamboo_ssd` (SSD aggregates)
- `stats show wafl:wafl:read_io_type.hya_hdd` (Flash Pool aggregate HDDs)
- `stats show wafl:wafl:read_io_type.hya_cache` (Flash Pool aggregate SSD cache)

If most of the workload is being handled by system memory, then adding cache—either Flash Pool SSDs or Flash Cache cards—will provide little or no improvement in overall system performance. On the other hand, if a high percentage of total read I/O is coming from disks, then adding cache will likely improve performance if a large part of the read workload is random (as reflected by the readahead counters).

If the workload includes a significant amount of random writes (for example, a transactional database application), it is important to know whether the write transfer are 16KB or smaller. Only random writes, of that size are candidates to be inserted via the write caching policy for Data ONTAP releases prior to 8.3. The protocol-specific “write size” histogram can be used to display the average block or transfer size. For example, to view the histogram for NFS writes, use the `show stats nfsv3:nfs:nfsv3_write_size_histo` counter.

6 Interoperability

Can NetApp storage efficiency features be used with volumes provisioned on a Flash Pool aggregate?

Answer: Yes. All storage efficiency features are supported for use with volumes provisioned on a Flash Pool aggregate.

Deduplicated and cloned blocks that match the Flash Pool cache insertion policies for a volume are inserted into Flash Pool cache. If a block with the same data already exists in Flash Pool cache, the logical reference is incremented and the duplicate block is not inserted.

Volumes that have data compression enabled are supported on Flash Pool aggregates. Starting with the Data ONTAP 8.3 release, a random read of a block that is within a compression group results in the entire compression group—in compressed format—getting inserted into Flash Pool cache at the same time the compression group is copied into controller memory for decompression. With earlier Data ONTAP releases, blocks within a compression group bypass Flash Pool cache, that is they are never inserted into the cache.

Starting with Data ONTAP 8.3.1, adaptive inline compression of inbound write I/Os is supported as a storage efficiency policy option for volumes in Flash Pool aggregates. Write I/Os that are 8KB in size —and aligned 8KB segments of larger I/Os—that are 50% compressible are written to HDD, or inserted into Flash Pool cache if the overwrite criteria are met. The I/Os are written in compressed fomat in a 4KB block. Enabling adaptive inline compression fits best with 8KB or 16KB random read-write workloads, such as database applications, where reducing storage space for the dataset by up to 50% is important.

How do Snapshot copies work with Flash Pool?

Answer: All of the standard Snapshot® copy functionality works with Flash Pool aggregates. However, if a block that belongs only to a Snapshot copy is read, it is not inserted into cache. Therefore, access to data that exists only in a Snapshot copy, and not also in the active file system, is not accelerated.

How does Flash Pool work with SnapMirror and SnapVault?

Answer: A source volume that is provisioned in a Flash Pool aggregate can be replicated to a destination volume using SnapMirror® or SnapVault®. Both asynchronous and synchronous SnapMirror are supported.

A SnapMirror or SnapVault destination volume can be provisioned in either an all-HDD aggregate or a Flash Pool aggregate. Blocks that reside only in Flash Pool cache—that is, blocks that were inserted via the write caching policy— are replicated from the source volume to the destination volume. If the destination volume is in a Flash Pool aggregate, the write cached blocks from the source volume are copied to HDDs— they are not inserted into the destination aggregate's cache.

Blocks that were inserted into cache via the read caching policy are copies of blocks that are stored on HDDs. Only the block copies on HDDs are replicated from the source volume; blocks from source volumes are never inserted into the destination aggregate's Flash Pool cache as a result of a replication.

Random blocks that are read or written as part of SnapMirror or SnapVault transfer operations are not inserted into Flash Pool cache.

Starting with the Data ONTAP 8.3 release, blocks that are randomly read from a read-only volume—a SnapMirror or SnapVault destination volume— are inserted into Flash Pool cache. With releases earlier than Data ONTAP 8.3, blocks read from read-only volumes are not inserted into cache.

Note: With releases earlier than Data ONTAP 8.3, if the caching blocks read from a SnapMirror or SnapVault destination volume is needed, the read-only volume can be provisioned on a Flash Pool aggregate and then cloned (using FlexClone®). The cloned volume (which is a read-

write volume) can be then used for read-only operations, and randomly read blocks will be inserted into cache. One note of caution: Since the cloned volume is read-write, it is modifiable; any write operation that occurs will modify the cloned volume, but not the origin volume which remains read-only.

An alternative to cloning a read-only volume in Flash Pool aggregate to cache reads is to put the destination SnapMirror and SnapVault volume in a HDD-only aggregate and use Flash Cache on the destination controller.

Are Flash Pool aggregates supported on NetApp Storage Encrypted (NSE) systems?

Answer: Yes, for systems running Data ONTAP 8.2.2 or a later release. Data ONTAP 8.2.2 is the first release that supports an encrypting SSD (800GB). All drives on a NSE HA pair – both HDDs and SSDs – must be encrypting drives. Flash Pool aggregates on NSE HA pairs are supported with clustered Data ONTAP and with Data ONTAP operating in 7-Mode.

Can a Flash Pool aggregate be configured with SyncMirror RAID protection?

Answer: Yes. The RAID groups in each plex of a Flash Pool aggregate that is protected with SyncMirror must be configured identically. The SSD RAID groups in each plex must have the same RAID type, drive count, and drive capacity, and those requirements also apply to the HDD RAID groups in each plex. The SSD and HDD RAID policies can be different from each other, if the system is running Data ONTAP 8.2 or a later release. If the system is running Data ONTAP 8.1.1 or a later 8.1.x release, the SSD and HDD RAID groups must have the same RAID policies.

Are Flash Pool aggregates supported on MetroCluster systems?

Answer: Yes. Flash Pool aggregates are supported on MetroCluster systems running Data ONTAP operating in 7-Mode or clustered Data ONTAP. Special storage subsystem requirements apply when Flash Pool aggregates are configured on a MetroCluster system:

- DS14mk4 FC shelves are supported on 7-Mode MetroCluster systems only; they are not supported on clustered Data ONTAP MetroCluster systems.
- DS4246 and DS2246 shelves can be configured in the same SAS storage stack; DS4243 shelves must be configured in a separate, dedicated stack.
- DS4486 shelves are not supported on MetroCluster systems.
- In a SAS storage stack that is directly connected to controllers (without FibreBridge), the maximum number of SSDs is 48, and the maximum number of shelves (SSD plus HDD) is 4.
- In a SAS storage stack that is connected to controllers via FibreBridge 6500, the maximum number of SSDs is 48, and the maximum number of shelves depends on the number of SSDs on the stack:
    - With 24 or fewer SSDs, the maximum number of shelves per stack is 7.
    - With 25 to 48 SSDs, the maximum number of shelves per stack is 4.

In a 7-Mode MetroCluster, do the SSDs in both plexes of a Flash Pool aggregate count toward the maximum Flash Pool cache size for the system?

Answer: No, the Flash Pool cache of only one plex counts toward the cache maximum for a 7-Mode MetroCluster system, or 7-Mode SyncMirror configured aggregate.

What is the maximum Flash Pool cache size in a 4-node clustered Data ONTAP MetroCluster?

Answer: The total amount of Flash Pool cache that can be configured on a 4-node clustered Data ONTAP MetroCluster—taking into account the cache on all 4 nodes—is equal to the maximum Flash Pool cache size per HA pair as shown in table 12.

A 4-node clustered Data ONTAP MetroCluster consists of two 2-node clusters with each node's aggregates mirrored to a peer on the other cluster. The effect of this is the maximum Flash Pool cache

per HA pair in a clustered Data ONTAP MetroCluster is one-half of the maximum supported for a standard (non-MetroCluster) HA pair.

Can FlexShare be used to prioritize how volumes are cached in a Flash Pool aggregate?

Answer: No. FlexShare® should never with Flash Pool aggregates. Flash Pool caching policies can be modified on a per volume basis to tune cache usage. The node-level “priority hybrid-cache” command, or cluster-shell “volume modify” are used to change Flash Pool read and write caching policies. Both commands are discussed in the Configuration section of this FAQ.

Is SnapLock software supported with Flash Pool aggregates?

Answer: No. A Flash Pool aggregate cannot use the SnapLock® software feature. Likewise, an aggregate with SnapLock volumes may not be converted into a Flash Pool aggregate.

Can the constituents of an Infinite Volume be provisioned on Flash Pool aggregates?

Answer: Yes. However, testing has not shown a performance benefit when provisioning the namespace constituent or data constituents in Flash Pool aggregates.

7 Flash Cache and Flash Pool Maximum Cache Sizes

Table 11) Data ONTAP 8.1.4P1 and later[1] Flash Cache, Flash Pool, and combined maximum cache sizes.

Data ONTAP 8.1.4P1 and Later: Clustered Data ONTAP and Data ONTAP Operating in 7-Mode						
Platform Model	Flash Cache per Node	Flash Cache per HA Pair	Flash Pool per Node	Flash Pool per HA Pair	Combined per Node	Combined per HA Pair
FAS/V6290	8 TiB	16 TiB	6 TiB	12 TiB	8 TiB	16 TiB
FAS/V6280	8 TiB	16 TiB	6 TiB	12 TiB	8 TiB	16 TiB
FAS/V6250	4 TiB	8 TiB	6 TiB	12 TiB	6 TiB	12 TiB
FAS/V6240	3 TiB	6 TiB	6 TiB	12 TiB	6 TiB	12 TiB
FAS/V6220	3 TiB	6 TiB	2 TiB	4 TiB	3 TiB	6 TiB
FAS/V6210	1.5 TiB	3 TiB	2 TiB	4 TiB	2 TiB	4 TiB
FAS/V6080	2 TiB	4 TiB	2 TiB	4 TiB	2 TiB	4 TiB
FAS/V6070	2 TiB	4 TiB	2 TiB	4 TiB	2 TiB	4 TiB
FAS/V6040	1 TiB	2 TiB	1 TiB	2 TiB	1 TiB	2 TiB
FAS/V6030	Not supported	Not supported	1 TiB	2 TiB	Not supported	Not supported
FAS/V3270	1 TiB	2 TiB	1 TiB	2 TiB	1 TiB	2 TiB
FAS/V3250	1 TiB	2 TiB	1 TiB	2 TiB	1 TiB	2 TiB
FAS/V3240	0.5 TiB	1 TiB	0.59 TiB	1.17 TiB	0.59 TiB[1]	1.17 TiB[2]
FAS/V3220	0.5 TiB	1 TiB	0.59 TiB	1.17 TiB	0.59 TiB	1.17 TiB
FAS/V3210	256 GiB[3]	512 GiB[3]	Not supported	Not supported	Not supported	Not supported
FAS/V3170	1 TiB	2 TiB	1 TiB	2 TiB	1 TiB	2 TiB
FAS/V3160	0.5 TiB	1 TiB	0.59 TiB	1.17 TiB	0.59 TiB	1.17 TiB
FAS/V3140	256 GiB[4]	512 GiB[4]	Not supported	Not supported	Not supported	Not supported
FAS/V3070	0.5 TiB	1 TiB	Not supported	Not supported	Not supported	Not supported
FAS2240	Not supported	Not supported	400 GiB[5]		Not supported	Not supported
FAS2220	Not supported	Not supported	400 GiB[5]		Not supported	Not supported

Note: (1) See Hardware Universe for maximum cache values on earlier Data ONTAP 8.1.x releases

Note: (2) When configuring Flash Cache and Flash Pool on a FAS/V3240, only the 256GB Flash Cache card can be used with up to two 100GB SSDs or one 200GB data SSD per node.

Note: (3) Only FAS/V3210 systems using Data ONTAP 8.1.1 (and later) operating in 7-Mode, and that had Flash Cache cards installed before January 2012, are supported.

Note: (4) Only FAS/V3140 systems using Data ONTAP 8.1.1 (and later) operating in 7-Mode, and that had Flash Cache cards installed before January 2012, are supported. The 256GB Flash Cache card is supported with a maximum of 128GiB of data cached per node.

Note: (5) For FAS2240 and FAS2200, the maximum Flash Pool cache capacity per HA pair and per node (including single-controller systems) is the same; that is, the maximum cache capacity can be configured on one node of an HA pair with no cache on the other node.

Table 12) Data ONTAP 8.2.2 and later[1] Flash Cache, Flash Pool, and combined maximum cache sizes.

Data ONTAP 8.2.2 and Later: Clustered Data ONTAP and Data ONTAP Operating in 7-Mode				
Platform Model	Flash Cache per Node	Flash Cache per HA Pair	Flash Pool per Node and per HA Pair[2]	Combined Flash Cache and Flash Pool per HA Pair
FAS8080 EX	8 TiB	16 TiB	36 TiB	36 TiB
FAS/V6290	8 TiB	16 TiB	24 TiB	24 TiB
FAS/V6280	8 TiB	16 TiB	18 TiB	18 TiB
FAS8060	4 TiB	8 TiB	18 TiB	18 TiB
FAS/V6250	4 TiB	8 TiB	18 TiB	18 TiB
FAS/V6240	3 TiB	6 TiB	12 TiB	12 TiB
FAS/V6220	3 TiB	6 TiB	12 TiB	12 TiB
FAS/V6210	1.5 TiB	3 TiB	4 TiB	4 TiB
FAS/V6080	2 TiB	4 TiB	4 TiB	4 TiB
FAS/V6040	1 TiB	2 TiB	2 TiB	2 TiB
FAS8040	2 TiB	4 TiB	12 TiB	12 TiB
FAS/V3270	1 TiB	2 TiB	2 TiB	2 TiB
FAS/V3250	1 TiB	2 TiB	4 TiB	4 TiB
FAS8020	1.5 TiB	3 TiB	6 TiB	6 TiB
FAS/V3240	0.5 TiB	1 TiB	1.17 TiB	1.17 TiB[3]
FAS/V3220	0.5 TiB	1 TiB	1.56 TiB	1.56 TiB
FAS/V3170	1 TiB	2 TiB	2 TiB	2 TiB
FAS/V3160	0.5 TiB	1 TiB	1.17 TiB	1.17 TiB
FAS2554	Not supported	Not supported	4 TiB	Not supported
FAS2552	Not supported	Not supported	4 TiB	Not supported
FAS2520	Not supported	Not supported	4 TiB	Not supported
FAS2240	Not supported	Not supported	800 GiB	Not supported
FAS2220	Not supported	Not supported	800 GiB	Not supported

Note: (1) See Hardware Universe for maximum cache values on earlier Data ONTAP 8.2.x releases

Note: (2) The maximum Flash Pool cache capacity per HA pair and per node (including single-controller systems) is the same. For example, the maximum cache capacity can be configured on one node of an HA pair with no cache configured on the other node.

Note: (3) When configuring Flash Cache and Flash Pool on a FAS/V3240, only the 256GB Flash Cache card can be used with up to two 100GB SSDs or one 200GB data SSD per node. Both nodes in the HA pair should have the same (that is, symmetric) configuration.

Table 13) Data ONTAP 8.3 and later[1] Flash Cache, Flash Pool, and combined maximum cache sizes.

Clustered Data ONTAP 8.3				
Platform Model	Flash Cache per Node	Flash Cache per HA Pair	Flash Pool per Node and per HA Pair[2]	Combined Flash Cache and Flash Pool per HA Pair
FAS8080 EX	12 TiB	24 TiB	144 TiB	144 TiB
FAS/V6290	8 TiB	16 TiB	96 TiB	96 TiB
FAS/V6280	8 TiB	16 TiB	72 TiB	72 TiB
FAS8060	4 TiB	8 TiB	72 TiB	72 TiB
FAS/V6250	4 TiB	8 TiB	72 TiB	72 TiB
FAS/V6240	3 TiB	6 TiB	48 TiB	48 TiB
FAS/V6220	3 TiB	6 TiB	48 TiB	48 TiB
FAS/V6210	1.5 TiB	3 TiB	16 TiB	16 TiB
FAS8040	2 TiB	4 TiB	48 TiB	48 TiB
FAS/V3270	1 TiB	2 TiB	8 TiB	8 TiB
FAS/V3250	1 TiB	2 TiB	16 TiB	16 TiB
FAS8020	1.5 TiB	3 TiB	24 TiB	24 TiB
FAS/V3240	0.5 TiB	1 TiB	3.13 TiB	3.13 TiB
FAS/V3220	0.5 TiB	1 TiB	6.25 TiB[3]	6.25 TiB[3]
FAS2554	Not supported	Not supported	16 TiB	Not supported
FAS2552	Not supported	Not supported	16 TiB	Not supported
FAS2520	Not supported	Not supported	16 TiB	Not supported
FAS2240	Not supported	Not supported	800 GiB	Not supported
FAS2220	Not supported	Not supported	800 GiB	Not supported

Note: (1) See Hardware Universe for maximum cache values on earlier Data ONTAP 8.2.x releases

Note: (2)The maximum Flash Pool cache capacity per HA pair and per node (including single-controller systems) is the same. For example, the maximum cache capacity can be configured on one node of an HA pair with no cache configured on the other node.

Note: (3) Maximum Flash Pool and combined Flash Cache plus Flash Pool cache size for the FAS3220 in Data ONTAP 8.3 RC1 and RC2 releases is 1.56 TiB. Those maximum cache sizes are 6.4 TiB starting with the Data ONTAP 8.3 GA release.

8 Additional Resources

- Technical Report TR-4070: NetApp Flash Pool Design and Implementation Guide
- NetApp Flash Pool SE Technical Presentation
- Storage Subsystem Technical FAQ
- NetApp Hardware Universe
- System Performance Modeler (SPM)

9 Contact us

Do you have a question that you think we should include in this FAQ?

Ask us at docfeedback@netapp.com.

Include Flash Pool Technical FAQ in the subject line.

Refer to the Interoperability Matrix Tool (IMT) on the NetApp Support site to validate that the exact product and feature versions described in this document are supported for your specific environment. The NetApp IMT defines the product components and versions that can be used to construct configurations that are supported by NetApp. Specific results depend on each customer's installation in accordance with published specifications.

Copyright Information

Copyright © 1994–2015 NetApp, Inc. All rights reserved. Printed in the U.S. No part of this document covered by copyright may be reproduced in any form or by any means—graphic, electronic, or mechanical, including photocopying, recording, taping, or storage in an electronic retrieval system—without prior written permission of the copyright owner.

Software derived from copyrighted NetApp material is subject to the following license and disclaimer:

THIS SOFTWARE IS PROVIDED BY NETAPP "AS IS" AND WITHOUT ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE, WHICH ARE HEREBY DISCLAIMED. IN NO EVENT SHALL NETAPP BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

NetApp reserves the right to change any products described herein at any time, and without notice. NetApp assumes no responsibility or liability arising from the use of products described herein, except as expressly agreed to in writing by NetApp. The use or purchase of this product does not convey a license under any patent rights, trademark rights, or any other intellectual property rights of NetApp.

The product described in this manual may be protected by one or more U.S. patents, foreign patents, or pending applications.

RESTRICTED RIGHTS LEGEND: Use, duplication, or disclosure by the government is subject to restrictions as set forth in subparagraph (c)(1)(ii) of the Rights in Technical Data and Computer Software clause at DFARS 252.277-7103 (October 1988) and FAR 52-227-19 (June 1987).

Trademark Information

NetApp, the NetApp logo, Go Further, Faster, ASUP, AutoSupport, Campaign Express, Cloud ONTAP, Customer Fitness, Data ONTAP, DataMotion, Fitness, Flash Accel, Flash Cache, Flash Pool, FlashRay, FlexArray, FlexCache, FlexClone, FlexPod, FlexScale, FlexShare, FlexVol, FPolicy, GetSuccessful, LockVault, Manage ONTAP, Mars, MetroCluster, MultiStore, NetApp Insight, OnCommand, ONTAP, ONTAPI, RAID DP, SANtricity, SecureShare, Simplicity, Simulate ONTAP, Snap Creator, SnapCopy, SnapDrive, SnapIntegrator, SnapLock, SnapManager, SnapMirror, SnapMover, SnapProtect, SnapRestore, Snapshot, SnapValidator, SnapVault, StorageGRID, Tech OnTap, Unbound Cloud, and WAFL are trademarks or registered trademarks of NetApp, Inc., in the United States and/or other countries. A current list of NetApp trademarks is available on the Web at http://www.netapp.com/us/legal/netapptmlist.aspx.

Cisco and the Cisco logo are trademarks of Cisco in the U.S. and other countries. All other brands or products are trademarks or registered trademarks of their respective holders and should be treated as such.

NetApp

www.netapp.com