JetStor®
Storage. Solutions. Support.

How to Deploy SSD Cache in JetStor Unified Storage

JetStor SAS 716U
JetStor SAS 724U
JetStor SAS 716U 10G
JetStor SAS 724U 10G

Version 1.0
September 2012

Copyright

Trademarks

Advanced Computer & Network Corporation
5001 Baum Blvd., Ste 680
Pittsburgh, PA 15213

Tel: 412-683-9010
Fax: 412-683-9070

Email: sales@acnc.com
Website: www.acnc.com

Introduction

JetStor unified storage system is a powerful yet simple storage system that provides consolidated solutions for both file sharing access and block level access at the same time. Powerful and efficient storage pool design includes enterprise storage functions such as thin provisioning, deduplication, compression and SSD caching. In this article we will explain how to deploy SSD cache and where SSD cache will make the most difference in performance.

In JetStor unified storage system, SSD cache is dedicated to each storage pool and only SSD drive can be assigned as SSD cache. In JetStor ZFS technology, there are two types of SSD cache:

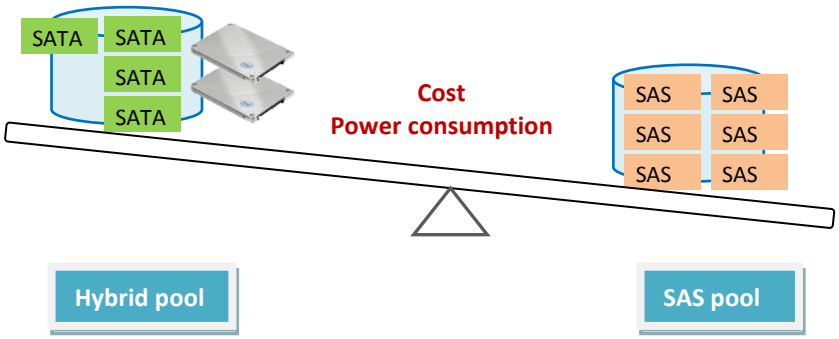
- Read cache: it is also called L2ARC (Layer 2 Adaptive Replacement Cache).
- Write cache: it is also called Log Device and is used by ZIL (ZFS intent log).

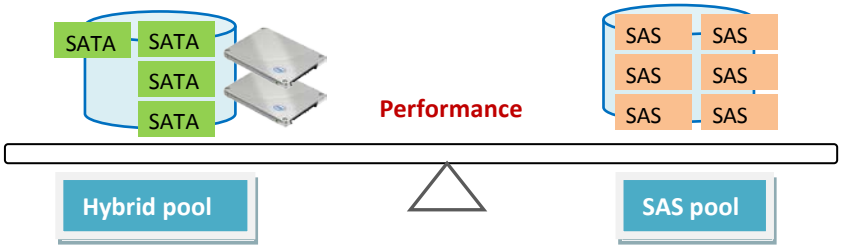
From data I/O viewpoint, SSD cache sits between main memory (DRAM) and the hard drives and acts as a buffer to keep frequently accessed data that are not in the main memory. Because the price per GB for SSD drives keeps dropping, using fast SSD cache with conventional SATA drives can create a hybrid storage pool solution that is cost-effective, power-efficient and high-performing by capitalizing the strengths of fast SSD drives. However, Adding SSD cache to the storage pool doesn't always guarantee improvement in performance especially when the storage pool has already been populated with high performance hard drives such as 15K rpm SAS, 10K rpm SAS and even SSD drives or the storage pool contains a large number of hard drives for example over 20 drives.

Below is the price per GB comparison of different storage media from the market.

SAS 15K rpm	SATA SSD (enterprise grade)	SATA 7.2K rpm
$3/GB	$1.5/GB	$0.12/GB

For example, compare a 6 SAS drives (300GB each) solution with a 4 SATA drives (1TB each) plus 2 SSD drives (240GB each) hybrid solution. Full SAS solution costs you $5400, while hybrid solution costs you about $1200. Full SAS solution gives you 1.5TB (RAID5) capacity and the hybrid solution offers you 3TB (RAID5) capacity.





By using read-optimized SSD and write-optimized SSD, the overall performance of hybrid SATA storage pool can be close to or better than that of a SAS storage pool in two distinct ways - synchronous write applications and deduplication. The extra benefits are lower costs, more capacity, and lower power consumption if compared with full SAS solution.

The following sections will show you how to find the proper SSD drives and deploy them to strike the right balance of cost and performance. We will show you how to deploy read cache and write cache to benefit the most from them.

Read cache (L2ARC) with deduplication

Read cache is specifically helpful when deduplication is turned on. It is because when using deduplication, there will be frequent access to DDT (deduplication table) where it stores the index of data blocks being deduplcated. Without L2ARC, DDT resides in main memory with limited space. DDT size keeps growing as more data being deduplicated and finally it gets pushed out from main memory to the slow disk pool. This is where performance becomes worse.

Let's give you an idea of how much performance drop it will be when deduplication function is enabled. Three SATA drives in RAID 5 are used in this test. Chunks of data are copied starting from 10GB and accumulating to 160GB in total. Figure 1 shows that the performance drops as more data is processed by deduplication.

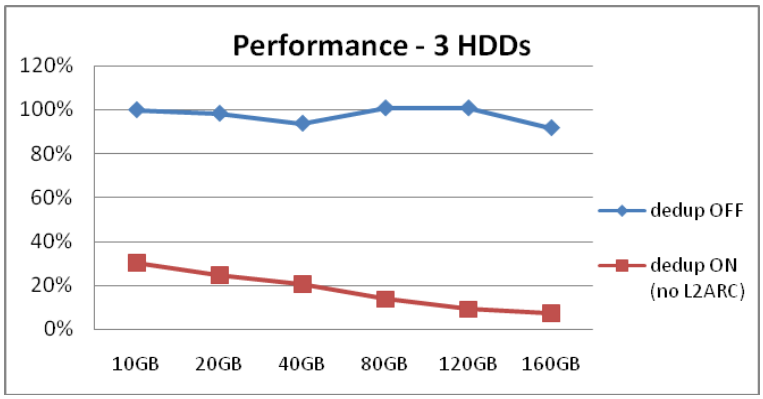


Figure 1 Write files from 10GB to 160GB with deduplication OFF and ON

With L2ARC added, JetStor ZFS technology allows to move DDT to L2ARC read cache and really cut back the access time of fetching data from the disks. If you want to use read cache to improve performance, we suggest using read cache in deduplication applications such as data backup. Let's see how much performance can be improved by adding read cache as L2ARC. Figure 2 shows that we continue the test in figure 1 by copying more files with L2ARC read cache added. JetStor ZFS technology supports up to 4 SSD drives as read cache. Multiple SSD drives can form as RAID 0 to add extra performance to L2ARC read cache.

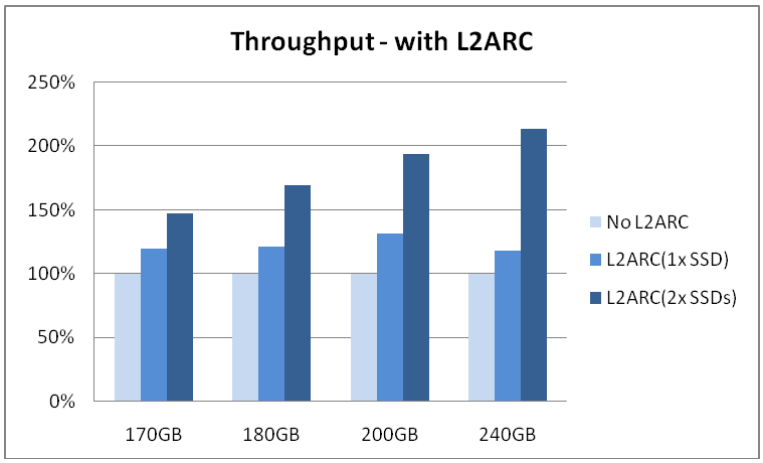


Figure 2 The faster the L2ARC, the better the deduplication performance

Compare hybrid pool solution and SAS pool solution in deduplication

In hybrid pool, we use 4x SATA drives, 1x SSD drive as read cache, and 1x SSD drive as write cache. In SAS pool, we use 6x SAS 15K rpm drives. Figure 3 below shows that hybrid pool can match the performance of SAS pool in deduplication application.

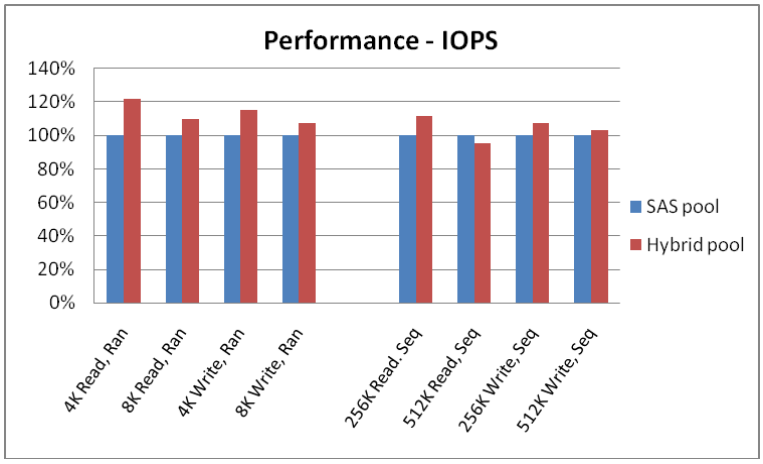


Figure 3 Hybrid pool & SAS pool with deduplication

Write cache (ZIL) with synchronous write

Synchronous write is often used in database and NFS applications where data security and consistency is top priority. JetStor ZFS technology uses ZFS intent log (ZIL) - a logging mechanism to conform to POSIX file system requirements for synchronous writes. ZFS intent log saves transaction records of system calls that change the file system with enough information to replay them. JetStor ZFS technology provides the use of separate intent log device where synchronous writes can be quickly written and acknowledged to the system calls before the data is written to the disk storage pool. By using a fast SSD drive as a ZFS log device, the write cache can improve synchronous write performance significantly.

Figure 4 shows both random write and sequential write performance can be improved over 2 to 3 times by using 1x SSD drive as write cache.

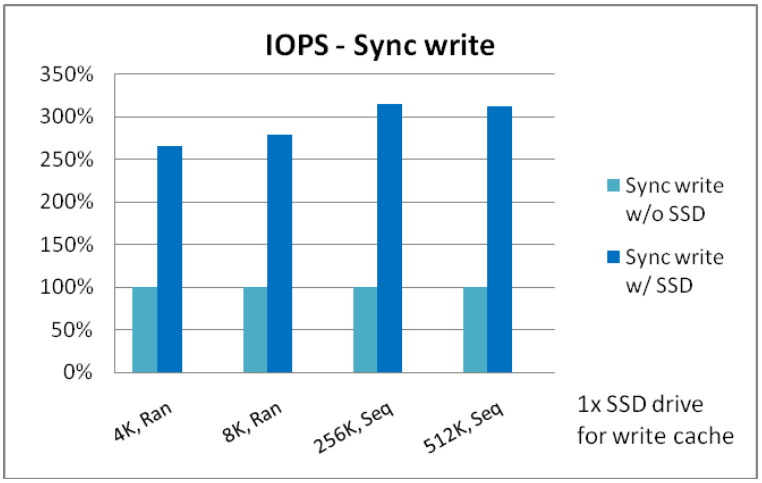


Figure 4 Compare write performance using 1x SSD drive as write cache

Compare hybrid pool solution and SAS pool solution in synchronous write

In hybrid pool, we use 4x SATA drives, 1x SSD drive as read cache, and 1x SSD drive as write cache. In SAS pool, we use 6x SAS 15K rpm drives. Figure 5 below shows that hybrid pool can even outperform SAS pool by at least double.

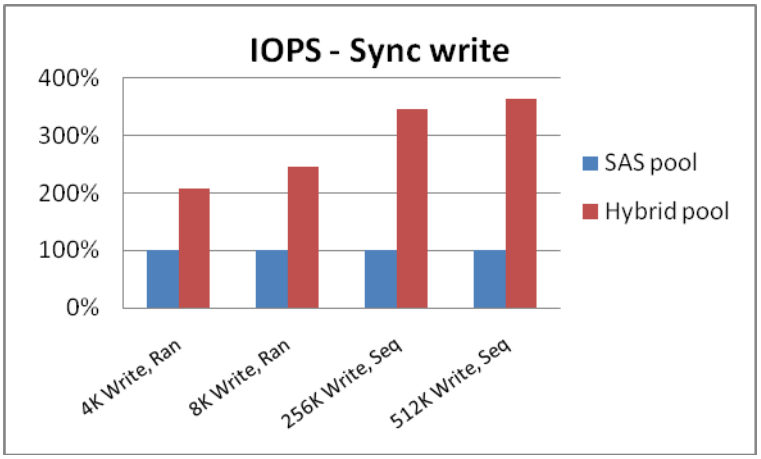


Figure 5 Hybrid pool and SAS pool with synchronous write

Because ZIL write cache contains important transaction records of system calls. If SSD drive for write cache fails, data will be lost but the storage pool still maintains its consistency and integrity for sure. We recommend using SLC type SSD drive as write cache. If you can spare two SSD drives for write cache, you may configure them as RAID 1 mirror for better data protection.

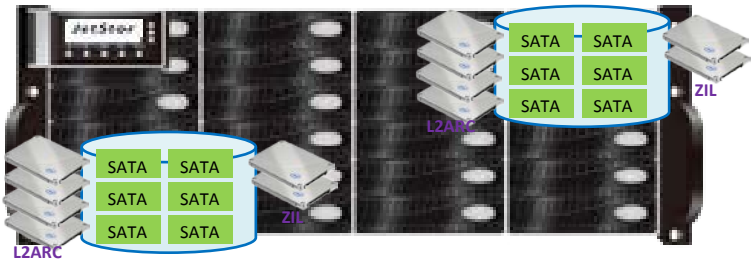
SSD cache rules and deduplication size limit

Starting with FW1.0.3 certain rules and limits need to be followed in order to set up SSD cache properly.

Maximum number of SSD drives

Each storage pool can have read cache (L2ARC) up to 4 SSD drives and write cache (ZIL) up to 2 SSD drives. A total of 6 SSD drives for each storage pool. For write cache, the two SSD drives can be configured as RAID 0 for better performance or as RAID 1 mirror for better data protection.

The number limit is applied to individual pool ONLY. There is no SSD drive limit on the storage system. You may create many storage pools using JBODs and apply 6 SSD drives to each pool as SSD cache.



Deduplication limit

JetStor unified storage system provides block level, inline, synchronous deduplication. A redundant block will be removed and replaced by an index in the deduplication table (called DDT), which keeps track of all the redundancy information of a specific storage pool. The DDT resides in main memory when there is no SSD read cache deployed. The more deduplicated data, the bigger the DDT becomes. After a certain threshold is reached, the DDT will be moved from memory to storage pool and that's when the performance degenerates very severely. To avoid unbearable performance drop, when deduplication size limit is reached, deduplication function on ALL pools of the system will be forced to turn OFF automatically unless SSD read cache is added to individual pool to unlock deduplication function of that specific storage pool.

Memory size	Deduplication size limit
8GB	137GB
16GB	371GB

Conclusion

In JetStor ZFS technology, SSD cache function (L2ARC & ZIL) is purposely built to serve for deduplication and synchronous write applications. It enables the use of hybrid storage pool that combines high capacity, low cost SATA drives with low I/O latency, power saving SSD drives to form a cost effective solution that can competes with pure SAS solution in price, power consumption and performance. L2ARC read cache is typically suitable for deduplication applications such as data backup. ZIL write cache can give you an edge in synchronous write applications such as database and NFS.

SSD cache is not a general-purpose tool that can boost performance under all circumstances. Please take precaution to apply it to proper applications to benefit from this evolving technology.

Reference

1. All tests are conducted using JetStor SAS 724U, 7.2K rpm SATA drive, 15K rpm SAS drive and enterprise SSD.
2. Can flash memory become the foundation for a new tier in storage hierarchy? By Adam Leventhal, COMMUNICATIONS OF THE ACM, No 7, Vol 51, July 2008