Synology SSD Cache White Paper for DSM 7.1

Last updated: Nov 22, 2022

Table of Contents

Introduction

Data storage performance challenges

High information system productivity demands low latency, and the requirements for input/output (I/O) latency are most stringent when running mission-critical applications. For any IT deployment, the greatest challenge is achieving a balance between low latency, high efficiency, and optimal system utilization rate.

The degree of latency in a storage system is determined by two factors: workload patterns and storage media capabilities. Most business applications (e.g., OLTP databases or email services) involve random **input/output operations per second (IOPS)**, which access non-contiguous data stored on system drives. Since these bits of data are not physically close enough to one another, the system generates numerous search operation processes, thereby increasing latency.

To address the high latency caused by random IOPS workloads, the best way is to use an all-flash, solid-state drive (SSD) storage solution. SSDs utilize NAND flash memory, which delivers high read/write performance for random data requests. However, due to its high deployment costs, sustaining an all-flash storage environment can be challenging.

SSD cache as a solution

SSD cache offers a high-performance solution to latency and storage cost challenges. It increases read and write speeds by only using a limited number of SSDs.

On read operations, statistically, only a small percentage of data in any particular storage system is accessed frequently. As a result, system performance may be increased by storing frequently accessed data on the SSD cache to establish a read buffer while keeping the total cost of ownership to a minimum.

In terms of write operations, small, random write operations are common in enterprise workloads. SSD read-write cache can speed up the performance of volumes and LUNs, reduce the latency of random write operations, and drastically decrease the impact on data transfer performance.

Refer to the list of supported models for details.

Data Operations and SSD Caching

Data read/write operations

Typically, when a read request is received, the servers check if the relevant data are stored in the system memory cache known as **random access memory (RAM)**. As shown in Figure 1, RAM saves the most recently accessed information. If the requested data are not found on the RAM, the system moves on to access the data on drives. Since the RAM size is significantly limited compared to the working data set, most retrieval requests must be read from the drives, which can result in increased latency.

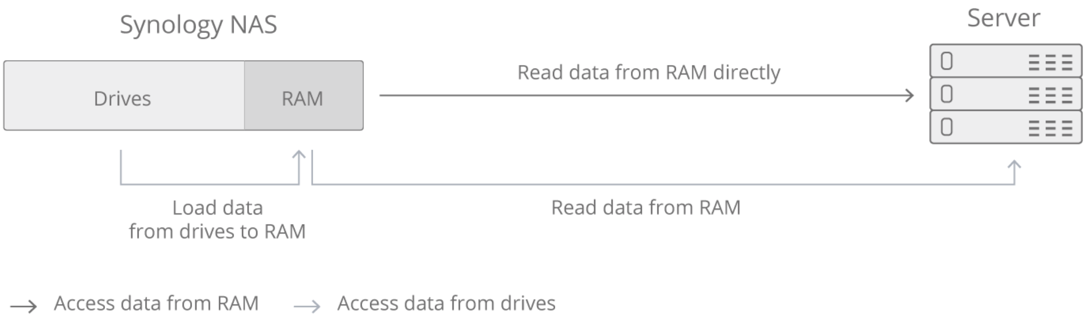


Figure 1 - Accessing data from the RAM and drives

In most applications, there are noticeable patterns in data retrieval and workload due to the specific read/write characteristics of the application's behavior. For instance, in an Online Transactional Processing (OLTP) database workload, some tables in the database are accessed more frequently than others. These frequently accessed data are referred to as **hot data**.

Of this hot data subset, the most recent data has an even higher probability of being frequently accessed, which raises the need for quick retrieval.

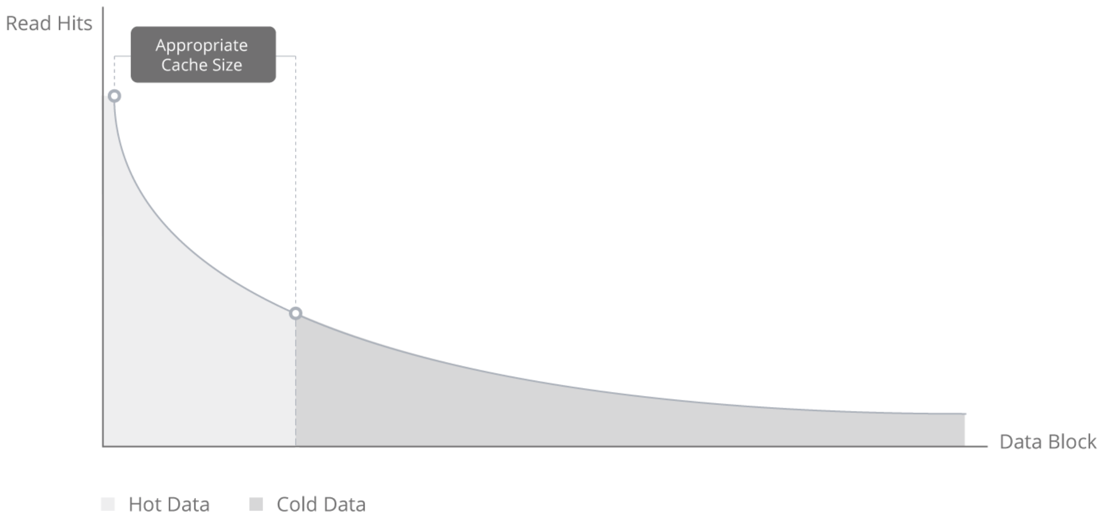


Figure 2 - Hot data, cold data, and appropriate cache size

As shown in Figure 2, even though hot data is only a portion of the whole data set, they have the most intensive read/write requests. A small number of SSDs can then be used to cache all hot data. By leveraging the superior read/write capabilities of SSDs, the system performance will significantly improve.

Synology SSD caching technology

Synology offers read-only cache and read-write cache. While a read-only cache can be created with a single SSD, a read-write cache must be created with at least two SSDs of the same type (i.e., NVMe or SATA).

The capacity of an SSD doesn't have to be wholly reserved for one SSD cache. A single set of SSDs can be used to create multiple caches through **SSD cache group**. Since each cache is mounted to a different volume, SSD cache group makes multi-volume caching possible. This means that administrators can flexibly allocate and distribute SSD storage capacity to ensure maximum, efficient use of available SSDs.

SSD cache can also be equipped with different RAID types to set fault tolerance and improve the random read-write performance. The RAID type set on an SSD cache group is applied to all its SSD caches. For example, if the RAID type of the cache group is set to RAID 5, each SSD cache has fault tolerance for one SSD failure. For detailed information, refer to Choose a Cache RAID Type.

Once a volume is equipped with SSD cache, any LUN or shared folder on the volume will benefit from the increased performance.

SSD Caching Methods

SSD cache creation and population

Once the SSD cache is created, data requested from the drives will be continuously copied and stored onto the SSD cache. A data block mapping table will be created in the system RAM to record which data blocks are on the SSD cache. Therefore, the SSD cache size and system RAM size are proportionally-correlated.

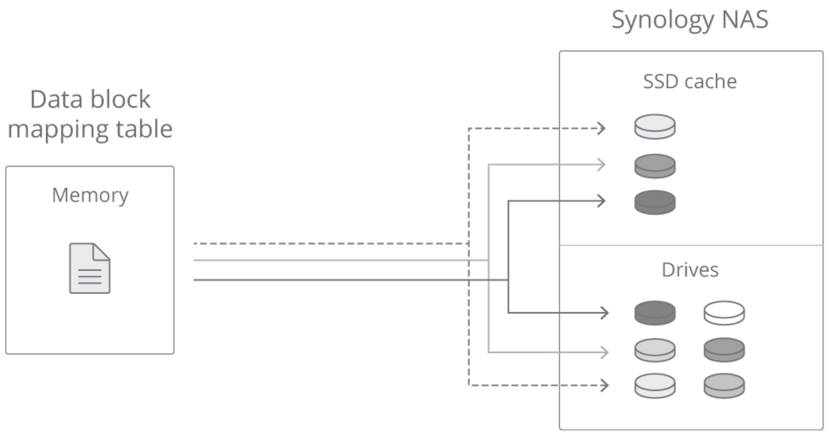


Figure 3 - Data block mapping table

If more requests for the same data are generated, read operations will be performed on the SSDs, which is referred to as a **cache hit**. The read speed enhances when data are retrieved from the cache.

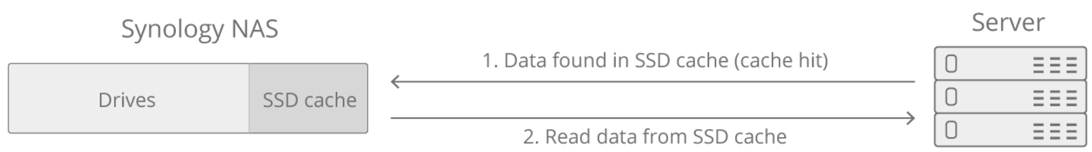


Figure 4 - Data access from cache as "cache hit"

If the data are not found on the SSD cache, it is referred to as a **cache miss**. After a cache miss, read operations will be performed on drives. Once the data is retrieved, the SSD cache will save a copy of the data to speed up future requests.

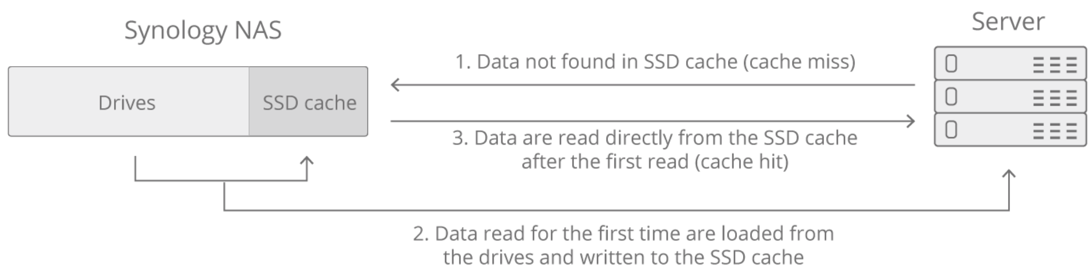


Figure 5 - Data access from drives as "cache miss"

The SSD cache controls all data at the block-level. For example, when reading a 400 KB chunk of data from a 4 GB file, only the most-relevant 400 KB will be accessed. If this 400 KB of data is absent, the system will read the data from the drives and transfer the information to the SSD cache.

Cache hit rate and warm-up period

The **hit rate** for read operations is the percentage of cache hits each time the SSD cache performs data requests. A higher hit rate means fewer read operations on drives, resulting in a faster response time and lower latency. A lower hit rate indicates that the majority of data are being read from the drives, and the response time is similar to that of reading data from volumes without an SSD cache.

Since the SSD cache is empty at the start, almost every data reading operation will cause a cache miss. Meanwhile, copies of data will be continuously added to the SSD cache. This period is called **warm-up** and is mainly composed of copy operations. Figure 6 illustrates how the hit rate increases throughout the warm-up period. A higher cache hit rate indicates that the SSD cache is being fully utilized.

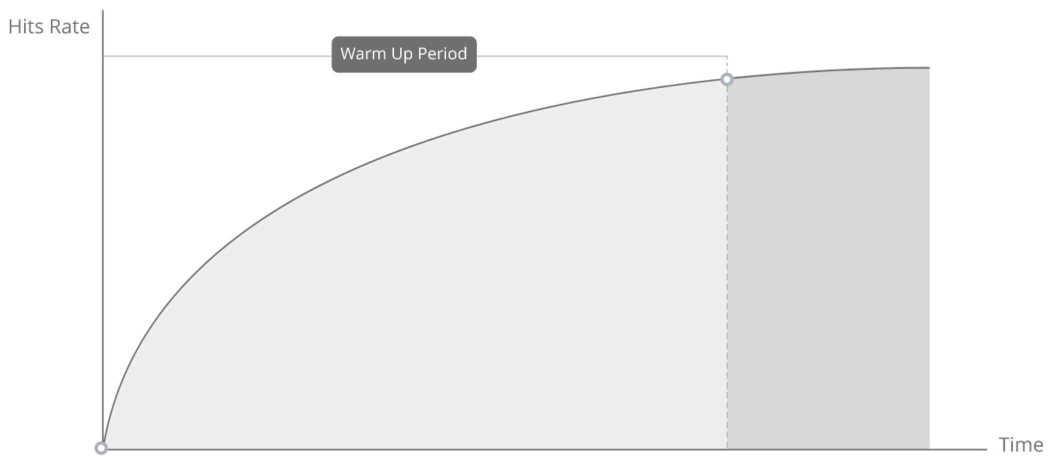


Figure 6 - Warm-up period and cache hit rate

The warm-up period can also occur if there are drastic changes to the frequently accessed data. This means the new “hot” data will take longer to retrieve at first.

SSD write-back caching

Synology SSD cache uses a write-back mechanism to accelerate write operations. As shown in Figure 7, data are first written to the SSD cache. Once complete, the write will be acknowledged (ack write) so that further data can be written to the SSD cache.

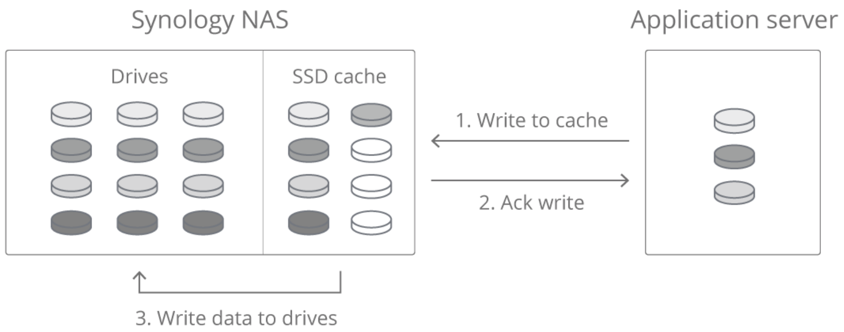


Figure 7 - Write-back caching process

Eventually, data on the SSD cache will be written back to the drive to make space for more frequently accessed data.

SSD cache implements a **Least Recently Used (LRU)** replacement mechanism to determine which data blocks should remain and which can be overwritten by new ones. LRU bases its decisions on the two following factors: the access sequence and the latest access time. Only the more frequently accessed (i.e., hot) data will remain on the SSD cache.

Automatic protection mechanism

In the event of cache RAID degration, the cache becomes a high-risk zone for data storage. In this case, the system will trigger the **automatic protection mechanism**, an automated write-back mechanism, to protect the data.

The automatic protection mechanism reduces the risk of data loss by halting all write operations on the degraded SSD cache. During this period, the system bypasses SSD cache and routes incoming write operations directly to drives. Data in the SSD cache are automatically sorted in sequential order and are fast-tracked back to the drives. By transferring the data to a safe haven in the shortest time possible, it keeps the chance of data loss to a minimum.

Recommended Deployment

SSD cache size

The recommended SSD cache size should be just enough to cover the size of frequently accessed data. For example, in a 2 TB file server with 100 GB of frequently accessed data, the recommended cache size would be slightly above 100 GB.

With that in mind, configuring more cache space than required does not guarantee better cache performance. Continuing with the example above, increasing the SSD cache size to 500 GB for 100 GB of hot data will not result in significant performance increase. Excess cache space will only be used to store cold data. Hence, using SSD cache space slightly larger than the actual hot data size is sufficient.

Memory requirements

RAM size is another factor that determines the SSD cache size a system can produce. A data block mapping table resides in the RAM module to track the SSD cache data. The larger the cache size, the more system memory required.

In general, every 1 GB of SSD cache requires approximately 400 KB of system memory. However, to maintain the stability of Synology NAS, the system will only use up to a quarter (25%) of the pre-installed system memory for SSD cache creation.

This is with exception to Synology NAS models with an Alpine CPU, where the total size of the SSD cache is limited to 930 GB.

Learn about which CPU your NAS has.

SSD Cache Advisor

The aggregate size of hot data and appropriate cache size can be determined with the help of Storage Manager's **SSD Cache Advisor**. SSD Cache Advisor looks at your data usage pattern (such as the size and number of recently-accessed files) for a period of time before recommending a suitable size for an SSD read-write cache.'

Consider the following example:

The amount of frequently accessed files on a selected volume is around 100 GB per day for one month, except for one day where there is an outburst of 300 GB. In this case, SSD Cache Advisor will still consider your average data usage to be 100 GB and recommend creating a read-write cache based on this size.

The initial analysis takes at least 7 days and automatically ends after 30 days. Afterwards, SSD Cache Advisor will recommend an appropriate cache size based on its analysis.

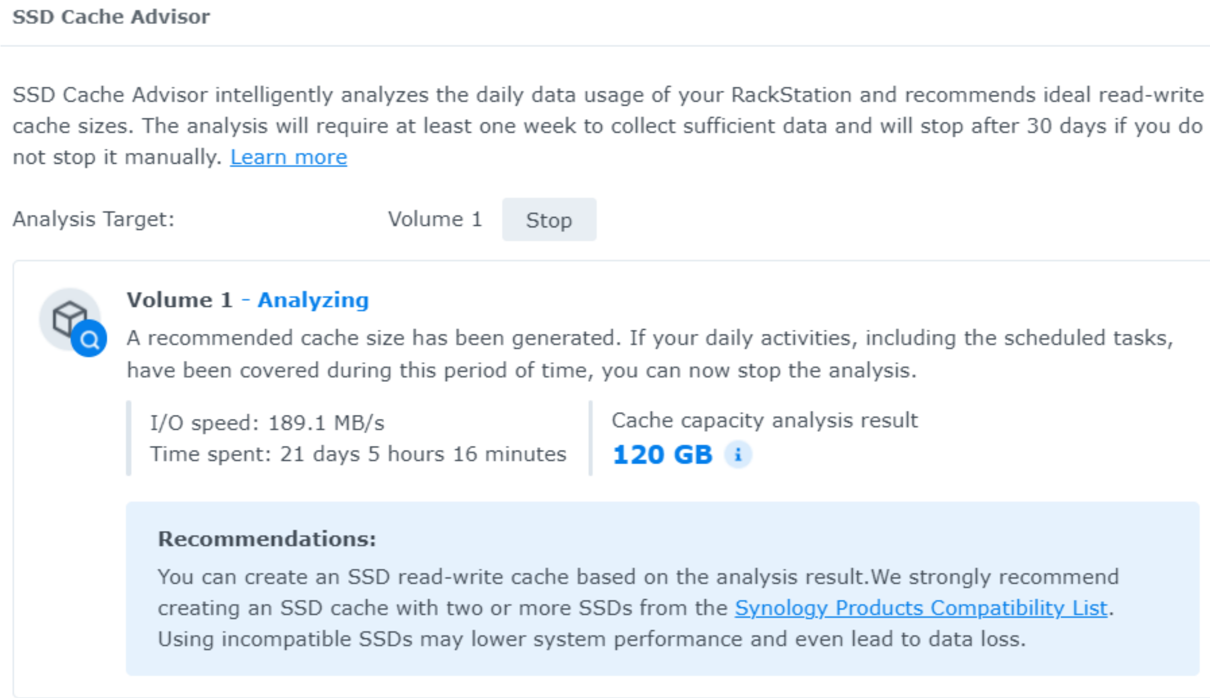


Figure 8 - DSM's built-in SSD Cache Advisor

Btrfs metadata pin feature

Increasing the cache hit rate on metadata reduces indexing time and improves the overall performance of read/write operations. This is because all file operations start with locating a file's metadata.

You can achieve this by enabling the "pin all metadata" feature. This allows SSD cache to store the metadata of the Btrfs volume where it is mounted. The system is then able to access metadata more quickly, reducing indexing time and improving the performance of read/write operations. Pinning all Btrfs metadata to an SSD cache enhances small-file access efficiency in Active Backup, Hyper Backup, and snapshot operations.

Services, protocols, and applications

SSD cache enhances performance where read/write operations require **frequent access to randomly placed small blocks of data**.

SSD cache is likely to boost performance if you use the following:

- CIFS/SMB and iSCSI
- Virtualization applications
- Cloud computing, OLTP databases, mail, and file services

SSD cache will not improve performance in scenarios involving **sequential access patterns**.

Performance gains from SSD cache are expected to be minimal if you use the following:

- Single channel HD video streaming
- Surveillance video recording

To learn more about which applications can benefit from using SSD cache, refer to SSD cache considerations.

Performance Evaluation

Testing environment

The objective of this test was to compare the performance between an environment with SSD cache and an environment without SSD cache.

In this scenario, each test was performed three times, each with a different set of configurations. Each test included a round of 100% random read and another round of 100% random write. The third round was implemented to represent a more realistic scenario, since typical workload patterns may involve simultaneous 70% read and 30% write operations.

Each test used a device with pre-populated hot data on the SSD caches. Excluding the warm-up period from the testing process helps to show the actual performance of the SSD caches and keep the results as consistent as possible. These tests were implemented in a controlled environment using the specific configurations of our testing facility. Performance results may vary in other conditions.

Testing configurations

Tests were conducted with the following configurations:

Test A Configurations

Item	Description
Server model	SA3600
Hard drive	ST6000NE0023 x 12
Network Adapter Card	E10G18-T2 x 1
NVMe Adapter Card	M2D20
SSD model	Synology SNV3500-400G M.2 NVMe SSD x 2
SSD cache RAID type	RAID 1
Volume RAID type	RAID 5
System memory	D4RD-2666-16G x 2

Test B Configurations

Item	Description
Server model	SA3600
Hard drive	ST6000NE0023 x 10
Network Adapter Card	E10G18-T2 x 1
SSD model	Synology SAT5200-960G SSD x 2
SSD cache RAID type	RAID 1
Volume RAID type	RAID 5
System memory	D4RD-2666-16G x 2

We used this opportunity to demonstrate the performance difference of NVMe and SATA SSDs. In Test A, we used NVMe SSD. In Test B, we used SATA SSD. Due to the SA3600's form factor, the NVMe SSDs are placed on the PCIe expansion slots. On the other hand, the SATA SSDs must occupy two drive slots, which reduces the number of accessible hard disk spaces from 12 to 10.

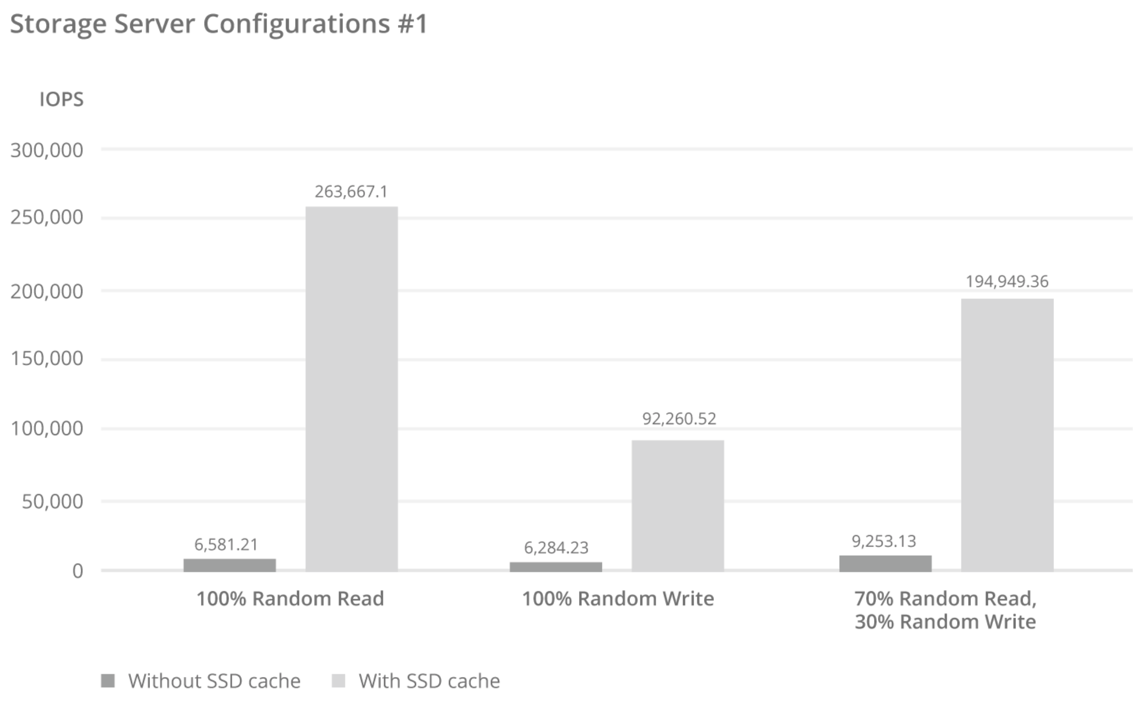
Results and analysis

IOPS is a standard performance benchmark for storage devices. It measures how many read/write operations a storage device (in this case, an SSD cache) can handle every second. Generally, the higher the number of IOPS number, the faster the data can be retrieved.

The achieved IOPS, including read, write, and read/write operations, are shown in the following figures:

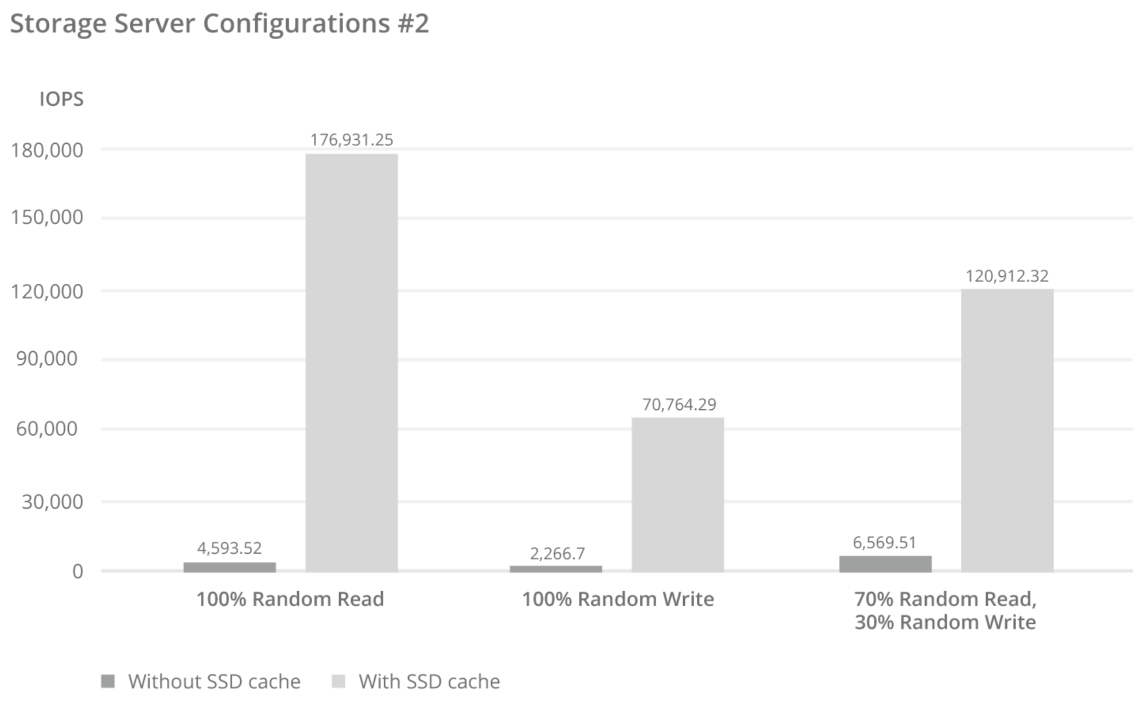
Test A

Round	Without SSD cache	With SSD cache
100% Random Read	6,581.21	263,667.10
100% Random Write	6,284.23	92,620.52
70% Random Read, 30% Random Write	9,253.13	194,949.36



Test B

Round	Without SSD cache	With SSD cache
100% Random Read	4,593.52	176,931.25
100% Random Write	2,266.7	70,764.29
70% Random Read, 30% Random Write	6,569.51	120,912.32



Analysis

As shown in the tables and charts above, the IOPS for read, write, and read/write performance significantly improves in environments that use SSD cache. This shows that the speed at which users can retrieve their frequently accessed data is increased by 30x when using SSD cache, even without the use of an all-flash storage environment to boost IOPS capabilities.

Conclusion

In this document, we discussed why SSD cache is an economical choice for increasing server performance. In Synology's tests, the random read performance increased by 30x when the servers were equipped with Synology SSD cache. All in all, SSD cache is a simple and affordable way to build a storage system with low latency and high efficiency.

Synology SSD cache includes the following benefits:

- **SSD Cache Advisor**: Helps to recommend an ideal SSD cache capacity based on your actual usage.
- **Automatic protection mechanism**: Keeps your data protected in case of SSD malfunction.
- **SSD cache group**: Grants you the convenience of flexibly allocating and distributing SSD storage capacity.