Maximizing the Benefits from Flash In Your SAN Storage Systems

Dot Hill Storage Hybrid Arrays integrate flash and

HDDs in Optimal Configurations

INTRODUCTION – EXECUTIVE SUMMARY

Virtual server environments place demanding workloads on storage systems. VMware and Hyper-V deployments have become common in many organizations and managing storage is an increasing concern among IT managers who are responsible for those systems. The foundation of an efficient scalable virtualized infrastructure is having the right storage strategy. The storage layer must deliver reliable pools of storage that meet performance demands and integrate with virtualization and data protection solutions, to ensure the business can survive outages and scale painlessly as the company grows.   Dot Hill and Veeam offer an integrated backup/replication/recovery solution for virtual server environments that is high performance, high availability, highly scalable and supports both VMware and Hyper-V from a single console.

CONTEXTUAL BACKGROUND

Data storage workloads have changed. While some changes can be attributed to technological improvements, there is another cause that explains the growing unpredictability of today's workloads. That is the human element: the role each user plays in accessing information and how storage systems respond.

Traditionally, storage performance has been measured in terms of throughput and transactions per second. Now that has changed. Today, responsiveness to data has become an additional metric for successful IT operations. Consider the impact that social media and mobile devices have had on consumer demand for information. Storage systems that hold this data are directly impacted. Their active workloads are characterized by two factors: time and locality of reference. Locality of reference refers to how often the same data, or data in adjacent locations, are accessed. For example, in a data set that involves records for class registration at a university, all of the records related to students registered for Psych 101 would be related and would become active at the time that Freshmen are registering for that class.

Human interaction with information is what creates changing data over time and therefore makes the storage workload unpredictable. This unpredictability is responsible for a third factor that affects performance: how the workload changes over time. To be effective and efficient, storage systems must respond to the challenges posed by all three factors. Flash addresses all three factors well and therefore is suitable for devices such as smart phones and other consumer devices.

Flash is also an important component in the enterprise. SAN architectures are still evolving to meet the need for increasing storage capacity while generating faster access and recall. Flash will play an important role in this evolution as long as SAN architectures dominate enterprise data center storage.

FACTORS TO CONSIDER WITH FLASH AND HDDs

Enterprise usage of SSDs is expected to approach 10 million units by 2016. Despite its growing use, the cost of flash remains its biggest detriment. Forecasts within the industry call for the price of flash over the next three years to fall to four times that of mission-critical HDDs and 25 times higher than business-critical HDDs. Even as flash prices decline, the cost of an SSD array is too high for most general applications. Another factor that will limit complete adoption of flash is that worldwide production of flash is not sufficient to replace the worldwide demand for storage currently provided by HDDs. Flash media may also wear out over time when subjected to very high numbers of write cycles. HDDs may have a longer lifespan if the mechanical components of the drive do not fail.

In a report issued by Gartner, the global information technology and research company, researchers forecast that "SSD drives will complement, not replace hard disk drives in data centers."[1] The report rejected predictions that SSDs will replace HDDs over the next 5-10 years, calling them "practically impossible" and recommended against elimination of HDDs in the IT infrastructure. Until cost, production and reliability issues can be addressed, flash is not likely to replace HDD storage systems.

> ### History of Flash
>
> Storage caches have been a part of networks for nearly 25 years and their value is unquestioned. However, caching has had practical limits since its inception.
>
> It can only retain the latest and most recent data. Older data that have been squeezed out must be accessed via disk, a process 1,000 times slower than the use of SSD.
>
> Given today's workloads, which have become significantly more dynamic in the last decade, the need to increase performance beyond today's cache limitations is urgent.

THE HYBRID ARRAY SOLUTION

Fortunately, IT managers don't have to choose between SSD and HDD technologies. There is an alternative: the hybrid array, which is more cost-effective than an all-flash SSD array. The hybrid array combines traditional HDDs with flash, which can be implemented as a large cache or a storage tier. The hybrid array is the optimal solution because it combines SSD storage for high transaction performance with HDD storage for high capacity at a much lower cost. A hybrid system can be designed with only 5-10 percent of the capacity of traditional HDD storage, but is still able to deliver 80 percent of the I/O.

Understanding the operation of real-time data migration and batch data migration in the array are important considerations when comparing tiered storage systems. Storage systems that utilize batch data migration are less effective at responding to dynamic workloads that blend random and sequential access, because batch data migration optimizes data placement for yesterday's workload, not today's. Today's hot data is dynamic and constantly changing. Optimizing data placement in a batch process will always lag the workload.

Real-time tiering resolves the batch migration problem by taking into account multiple factors such as:

---

[1] Gartner Symposium ITxpo. http://www.gartner.com. April 2013.

- Frequency of data access
- Identification of I/O pattern
- Ability to distinguish streaming data from a random data pattern.

In the Dot Hill AssuredSAN^{TM} storage array there may be up to three tier levels in the hybrid array technology with real-time tiering. Hot data are moved to the top tier level (Performance SSD) for ready access. Other incoming data may be moved to the middle tier (Enterprise Class HDD) if they are warm or to the bottom level (Nearline HDD) if the data are cold. Unlike batch tiering that migrates data about once per day, the data in hybrid systems with real-time tiering move continuously. The hybrid system can improve business competitiveness by eliminating the problem most associated with batch operations: limitations on data migration.

Dot Hill hybrid storage arrays with real-time tiering offer the following benefits:

- They're solid. 99.9999 percent reliability is the standard. In other words, "It just works."
- They're fast. With massive data stores, performance requires a combination of access speed and transfer speed for randomized sequential I/O.
- They're smart. Dot Hill systems are responsive to dynamic and unpredictable workloads and manage sequential and transactional data on appropriate tiers.
- They're simple. Dot Hill hybrid storage arrays are compatible with workstations, networks, and file systems across multiples OSs and virtual server environments—and they're easily expandable to handle storage growth that is doubling every 18-24 months.

What is the difference between caching and Real-Time Tiering?

Using flash in a tiered architecture with real-time data migration solves this issue. SSD flash drives can expand the size of the cache at a much lower cost than DDR memory.

Real-time tiering systems are better at managing workloads due to their sophisticated algorithms, which do a better job of optimizing performance than caches.

While traditional tiering can be costly and somewhat difficult to manage, hybrid array tiering accesses data more quickly and efficiently than using just one class of storage, and at less cost.

CONCLUSION

A hybrid of SSDs and HDDs for SAN storage systems is more cost-efficient and therefore a better alternative for enterprise applications. A hybrid array containing algorithms enabling real-time tiering for data migration is preferable not only because of affordability, particularly for mid-markets, but also due to its capacity for quick responses to the dynamic nature of data.

By taking hot data and seamlessly moving it to the top tier while lowering cold data to the lower cost tier of HDDs, hybrid systems are built to optimize the benefits of both SSDs and HDDs. Real-time tiering in a hybrid array ensures uptimes that resolve the slowdown situation prevalent in a traditional batch

tiering process. With its combination of flash capacity and the storage capability of HDDs, the hybrid array is configured for today and tomorrow's ever-changing and dynamic data environment.

About Dot Hill Systems

Dot Hill Systems, Incorporated, based in Longmont, Colo., creates hardware and software solutions for storing, sharing, protecting and managing data. Dot Hill has deployed more than 600,000 data storage systems in its 30-year history. The company's products and services are especially designed for demanding workloads in verticals such as telecom, media and entertainment, oil and gas as well as general purpose server virtualization environments. Dot Hill has pioneered and developed innovative systems for SAN storage including RealStor™ for real-time tiering.

For more information, call (877) 872-2783 or visit our website: www.dothill.com.