

Introduction to PowerVault Auto-Tiering and SSD Read Cache Technology

This white paper introduces and describes Dell EMC™ PowerVault ME4 Series 3-Level Auto-Tiering and SSD Read Cache technologies that combine SSDs and HDDs to accelerate your application performance and efficiency with automated sub-LUN tiering and Sub-LUN SSD Read Cache capabilities.

January 2019

The information in this publication is provided “as is.” Dell Inc. makes no representations or warranties of any kind with respect to the information in this publication, and specifically disclaims implied warranties of merchantability or fitness for a particular purpose.

Use, copying, and distribution of any software described in this publication requires an applicable software license.

Copyright © 2019 Dell Inc. or its subsidiaries. All Rights Reserved. Dell EMC, and other trademarks are trademarks of Dell Inc. or its subsidiaries. Other trademarks may be the property of their respective owners. Published in the USA [09/12] [White Paper] [HXXXX]

Dell EMC believes the information in this document is accurate as of its publication date. The information is subject to change without notice

Introducing the PowerVault ME4 Series Storage Arrays

The PowerVault ME4 Series is a next gen entry-level block storage array that's purpose-built and optimized for price-sensitive SAN & DAS environments with the simplicity, performance and affordability you'd expect from Dell EMC. This block storage array brings the essential features of high-end storage to small-medium businesses — simply and affordably. With its flexible connectivity options, the PowerVault ME4 arrays offer significant advantages for SAN/DAS workloads with outstanding IOPs, massive bandwidth, scale and low latency. An all-inclusive software model helps to reduce OPEX and improve operational efficiencies — making the ME4 Series hard to beat. If you're looking for raw power and absolute simplicity in a small footprint; if you're cost-conscious and need storage availability and reliability, then Dell EMC PowerVault ME4 Series – starting in either 2U or 5U – is for you. Designed to deliver the performance of flash with the best economics of disk, the PowerVault ME4 systems are the perfect fit for small-scale storage consolidation, application acceleration, and business continuity.

Choosing the right balance of HDD and SSD varies by business use case

Today, businesses of every size and in virtually every industry and discipline need to solve a common problem – where (and how) to store their business data. Data storage technology has advanced significantly, and people who are looking for their next storage solution are faced with a barrage of information about which technology is best. A quick search on the subject brings up pages of conflicting opinions, but one trend is obvious – there is a lot of hype around the benefits that solid-state drives (SSDs) provide. SSDs provide increased performance measured in IOPs (input/output operations per second) and latencies (application response times) – especially compared to traditional HDD's.

Auto-Tiering that combines flexible hardware options with powerful software

From a hardware perspective, Dell EMC offers a variety of HDD and SSD sizes and models, as well as different enclosures to allow for 2.5" form-factor SSDs and HDDs, and also 3.5" form-factor high capacity HDDs. Customers can design systems with different combinations for the RAID head and expansion JBODs (EBODs/DAEs) to reach their capacity and performance goals. For example, one customer might want to start with a 2u24 RAID enclosure and add some SSDs, HDDS, or some combination thereof, and then increase their overall capacity by adding a 2u12 or 5u84 expansion with 10TB drives.

To leverage these hardware options effectively and economically requires intelligent software. The ME4 Series provides an intelligent software layer which does automatic data tiering between SSDs and HDDs in real-time fashion. The software can recognize incoming data patterns, and place data on the most appropriate tier to provide optimal performance.

Before discussing ME4 software in more detail, it is important to understand a basic premise of hybrid array solutions – not all customer data is “hot” all the time and requires the performance that SSDs can offer. As an example, figure 1 below shows a 3D graph depicting host I/O's across an entire pool of available capacity.

At first glance it might appear there is an even distribution of I/O across the entire LBA range; however, if we graph this data as a plot of cumulative IOPs vs. the total capacity (figure 2) we can see that 80% of the IOPs are coming from only 6.6% of the total capacity. In this specific example – the overall capacity of the storage pool was 30TB. 6.6% of that number equates to 1.98TB.

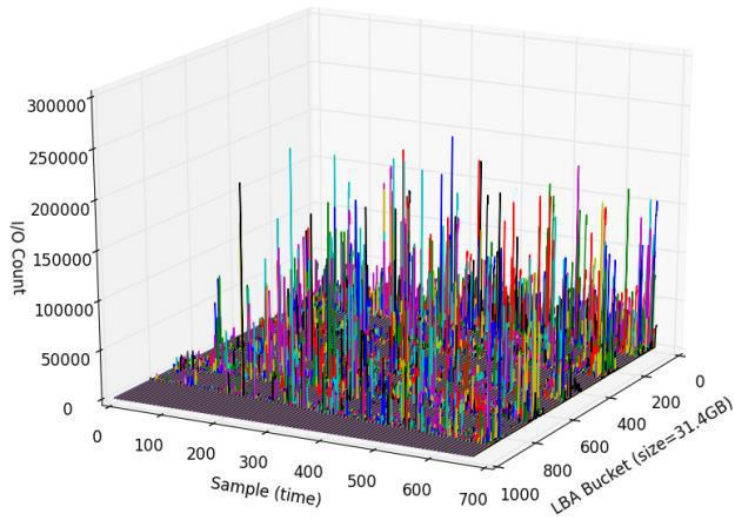


Figure 1: 3D I/O plot

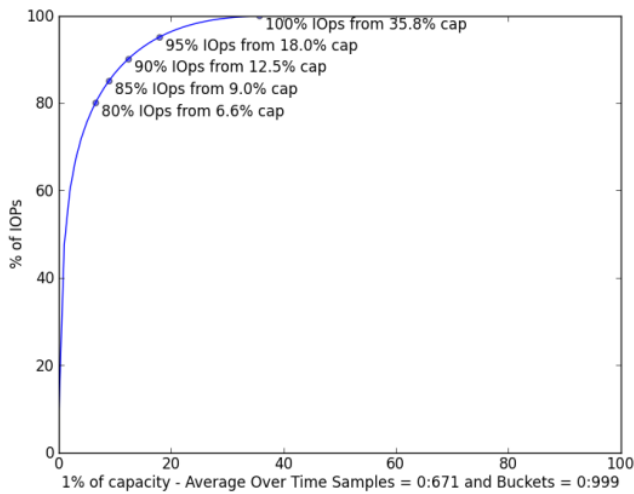


Figure 2: CDF plot of I/Os over entire pool capacity

What this tells us is that we can service a large percentage of the host I/O in just 1.98TB of space, so why not size our SSD purchase to match this size and utilize more cost-effective HDDs for the rest of the data. This is a key premise to Dell EMC's Entry-Level Storage solutions – all empowered by unique software.

ME4 Auto-Tiering Enabling Technology

The ME4 software is based on a concept called “paging”. User volumes are logically broken down into small, 4MB chunks called pages. Pages are ranked based upon a sophisticated algorithm. The page rank

is used to very efficiently select ideal pages to move between tiers. The result is that pages can be migrated between tiers automatically such that I/O's are optimized in real-time. Figure 3 captures this concept. Note that 3 pages have been moved from the Enterprise SAS Tier to the SSD Tier, 1 page from the Midline SAS Tier to the SSD Tier, and a single page has been moved down from the SSD Tier to the SAS Tier.

This tiering is done at the sub-LUN level which provides efficient data movement with a minimum of CPU and memory resources. Movement can happen in real-time rather than in offline batch movements – providing a more reactive, efficient means of ensure the most active data is promoted to the highest available tier. Many competitive offerings require a batch migration done after a 24hr period. Most customers know that in today's world, a 24hr period is a long time in terms of what data is most active and critical to their business. Often with these batch migration approaches, by the time the data has been migrated – it is no longer as critical and therefore customers are never able to fully utilize their SSD investment.

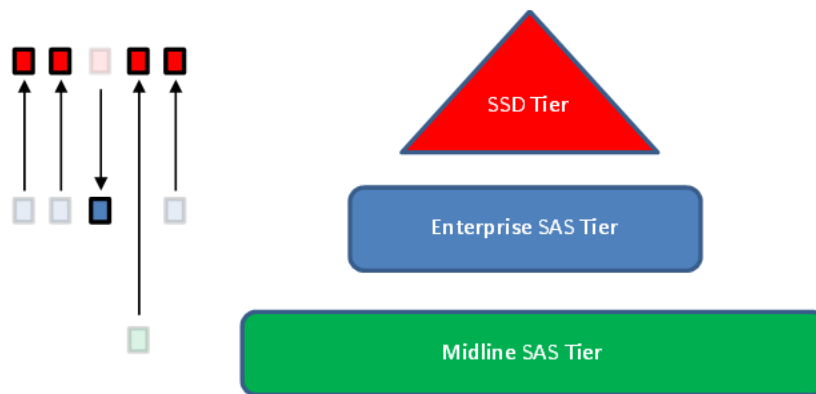


Figure 3: Pages moving between tiers

Key Concepts:

- The tiering algorithm runs every 5 seconds. The algorithm ranks, scans, and migrates pages.
- Pages are ranked according to access patterns
- A scan looks for highly-ranked pages
- Highly ranked pages are migrated up/down
- Pages are only migrated down if room is needed for a highly-ranked page
- Only 80MB migrated from any one 5-second interval to avoid degrading system throughput

ME4 SSD Read Cache

In addition to tiering, the ME4 offers a second method to leverage SSDs – SSD Read Cache. Read cache is a special-purpose disk group dedicated to hold volatile copies of data within one or more SSDs.

Advantages of a read cache include:

- Improved performance in read-centric workloads where maximum IOPS is crucial – note this does not benefit writes – and is intended for customer workloads that are highly read-intensive
- Requires fewer SSDs when compared to a dedicated SSD tier.
- A single SSD disk can be used as a read cache for a storage pool; whereas, an SSD tier requires multiple disk to provide fault tolerance.
- Takes advantage of tiering-like algorithms so that only active pages are copied into the read cache.
- Active pages can be from any volume within the pool.

The read cache becomes active when the controller cache is saturated. The controller cache software gives “hints” to the ME4 software about which pages it would like to keep but does not have space for in controller cache. The system then copies these pages into the SSD read cache. Reads are served from the SSD, and writes are written to both SSD and HDD.

Key Concept: The read cache keeps copies of data that already exist in the hard disk tier; therefore, it doesn’t require more SSD disks to be fault-tolerant. The SSD disk(s) provide the performance and the HDD disks provide the fault tolerance. This is a cost-effective method of providing performance.

Test Results

Work was done with a third-party vendor to conduct a series of tests to compare how its Hybrid array performance against an all HDD array. There were several workloads tested (Microsoft SQL Server, Microsoft Exchange, and a VMWare environment with mixed applications running. The baseline all-HDD test was performed first, followed by a Read-cache test, and then a full SSD tiering test.

Figure 4 shows the results obtained when comparing the baseline all HDD configuration (blue), to the Read-cache configuration (red) and the full SSD tier testing (Green). Adding the SSDs, along with the tiering software, resulted in increased database transactions per second, as well as lowered response times.

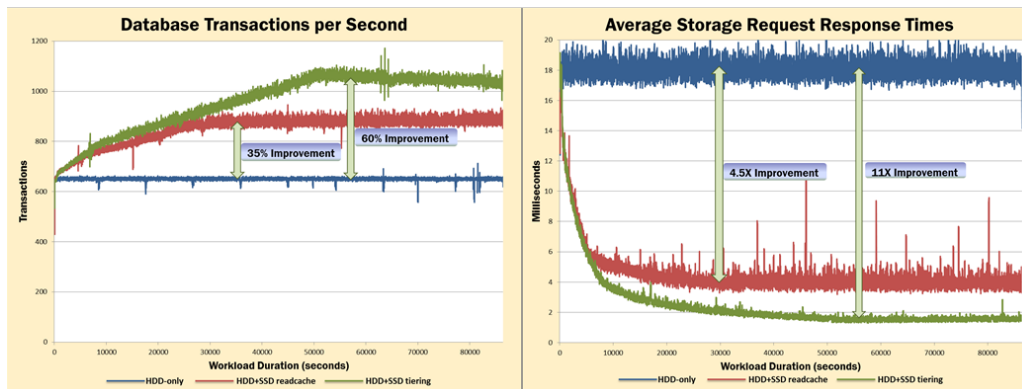


Figure 4: Results from Microsoft SQL OLTP testing

It should be noted, that aside from provisioning the SSDs, as either read-cache or an SSD-tier there is no additional user intervention required. ME4 Series software automatically analyzes the incoming workloads and makes intelligent decisions on where to place data to provide the maximum performance. As an example, most database workloads are small, random transactions. Those workloads are generally IOP intensive, and better suited for an SSD vs. an HDD. Therefore, when the ME4 array detects this workload, it automatically starts to migrate that data from HDD up to an SSD tier. As seen in Figure 4, this migration starts to happen immediately and reached a steady-state in less than 14 hrs. In comparison, competitive batch-migration solutions would require significantly more time before providing performance benefits.

Conclusion

The use of SSDs in enterprise arrays has clearly increased and led to a number of innovative solutions in the marketplace. Coupled with decreasing SSD costs, customers are finding it more practical to look at hybrid storage arrays.

The use of SSDs can enhance the performance of a system. Improved read and write performance is possible through automated tiering for array efficiency and performance with less hardware expense. Alternatively, the use of SSD Read Cache increases execution speed through caching frequently accessed read data.

The PowerVault ME4 software features such as Tiering and SSD Read Cache, coupled with flexible hardware options, enables customers to effectively implement cost-effective, high-performance arrays.