WHITE PAPER

Titan T2000
Microsoft SQL Server

CRS TITAN

POWERED BY DELL Technologies

Dell PowerVault ME5 Series: Microsoft SQL Server Best Practices

February 2022

H19049

White Paper

Abstract

This document provides best practices for deploying Microsoft SQL Server with Dell PowerVault ME5 Series storage, including recommendations and considerations for performance, availability, and scalability.

Dell Technologies

Copyright

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.

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

Contents

Executive summary....4

Introduction....5

Best Practices....6

SQL Server design considerations....7

Volume configuration....11

References....15

Executive summary

Overview

This paper provides guidance for using Dell PowerVault™ PowerVault ME5 storage systems in a Microsoft® SQL Server® environment. SQL Server is a robust product that can be used in a variety of solutions, allowing you to prioritize performance, manageability, and flexibility depending on your environment. This paper provides important considerations and recommendations to help meet your design goals, and builds upon the best practices in the PowerVault ME5 Administrator's Guide on Dell.com/support.

Note: While following the best practices in this document is strongly recommended by Dell Technologies, some recommendations may not apply to all environments. These recommendations are specific to SQL Server and PowerVault ME5. It is not a comprehensive SQL Server best practices guide for all environments. For questions about the applicability of these guidelines in your environment, contact your Dell Technologies representative.

Audience

This document is intended for PowerVault ME5 administrators, database administrators, architects, partners, and anyone responsible for configuring PowerVault ME5 storage systems. Some familiarity with Dell storage systems is assumed.

Revisions

Date	Description
February 2022	Initial release

Note: This document may contain language that is not consistent with Dell Technologies' current guidelines. Dell Technologies plans to update the document over subsequent future releases to revise the language accordingly.

We value your feedback

Dell Technologies and the authors of this document welcome your feedback on this document. Contact the Dell Technologies team by email.

Author: Doug Bernhardt

Introduction

About ME5

The PowerVault ME5 is the new gold standard for entry storage that is purpose-built and optimized for high-value SAN and DAS virtualized workloads. Available in 2U or dense 5U base systems, the low-cost PowerVault ME5 simplifies the challenges of server capacity expansion and small-scale SAN consolidation with up to 336 drives or 6PB capacity. It also comes with all-inclusive software, incredible performance, and built-in simplicity with an HTML5 management UI, PowerVault Manager. Connecting PowerVault ME5 storage to a PowerEdge server or to a SAN ensures that business applications will get high-speed and reliable access to their data — without compromise.

Product features include the following:

Simplicity: A web-based management UI (HTML5), installs and configures in 25 minutes, and easily deploys in 2U or 5U systems.

Performance: Compared to its PowerVault ME4 predecessor, PowerVault ME5 storage packs a lot of power and scale with Intel® processors. The PowerVault ME5 processing power delivers incredible performance gains over the ME4 Series, as well as increased capacity and bandwidth.

Connectivity: Robust and flexible connectivity starting with a 12Gb SAS back-end interface, and a front-end interface options including four 16/32Gb FC ports per controller, four 10Gb iSCSI ports per controller (BaseT), four 10/25 Gb iSCSI (Optical), or four 12Gb SAS ports per controller.

Scalability: Both 2U and 5U base systems are available, with the 2U system supporting either 12 or 24 drives and the 5U system supporting 84 drives. Each of the 2U (ME5012 and ME5024) and 5U (ME5084) base systems supports optional expansion enclosures of 12, 24, and 84 drives, allowing you to use up to 336 drives. Drive mixing is also allowed.

All-inclusive software: PowerVault ME5 software provides volume copy, snapshots, IP/FC replication, VMware® VCenter Server® and VMware Site Recovery Manager™ integration, SSD read cache, thin provisioning, three-level tiering, ADAPT (distributed RAID), and controller-based encryption (SEDs) with internal key management.

Management: An integrated HTML5 web-based management interface, PowerVault Manager is included.

For more information, see the PowerVault ME5 product page.

Best Practices

Overview

Use the following general steps to set up and configure a PowerVault ME5 system for SQL Server:

1. Capture the storage I/O performance characteristics and capacity requirements of your SQL Server workload.
2. Review the remaining sections of this document and apply the best practices that are applicable to your workload and environment. Because SQL Server workloads can vary, not all recommendations may apply.
3. Follow the deployment instructions for setting up a PowerVault ME5 system found in the PowerVault ME5 Deployment Guide on Dell.com/support.
4. Configure the PowerVault ME5 system using the PowerVault ME5 Administrator’s Guide, applying best practices for Microsoft Windows and SQL Server as recommended by Dell Technologies.

SQL Server design considerations

Introduction

The I/O storage system is a critical component of any SQL Server environment. Sizing and configuring a storage system without understanding the I/O requirements can result in poor performance, insufficient capacity, or both. Analyzing performance in an existing environment using a tool like Live Optics can help define the I/O requirements. Your Dell Technologies representative can assist with Live Optics data collection and analysis. For best results, capture performance statistics for a period of at least 24 hours that includes the system peak workload. Also include special operational circumstances such as critical business and maintenance activities.

OLTP workloads

While every environment is unique, an online transaction processing (OLTP) workload typically consists of small random reads and writes. A storage system for OLTP workloads is primarily sized based on capacity and the number of IOPS required.

OLAP/DSS workloads

An online analytic processing (OLAP) or decision support system (DSS) workload is typically dominated by large sequential reads. A storage system for OLAP/DSS workloads is primarily sized based on throughput. When designing for throughput, the performance of the entire path between the server and the drives in the PowerVault ME5 array needs to be considered. For best throughput, consider using 32Gb Fibre Channel (FC) or 25 Gbps iSCSI connectivity to the array. To meet high-throughput requirements, multiple physical paths may be required.

Mixed workloads

The most common scenario for a SQL Server environment is a mixed workload. Typically, SQL Server I/O patterns do not strictly fall into an OLTP or OLAP pattern. This is what can make SQL Server workloads challenging because no two workloads behave the same. In addition, the same SQL Server host or instance may be servicing multiple applications or transaction workloads. A mixed workload can also imply that multiple applications (in addition to SQL Server) are residing on the same host or accessing the same storage. The combined workload of these applications invalidates any typical application I/O usage pattern. For these reasons, it is important to gather actual performance metrics for best sizing results.

PowerVault ME5 configuration

Balanced configuration

Creating a balanced storage configuration is important because SQL Server workloads can vary greatly and I/O patterns can often fluctuate due to changes in the database environment, evolving data-access patterns, or data growth. For most SQL Server workloads, it is recommended to configure the PowerVault ME5 array using the Virtual storage type and ADAPT as the RAID level.

For best performance, a minimum of 24 drives should be used when using the ADAPT RAID level because this is the minimum number required to create two Virtual storage pools, one per controller in a dual-controller system. Start with 24 SSD drives and add additional drives as needed to achieve performance and capacity requirements. A PowerVault ME5 array with SSDs spread evenly across 2 Virtual storage pools (one per controller) configured with the ADAPT RAID level provides the best overall balance of performance, flexibility, capacity, and data protection. Detailed information on ADAPT can be found in the PowerVault ADAPT Software paper.

There is a variety of other RAID levels and storage configurations available for very specific workloads. Make sure the design tradeoffs are completely understood when choosing custom configurations and settings. In many cases, modifying the storage configuration for existing volumes will involve halting I/O on those volumes, resulting in a SQL Server outage. For detailed information on all available choices, including ADAPT, see the PowerVault ME5 Administrator's Guide.

RAID levels and performance

The performance characteristics of the PowerVault ME5 largely depend on the RAID level chosen for the storage configuration. While the ADAPT RAID level provides the most balanced configuration, other RAID levels are available for maximum performance or maximum capacity. The two additional levels discussed in this section are RAID5 and RAID10. For all available RAID levels, consult the PowerVault ME5 Administrator's Guide.

The performance numbers presented are maximum values on an PowerVault ME5 array with a minimum of 24 drives and two storage pools, one per controller. Note that performance among RAID levels only has significant variation in workloads such as OLTP where frequent writes occur. For read-only workloads such as OLAP, there is practically no performance benefit to select RAID levels based on performance. Use care when configuring multiple RAID levels on the same PowerVault ME5 array. Ensure that the number of drives in the storage pool is sufficient to provide the performance desired. Mixing RAID levels within a storage pool is not recommended. Remember that when selecting RAID levels, there are design considerations other than performance. RAID levels impact the data protection, capacity, and flexibility of the overall storage design. Therefore, the goal is to find the best balance of performance, data protection, capacity, and flexibility.

ADAPT

When using the ADAPT RAID level, the PowerVault ME5 can perform up to 170,000 IOPS using an OLTP workload[1] and up to 13 GB/sec for OLAP workloads[2].The solid OLTP performance and outstanding read performance, and the balance of performance, capacity, and data protection, make ADAPT a good choice for mixed-database workloads.

RAID5

When additional performance and capacity is required, RAID5 can be used. In a RAID5 configuration, some disk resources that were reserved for data protection are used for additional capacity and performance. Therefore, RAID5 offers less data protection than ADAPT. However, it can deliver up to 207,000 IOPS for OLTP[3] workloads

RAID10

For maximum OLTP performance, RAID10 can deliver up to 344,000 IOPS for OLTP[3] workloads. In a RAID10 configuration, the usable capacity is reduced to 50% of raw

[1] An OLTP workload is defined as having an 8k block size and a 70/30 read/write mix.

[2] An OLAP workload is defined as having a 128k+ block size with 100% reads.

[3] Based on Dell internal testing, January 2022.

capacity. Therefore, a RAID10 solution requires almost double the number of drives compared to more space-efficient RAID types.

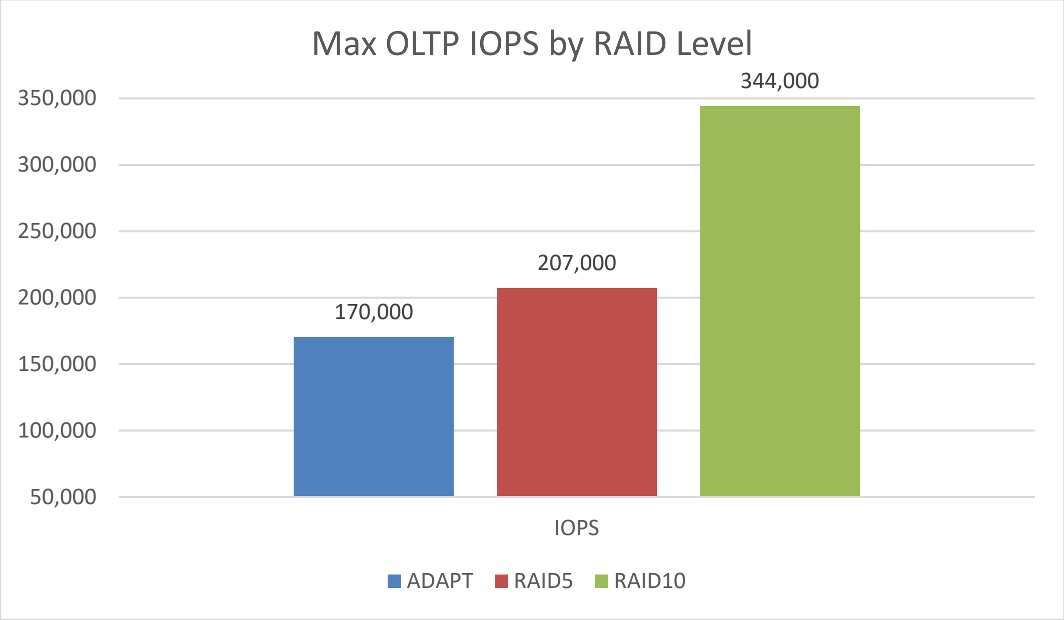

Figure 1. Maximum OLTP IOPS by RAID level

Validating the storage design

After the I/O requirements have been defined, it is easy to determine whether the hardware can provide the desired performance by running some simple tests. Diskspd is a free Microsoft utility that can simulate I/O patterns generated by SQL Server. There are several other utilities available as well. When selecting a utility to simulate I/O, verify that it meets the following requirements:

- Ability to configure block size
- Ability to specify number of outstanding requests
- Ability to configure test file size
- Ability to configure number of threads
- Support for multiple test files
- Does not write blocks of zeros during tests

Validating the I/O path

The first thing to test on a new configuration is the path between the server and the array. Running a large block sequential read test using small files should saturate the path between the server and the array. This test verifies that all paths are fully functional and can be used for I/O traffic. Run this test on a dedicated server and array; a live system could cause significant performance issues.

To validate the I/O path, run a large block sequential read test using the following guidelines:

- Create one LUN per storage controller.

- Format the volumes using a 64KB allocation unit.
- Use a block size of 512KB for the test.
- Configure the test for 32 outstanding I/Os.
- Use multiple threads. Eight is the recommended starting point.

If the displayed throughput matches the expected throughput for the number of HBA ports in the server, the paths between the server and PowerVault ME5 array are set up correctly.

Validating the drives

After the I/O path has been validated, the next step is to test the drives. For best results when testing drives on a PowerVault ME5 array, use the following guidelines when configuring the test:

- In a dual-controller system, use at least one volume per pool with each pool on a separate controller. This ensures that I/O will be distributed across both controllers. Using both controllers provides a more accurate simulation of real-world activity. For best results, use the same number of volumes on each controller.
- When performing I/O tests on any storage platform, it is important to use files that are larger than the controller cache. For more accurate results, use a file size that matches the amount of data being stored. In an environment where that is not practical due to a large data set, use a file size of at least 100GB.
- Avoid using test utilities to generate files full of zeros for drive validation. Some I/O test tools, including Diskspd, SQLIO, and IOMeter, can be used to write zeroes for drive validation, which causes inaccurate results when testing with files containing only zeros.

  The contents of the test file can be verified by viewing the test file with a hex editor after different stages of a test. For example, create a small test file and view it after the initial creation, as well as after the test has run for a few seconds. If the file is filled with zeros, check the configuration settings for the utility or select another utility. Diskspd and IOMeter initially create test files filled with zeros, and then write random characters when performing write tests. To properly initialize a Diskspd or IOMeter test file, run a sequential write test until the entire file has been overwritten with non-zero data. Unfortunately, SQLIO writes zeros during write tests and therefore is not recommended for drive validation.

The purpose of this drive testing is to validate that the storage design will provide the required throughput and IOPS with acceptable latency. It is important that the test does not exceed the designed capacity of the array. For example, an array designed for a workload of 5,000 IOPS is likely to perform poorly with a workload of 10,000 IOPS. If a test is generating a workload higher than the designed capacity, adjust the workload being generated by reducing the number of threads or outstanding I/Os. The results of the Live Optics analysis provide an I/O target to simulate using these tests. To estimate the performance capabilities of the array, run I/O tests with a range of I/O sizes commonly seen with SQL Server. When testing random I/O, test with an I/O size of 8KB and 64KB. When testing sequential I/O, start with I/O sizes of 8KB and 64KB. Because processes such as read-ahead scans and backups can issue much larger sequential I/O, it is also a good idea to test block sizes up to 1024KB.

Volume configuration

Creating volumes

Many types of files are part of a SQL Server instance. Those types of data often have different performance requirements. For performance-sensitive applications requiring maximum performance, Dell Technologies recommends creating at least five volumes for an instance of SQL Server as shown in the following table.

Table 1. Volume-provisioning requirements

File type	Number of volumes	Typical performance requirements
User DB data	At least 1 per instance	Lower performance may be acceptable
User DB transaction log	At least 1 per instance	High performance required
Data root directory (includes system DBs)	1 per instance	Lower performance may be acceptable
Tempdb data and transaction log	1 per instance	High performance required
Native SQL Server backup	1 per instance	Lower performance may be acceptable
Memory-Optimized Filegroup (if used)	At least 1 per instance	High performance required

Performance considerations

When there is one group of databases that require high performance and another group that does not, consider creating a set of volumes for each group of databases. This strategy will make it easier to adjust the storage configuration in the future. It also makes it easier to distribute the I/O load evenly across both controllers. Databases that have very high performance requirements can be spread across two or more data files on separate volumes to leverage resources on both controllers.

Flexibility, manageability, and virtualized environments

For ultimate flexibility, create a volume for each user database file. This allows for independently optimizing the storage for each individual database. With thin provisioning, there is no space penalty for creating numerous volumes. However, a large number of volumes can be difficult to manage, especially in virtualized environments. The DBA or storage administrator must find the right balance between flexibility and manageability when determining the number of volumes to create. Virtualized SQL Server environments are a good example where placing multiple file types on a single volume can make sense. Understanding the underlying storage layout and database I/O patterns is critical to making the best decisions. For VMware environments, see the PowerVault ME5 VMware vSphere Best Practices white paper.

Windows setup and configuration

Allocation unit size

Use a 64KB allocation unit size when formatting volumes that will contain database files (transaction log and data) or database backups. This is a Microsoft best practice for SQL Server and provides the best performance in various SQL Server workloads tested by Dell Technologies.

MPIO

PowerVault ME5 arrays support Asymmetric Logical Unit Access (ALUA), and when MPIO is configured, the default MPIO policy is Least Queue Depth. This is the recommended setting for all database volumes. This setting works best for most environments because it is easy to manage and performs very well. Use other MPIO policies with caution and remember to review custom MPIO policies when adding or removing volumes from the host.

SQL Server I/O reduction

Memory

Unnecessary I/O can be avoided, and performance can be increased by allocating the proper amount of memory to SQL Server. SQL Server performs all I/O through the buffer pool (cache) and therefore uses a large portion of its memory allocation for the buffer pool. Ideally, when SQL Server performs I/O, the data is already in the buffer pool and it does not need to go to disk. This type of I/O is referred to as logical I/O and is the most desirable because it results in the best performance. If the SQL Server data does not need to reside in the buffer pool, it will need to access disks, resulting in physical I/O. Proper memory allocation is critical to SQL Server performance and can improve storage performance as well. In many cases, SQL Server and storage performance can be further improved by adding memory. Adding memory generally improves performance, but there is a point of diminishing returns that is unique to each environment.

Buffer pool extension

The SQL Server buffer pool extension feature allows the SQL Server buffer pool to be extended to a file on the file system to provide additional space to cache data or index pages. Using this feature may provide performance benefits without adding memory to the database server in some cases. By caching more pages on the server, the I/O load on the array is reduced. When placing the buffer pool extension on the array, create a separate volume for the buffer pool extension and do not take snapshots of the buffer pool extension volume. The buffer pool data is repopulated by SQL Server when the instance is restarted, therefore data recovery does not apply.

Database compression

The overall I/O workload can be reduced by enabling database compression in SQL Server. Although there is a tradeoff in terms of CPU utilization on the database server, compression is still a viable option to consider and test in any environment. Database compression reduces I/O by reducing the amount of data that needs to be stored. The SQL Server data pages are compressed in memory before being written to disk, resulting in fewer pages needed to store the same number of rows and therefore less I/O.

Instant file initialization

By default, SQL Server writes zeros to the data file during the allocation process. The process of zeroing out the data files consumes I/O and acquires locks as the SQL Server data pages are written. This activity can occur for minutes or even hours depending on the file size. Although this may not seem important, writing zeros to these files can occur at critical periods when time and performance are critical, such as database auto growth, expanding a full data file, replication, or restoring a database as part of a disaster-recovery event.

When Instant File Initialization is enabled, SQL Server skips the process of zeroing out its data files when allocating space. Dell Technologies recommends enabling Instant File Initialization.

Resource Governor

The SQL Server Resource Governor allows database administrators to limit the CPU, memory, and I/O resources that a query can consume. For example, the Resource Governor can be used to reduce the impact of a user running an I/O-intensive report by limiting the maximum number of IOPS that that user can perform. While a query throttled by the Resource Governor will take more time to complete, overall database performance can be improved by reducing contention for bottlenecked resources or throttling lower priority workloads, such as reporting. For more information on the benefits and implementation of Resource Governor, see the Microsoft SQL Server Resource Governor page.

Database design considerations

Reducing SQL Server I/O requires a holistic approach. Many of the items in this section will require involvement from the whole team responsible for the SQL Server applications, including the business owner, architect, developer, database administrator, and system administrator. Decisions at the design level have a multiplied impact downstream because data is written and read multiple times and duplicated in various types of database copies. This includes databases copied for other uses, such as testing and reporting, replicated databases, replicated storage, and backups.

One of the most challenging aspects of SQL Server is that the I/O pattern and the amount of I/O that is generated can vary greatly depending on the application, even if those applications have databases of the same size. This is because the design of both the database and the data-access code control the SQL Server I/O. Database tuning can be one of the most cost-effective ways to reduce I/O and improve scalability.

At a high level, consider the tips in the following subsections when tuning a database to reduce I/O.

Database design

The database design determines the foundation of the entire database and the schema for how data will be stored and ultimately accessed. The database design should support both usability and efficient data access. This includes efficient table design and datatypes as well as indexes, partitioning, and other features that can improve efficiency. It is common for database design to be focused only on usability while performance and scale are overlooked.

Query design

How a query is written can greatly affect the amount of I/O SQL Server needs to perform when executing the query. Queries should return only the required amount of data in the most efficient manner possible. For best performance and scale, tune the queries responsible for consuming the most resources.

Application design

Consider how applications are using the data and how the data is requested. Sometimes code and component reuse can result in the same data being unnecessarily retrieved repeatedly. All data access should be purposeful. For example, when possible, SQL queries should avoid using “SELECT *” syntax and accessing or updating columns that are not used by the application.

Maintenance

SQL Server uses a cost-based optimizer to generate query plans for data access. These plans are based on the statistics regarding how data is distributed in the tables. If the statistics are inaccurate, bad query plans may result and unnecessary I/O will be performed. Proper database maintenance includes ensuring that statistics are up to date.

Frequent data modifications can also lead to fragmentation within SQL Server data files, producing unnecessary I/O. Fragmentation can be addressed through index reorganization or rebuilds as part of regular database maintenance. The database maintenance process itself can also have a large I/O impact. Typically, every table and index does not need to be rebuilt or reorganized every time maintenance is run. In addition, table partitioning strategies can also be leveraged to make the maintenance process more selective. Consider implementing intelligent maintenance scripts that use usage statistics to perform maintenance on an as-needed basis. For mission-critical databases, consider maintenance activities as part of the overall design. If maintenance is not considered as part of the overall process, issues can arise, such as unmanageable sizes and feature incompatibilities that limit available options and strategies.

Unnecessary index rebuilds will also cause updates to ME5 intelligent data access features such as read cache and tiering and therefore may cause unnecessary data shuffling.

References

Dell Technologies documentation

The following Dell Technologies documentation provides other information related to this document. Access to these documents depends on your login credentials. If you do not have access to a document, contact your Dell Technologies representative.

- *PowerVault ME5 Documentation*
- *Dell Technologies storage*
- *Dell Technologies SQL Server solutions*

Microsoft documentation

The following Microsoft documentation provides additional information.

- *Microsoft SQL Server*
- *SQL Server Labs and Workshops*