

65TB Data Warehouse Fast Track Reference Architecture for Microsoft SQL Server 2017 using Dell EMC PowerEdge R640 and Dell EMC PowerVault ME4024

Configuration and performance results

Abstract

Design principles and guidelines used to achieve an optimally balanced 65TB Data Warehouse Fast Track reference architecture for Microsoft® SQL Server® 2017 using Dell EMC™ PowerEdge™ R640 servers and Dell EMC PowerVault™ ME4024 storage.

September 2018

Revisions

Date	Description
September 2018	Initial release

Acknowledgements

Author: Doug Bernhardt

Special thanks to Jamie Reding at Microsoft.

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.

© 2018 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

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

Table of contents

Revisions.....	2
Acknowledgements.....	2
Table of contents	3
Executive summary.....	4
1 Microsoft SQL Server 2017: Data warehousing with improved column-store technology	5
1.1 Dell EMC Data Warehouse Fast Track reference architectures for SQL Server 2017	5
2 Recommended reference architectures	6
2.1 Single-server reference architecture	6
2.2 Highly available reference architecture	8
3 Hardware components	10
3.1 PowerEdge R640 server	10
3.2 QLogic QLE2662 16Gbps host bus adapter	10
3.3 PowerVault ME4024 array.....	10
4 Storage configuration	12
4.1 Cabling.....	12
4.2 Storage pools.....	12
4.3 Host setup.....	13
4.4 ME4 Series volumes.....	13
5 Server configuration	15
5.1 System BIOS	15
5.2 Host bus adapters	15
6 Windows Server 2016 configuration	16
6.1 Power plan.....	16
6.2 Lock pages in memory	16
6.3 Windows volumes.....	16
6.4 MPIO.....	17
7 SQL Server 2017 Enterprise Edition configuration	18
7.1 Grant perform volume maintenance task privilege.....	18
7.2 SQL Server maximum memory	18
7.3 Max degree of parallelism (MAXDOP)	18
7.4 Resource governor	18
7.5 Database configuration.....	18
7.6 Tempdb configuration.....	18
8 Additional considerations for the highly available reference architecture	19
9 DWFT for SQL Server 2017 certification.....	20
10 Summary	21
A Technical support and resources	22

Executive summary

Dell EMC and Microsoft, in cooperation, provide guidelines and principles to assist customers in designing and implementing a balanced configuration for Microsoft® SQL Server® data warehouse workloads to achieve out-of-the-box scalable performance. These database reference architectures enable each of the components in the database stack to provide optimal throughput to match the database capabilities of the specific setup. Innovative Dell EMC™ PowerEdge™ 14th-generation (14G) servers, along with robust and cutting-edge Dell EMC PowerVault™ ME4024 storage arrays, form efficient candidates for a high-performing data warehouse solution.

This paper describes the design principles and guidelines used to achieve an optimally balanced 65TB Data Warehouse Fast Track (DWFT) reference architecture for SQL Server 2017 using PowerEdge R640 servers and PowerVault ME4024 arrays. The configuration used to achieve the performance numbers for the reference configuration is presented in detail.

The target audience for this paper includes database administrators, business intelligence architects, storage administrators, IT directors, and data warehousing users seeking sizing and design guidance for business intelligence solutions with SQL Server 2017.

1 Microsoft SQL Server 2017: Data warehousing with improved column-store technology

Microsoft SQL Server 2017 has made significant improvements in data warehousing technologies and performance, including column-store features as well as many other improvements.

Column-store indices offer great advantages over traditional row stores for analytics and data warehousing queries. They are ideally suited for the star schemas, and tables with billions of rows which are commonly seen. Among their advantages for analytics are:

Up to 10X compression in data size: Data warehouses are very large by nature and the compression offered by column-store index technologies offers both space and cost savings as well as significantly increased performance. These benefits are possible due to the dramatically reduced I/O requirements given by the compression and coupled by the ability to only scan the specific columns required by each query. This compression also reduces the amount of memory required to hold a given number of rows from the source data warehouse.

Additional indices: SQL Server 2017 adds the capability to add B-tree indices to column store-based tables, which enables efficient single-row lookup.

In addition to these architectural features, Microsoft has further optimized the processing of queries in column-store indices in the following ways:

Operator pushdown: Pushdown refers to moving both filter and aggregation query operations closer to the data, so that many of the filters and calculations can be done in the scan operators, dramatically reducing the volume of data that needs to be handled further on in query processing.

Batch-mode processing: SQL Server 2017 includes enhancements in batch-mode processing that handles many rows at a time rather than serially doing calculations on each individual row. These batch operations are further optimized by leveraging Single Instruction Multiple Data (SIMD) vector processing CPU instructions in the Intel® architectures.

1.1 Dell EMC Data Warehouse Fast Track reference architectures for SQL Server 2017

Dell EMC and Microsoft have refreshed the DWFT reference architecture offerings with the latest technology advancements in database, server, and storage technology. Dell EMC PowerEdge 14G servers with Intel® Xeon® processors, Dell EMC PowerVault ME4024 arrays, and Microsoft SQL Server 2017 are the latest additions to the list of reference architecture components.

The Dell EMC DWFT reference architectures for SQL Server 2017 are engineered jointly by Dell EMC and Microsoft. The hardware and software optimizations are tested by Dell EMC and the performance results are crosschecked by Microsoft. This approach presents a fast time to value using integrated, balanced, and verified architectures.

2 Recommended reference architectures

The following subsections describe the two different DWFT reference architectures for SQL Server 2017, comprised of PowerEdge R640 servers and PowerVault ME4024 arrays.

2.1 Single-server reference architecture

Figure 1 illustrates the single-server reference architecture with its major elements and Table 1 lists the component details.

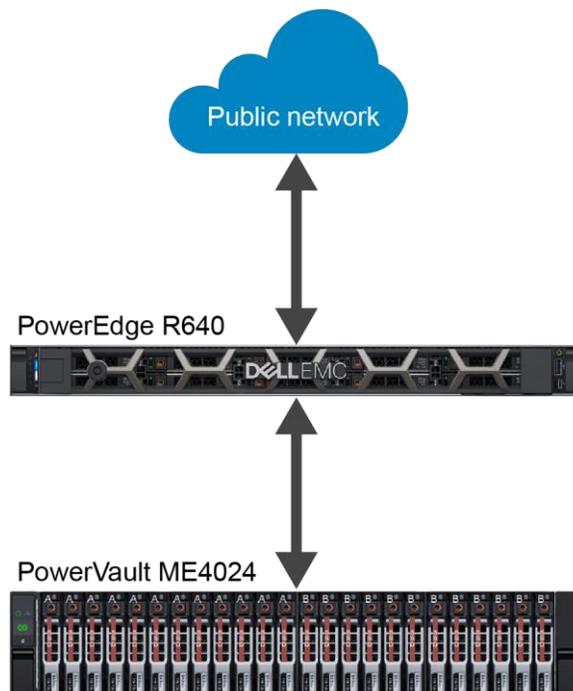


Figure 1 Single-server reference architecture

Table 1 Single-server reference architecture details

Component	Description	
Server	PowerEdge R640	
	Processors	Dual Intel® Xeon® Gold 6126 Processor (2.6GHz, 12 cores, 24 threads)
	Total cores	24
	Total logical processors	48 (Hyper-Threading Technology enabled)
	Total memory	1024GB
	Network adapters	Minimum of one network adapter (1Gbps or 10Gbps based on requirements) Recommended to have more than one network adapter with load balancing configured
	Host bus adapters	Two QLogic® QLE2662 dual-port 16Gbps FC HBAs
Software	Operating system	Microsoft Windows Server® 2016 Standard Edition
	Database software	SQL Server 2017 Enterprise Edition
Storage	Array	PowerVault ME4024 (firmware: GT275R0003-01)
	Front-end I/O ports	8 x 16Gbps FC ports (four per controller)
	Disk drives	24 x 1.92TB SSDs (2.5-inch SAS)

2.2 Highly available reference architecture

For database high availability, Windows® failover clustering is recommended. Using Microsoft clustering services, one database server is configured as the primary (active) server and the second server is configured as the secondary (passive) server. The secondary server should have exactly the same configuration as the primary server. Since the database is only active on a single server at any point of time, the performance of the database on the primary server (active) is comparable to the single server configuration (discussed earlier).

Figure 2 illustrates the highly available reference architecture with its major elements and Table 2 lists the component details.

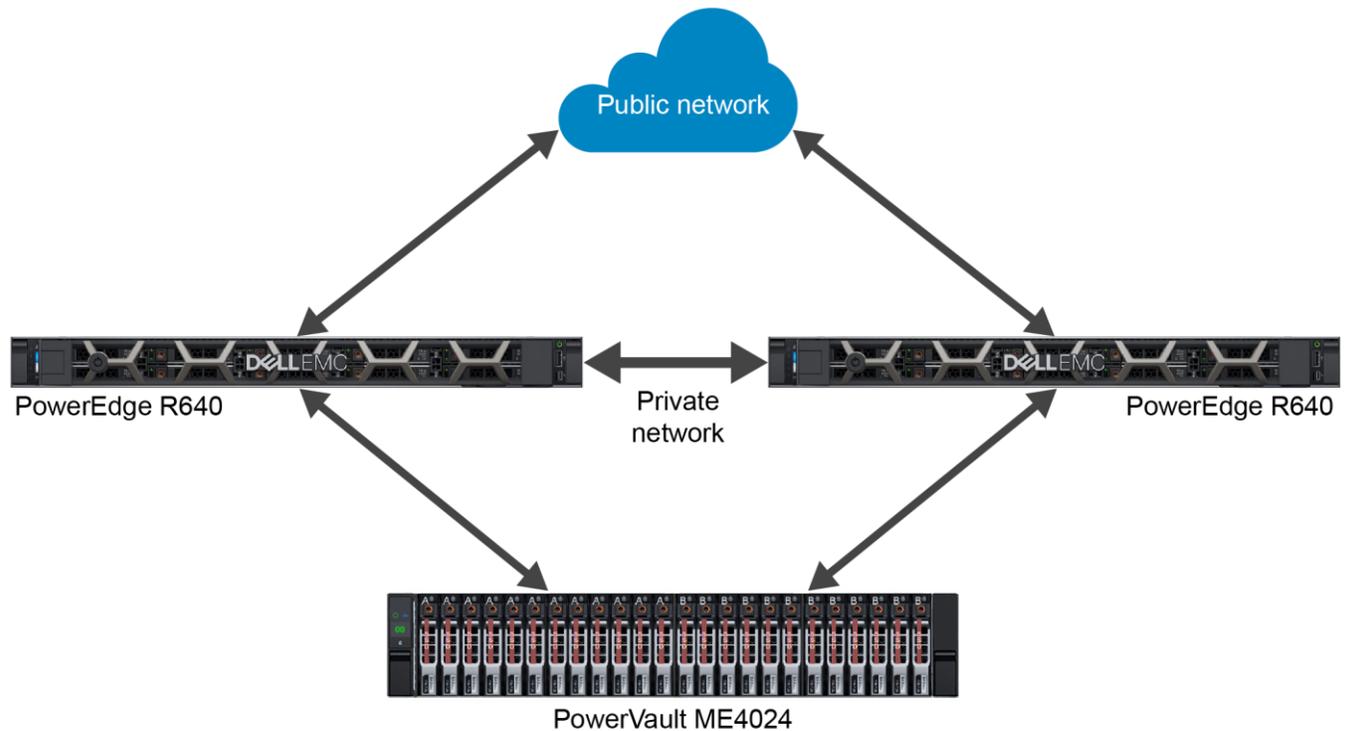


Figure 2 Highly available reference architecture

Table 2 Highly available reference architecture details

Component	Description	
Server	Two PowerEdge R640	
	Processors	Dual Intel® Xeon® Gold 6126 Processor (2.6GHz, 12 cores, 24 threads)
	Total cores	24 per server
	Total logical processors	48 per server (Hyper-Threading Technology enabled)
	Total memory	1024GB per server
	Network adapters	Public network: Minimum of two network adapters (1Gbps- or 10Gbps-based requirements) per server with load balancing configured Private network (cluster): Minimum of one network adapter with 1Gbps (or more) connectivity per server
	Host bus adapters	Two QLogic QLE2662 dual-port 16Gbps FC HBAs per server
Software	Operating system	Windows Server 2016 Standard Edition with Windows Failover Clustering
	Database software	SQL Server 2017 Enterprise Edition configured as a Failover Cluster Instance
Storage	Array	PowerVault ME4024 (firmware: GT275R0003-01)
	I/O cards	8 x 16Gbps FC ports (four per controller)
	Disk drives	24 x 1.92TB SSDs (2.5-inch SAS)

3 Hardware components

3.1 PowerEdge R640 server

The PowerEdge R640 server is a two-socket 1U rack server with impressive processor performance, a large memory footprint, extensive I/O options, and a choice of dense, high performance storage or low-cost, high-capacity storage. The R640 server offers simplified management, purposeful design, and energy efficiency with support for the latest Intel® Xeon® series processors as well as NVDIMM memory. The R640 server supports up to three Gen3 PCIe expansion cards. These features give it the ideal mix of performance, cost, and density for most data centers.

For more information, see the [PowerEdge R640 server product page](#).

3.2 QLogic QLE2662 16Gbps host bus adapter

The reference architecture makes use of two QLogic QLE2662 dual-port 16Gbps Fibre Channel (FC) to PCI Express host bus adapters.

For more information, see the QLogic FAQ document, [QLogic 2600 Series 16Gb Gen 5 FC HBAs for Dell PowerEdge Servers](#).

3.3 PowerVault ME4024 array

The Dell EMC PowerVault ME4 Series is next-gen, entry-level storage that is purpose-built and optimized for SAN and DAS virtualized workloads. Available in 2U or dense 5U base systems, the low-cost ME4 Series simplifies the challenges of server capacity expansion and small-scale SAN consolidation with up to 336 drives or 4PB capacity. It also comes with all-inclusive software, incredible performance, and built-in simplicity with a new web-based HTML5 management GUI, ME Storage Manager. Connecting the PowerVault ME4 Series 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: The ME4 Series includes a web-based management GUI (HTML5), installs in 15 minutes, configures in 15 minutes, and easily deploys in 2U or 5U systems.

Performance: Compared to the predecessor MD3 Series, the ME4 Series packs a lot of power and scale with the Intel® Xeon® processor D-1500 product family. The ME4 Series processing power delivers incredible performance gains over the MD3 Series, as well as increased capacity, bandwidth, and drive count.

Connectivity: The ME4 Series goes to the next level with robust and flexible connectivity starting with a 12Gb SAS back-end interface, and a front-end interface options including four 16Gb FC ports per controller, four 10Gb iSCSI ports per controller (SFP+ or BaseT), 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 (ME4012 and ME4024) and 5U (ME4084) 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: The ME4 Series 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 (ME Storage Manager) is included.

For more information, see the [ME4 Series product page](#).

4 Storage configuration

4.1 Cabling

The hardware components were connected using Dell EMC best practices.

In a direct-connect configuration, each HBA should have one port connected to the top storage controller and the other port connected to the bottom storage controller (see Figure 3).

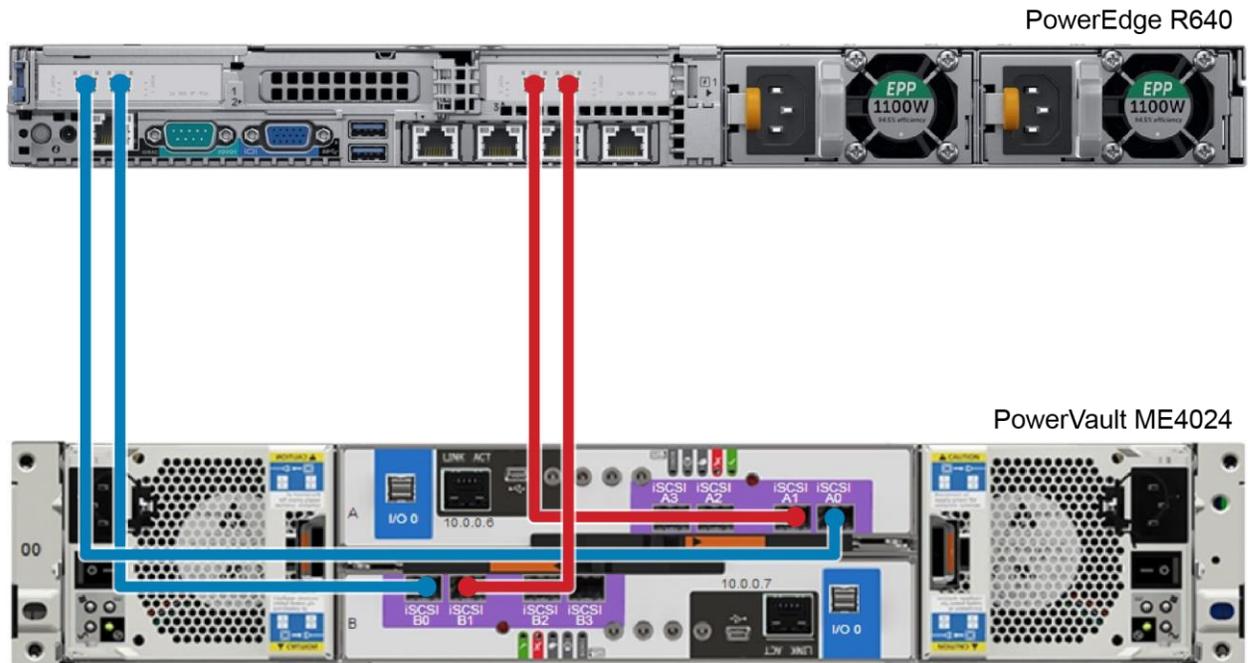


Figure 3 Single server configuration cabling diagram

4.2 Storage pools

Two virtual storage pools were created with twelve disks each and assigned a data protection type of RAID-5. RAID-5 data protection is used for this solution because it provides maximum capacity while still maintaining great read performance.

With the ME4 Series, a storage pool is owned by a single controller. Therefore, creating two storage pools allows the workload to be distributed across both controllers.

4.3 Host setup

The process of setting up a host object creates the I/O path (or paths) between a volume and a server. Each FC WWN will display as an initiator. In this configuration, all four discovered FC initiators (ports) are being added to the new host (see Figure 4).

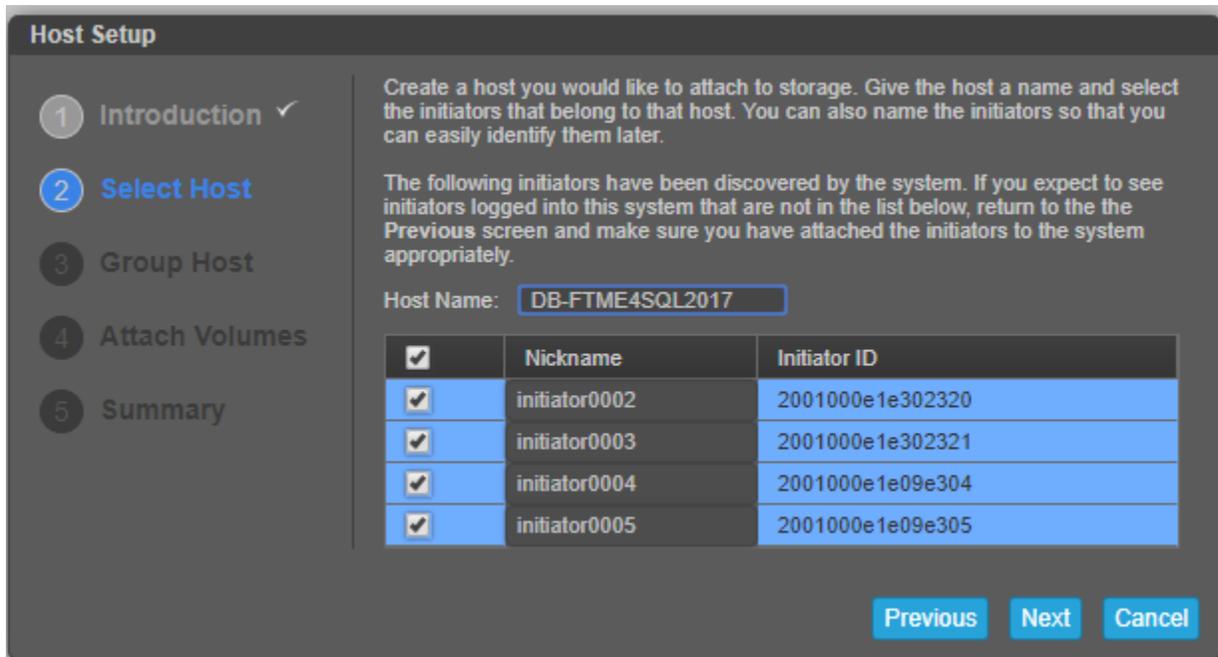


Figure 4 Host setup

4.4 ME4 Series volumes

None of the volumes were configured to use snapshots. This reference architecture does not include hardware resources to support the use of snapshots.

Eight volumes (LUNs) were created to store the SQL Server data files for the data warehouse, with four volumes assigned to each pool. Two volumes were created to store the tempdb data files, with one volume assigned to each pool. When volumes are created consecutively, the array automatically distributes the volumes evenly across the pools. Otherwise, the pool can be manually selected when creating volumes individually. To achieve optimal results, the SQL Server data and tempdb volumes must be evenly distributed across the pools, effectively balancing the workload across both storage controllers.

Table 3 ME4 Series volumes created for the reference architecture

Volume name	Pool	Volume purpose
MPHost	A	Mount point host
SQLSystem	B	SQL Server system databases and files
SQLLog	A	Transaction log files for tempdb and the data warehouse
SQLData01	A	Data files for the data warehouse
SQLData02	B	Data files for the data warehouse
SQLData03	A	Data files for the data warehouse

Volume name	Pool	Volume purpose
SQLData04	B	Data files for the data warehouse
SQLData05	A	Data files for the data warehouse
SQLData06	B	Data files for the data warehouse
SQLData07	A	Data files for the data warehouse
SQLData08	B	Data files for the data warehouse
SQLTempDBData01	A	Data files for tempdb
SQLTempDBData02	B	Data files for tempdb

5 Server configuration

5.1 System BIOS

The system profile is set to **Performance**. All other options, outside of iDRAC configuration, were left at their factory default settings. The **Logical Processor** option, under **Processor Settings**, is left at its default setting of **Enabled**. This enables Intel® Hyper-Threading Technology, which maximizes the number of logical processors available to SQL Server.

5.2 Host bus adapters

After each HBA port was reset to the factory defaults, several parameters were changed in accordance with Dell EMC best practices for ME4 Series storage arrays. Those changes are detailed in Table 4.

Table 4 QLogic HBA parameter changes

Parameter menu	Parameter	Setting
Adapter Settings	Connection Options	1 (Point to Point only)
	Login Retry Count	60
	Port Down Retry Count	60
	Link Down Timeout	30

6 Windows Server 2016 configuration

6.1 Power plan

To maximize performance, the server was configured to use the **High performance** power plan as shown in Figure 5.

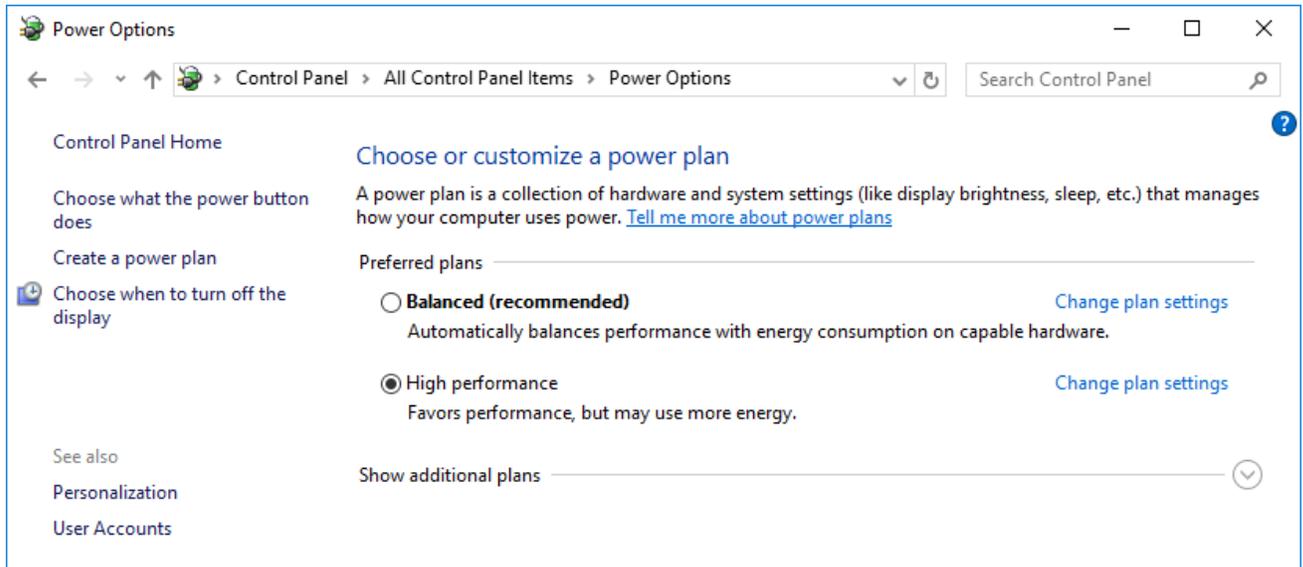


Figure 5 Windows power plan

6.2 Lock pages in memory

To prevent Windows from paging SQL Server memory to disk, the **Lock pages in memory** option was enabled for the SQL Server service account.

For information on enabling this option, visit <https://msdn.microsoft.com/en-IN/library/ms190730.aspx>.

6.3 Windows volumes

A single Windows volume was created on each ME4 Series volume. All volumes were formatted with the NTFS file system. The mount point host volume used the default allocation unit. All other volumes used an allocation unit of 64KB.

For DWFT reference architectures, Dell EMC recommends using mount points for the volumes instead of drive letters. It is highly recommended to assign appropriate volume and mount-point names to simplify troubleshooting and performance analysis. Ideally, the mount-point names should be assigned in a way that makes it easy to identify the ME4 Series volume for a given Windows volume.

Table 5 shows the volume labels and access paths used for the reference configuration.

Table 5 Windows volume details

ME4 Series volume name	Windows volume label	Access path
MPHost	MPHost	M:\
SQLSystem	SQLSystem	M:\ft\SQLSystem
SQLLog	SQLLog	M:\ft\SQLLog
SQLData01	SQLData01	M:\ft\SQLData01
SQLData02	SQLData02	M:\ft\SQLData02
SQLData03	SQLData03	M:\ft\SQLData03
SQLData04	SQLData04	M:\ft\SQLData04
SQLData05	SQLData05	M:\ft\SQLData05
SQLData06	SQLData06	M:\ft\SQLData06
SQLData07	SQLData06	M:\ft\SQLData07
SQLData08	SQLData08	M:\ft\SQLData08
SQLTempdb01	SQLTempdb01	M:\ft\SQLTempdb01
SQLTempdb02	SQLTempdb02	M:\ft\SQLTempdb02

6.4 MPIO

The MPIO policy for all volumes was left at the default setting. Since the ME4 Series supports Asymmetric Logical Unit Access (ALUA), the default MPIO setting is **Round Robin with Subset**.

7 SQL Server 2017 Enterprise Edition configuration

7.1 Grant perform volume maintenance task privilege

During installation of SQL Server 2017, the option to grant the **SQL Server Database Engine Service** the **Perform Volume Maintenance Task** privilege was selected.

7.2 SQL Server maximum memory

The maximum server memory for this reference architecture should be set to 864GB which leaves 32GB for the operating system.

7.3 Max degree of parallelism (MAXDOP)

The max degree of parallelism was set to **0**.

For more information, see the Microsoft article, [Configure the max degree of parallelism Server Configuration Option](#).

7.4 Resource governor

The resource governor was used to limit the maximum memory grant to **12 percent**.

For information about the resource governor, see the Microsoft article, [Resource Governor](#).

7.5 Database configuration

The data warehouse database was configured to use multiple file groups, each containing four files distributed evenly across the four data volumes. All files were allowed to grow automatically. The file groups were configured with the **AUTOGROW_ALL_FILES** option to help ensure that all files within a given file group remain the same size.

7.6 Tempdb configuration

The tempdb database was configured to use eight data files of equal size. The data files were evenly distributed across the two tempdb data volumes, with four files stored on each volume. The tempdb transaction log file was placed on the log volume. All files were expanded to the appropriate size and auto grow was enabled.

8 Additional considerations for the highly available reference architecture

The highly available (HA) reference architecture leverages Windows Failover Clustering to achieve high availability. When configuring a Windows failover cluster, there are additional storage considerations:

- The recommended quorum configuration is to allow all cluster nodes to have quorum votes and use a disk witness. An additional volume needs to be created and configured as the disk witness. Dell EMC recommends using a 2GB volume for the disk witness. For more information on quorum and voting configurations in a failover cluster, see the Microsoft article, [Configure and manage quorum](#).
- All volumes need to be mapped to each node of the cluster.
- All volumes need to be configured as a cluster resource and added to the SQL Server cluster resource group.

9 DWFT for SQL Server 2017 certification

DWFT Certification #2017-017	Dell PowerVault ME4024 with Dell PowerEdge R640 DWFT Reference Architecture			Report Date: 8/15/2018	
DWFT Rev. 5.4					
System Provider	System Name	Processor Type	Memory		
	Dell PowerEdge R640	Intel Xeon Gold 6126 2.6 GHz (2/24/48)	1024 GB		
Operating System		SQL Server Edition			
Windows Server 2016		SQL Server 2017 Enterprise Edition			
Storage Provider	Storage Information				
	24 x 960GB SSD for data, log, tempdb, and OS (RAID 5)				
Primary Metrics					
Rated User Data Capacity ¹ (TB)	Row Store Relative Throughput ²	Column Store Relative Throughput ³	Maximum User Data Capacity ¹ (TB)		
55	184	367	66		
Row Store					
Relative Throughput ²	Measured Throughput (Queries/Hr/TB)	Measured Scan Rate Physical (MB/Sec)	Measured Scan Rate Logical (MB/Sec)	Measured I/O Throughput (MB/Sec)	Measured CPU (Avg.) (%)
184	223	4,844	5,971	5,407	91
Column Store					
Relative Throughput ²	Measured Throughput (Queries/Hr/TB)	Measured Scan Rate Physical (MB/Sec)	Measured Scan Rate Logical (MB/Sec)	Measured I/O Throughput (MB/Sec)	Measured CPU (Avg.) (%)
367	2,676	2,214	N/A	N/A	96
<p>The reference configuration is a 2 socket system rated for 25TB using SQL Server 2014 and the DWFT V4 methodology</p> <p>¹ Assumes a data compression ratio of 5:1</p> <p>² Percent ratio of the throughput to the row store throughput of the reference configuration.</p> <p>³ Percent ratio of the throughput to the column store throughput of the reference configuration.</p> <p>[*] Reported metrics are based on the qualification configuration which specifies database size and SQL Server memory.</p>					

10 Summary

Dell EMC, in partnership with Microsoft, enables customers to deploy tested and validated data warehouse solutions using Data Warehouse Fast Track reference architectures for SQL Server 2017. These uniquely designed architectures ensure optimal business intelligence solutions. The end-to-end best practices and recommendations enable the customer to achieve enhanced return on investment, and faster time to value with a balanced data warehouse environment that can perform better than traditional data warehouse systems.

Dell EMC DWFT reference architectures provide the following benefits:

- Deliver a tested and validated configuration with proven methodology and performance behavior
- Deliver outstanding performance on the Dell EMC PowerEdge 14G server platform with blazing processor speeds and leading-edge, flash-based Dell EMC storage arrays
- Achieve a balanced and optimized solution at all levels of the stack by following best practices for both hardware and software components, achieving faster time to value, and lower total cost of ownership
- Avoid over provisioning of hardware resources
- Offer high availability at all levels of setup (host, switches, and storage)
- Offer a single point of contact and accountability for purchases, services, and support; SQL Server is available to purchase from Dell EMC worldwide
- Help customers avoid the consequences of an improperly designed and configured system
- Reduce future support costs by limiting solution rearchitect efforts due to scalability challenges

This paper describes a reference architecture using a PowerEdge R640 server with a PowerVault ME4024 storage array. By implementing Data Warehouse Fast Track for SQL Server 2017 design principles, this configuration achieved a 65TB rating.

A Technical support and resources

[Dell.com/support](https://dell.com/support) is focused on meeting customer needs with proven services and support.

[Storage solutions technical documents](#) provide expertise that helps to ensure customer success on Dell EMC storage platforms.

The following ME4 Series publications and additional resources are available at [Dell.com/support](https://dell.com/support).

- Administrator's Guide
- Deployment Guide
- CLI Guide
- Owner's Manual
- Support Matrix

See the following additional resources:

- [Dell EMC products](#)
- [Dell SQL Server solutions](#)
- [Dell EMC PowerVault ME4 Series](#)
- [Dell EMC PowerEdge R640](#)