CISCO

PURESTORAGE®

REFERENCE ARCHITECTURE

Microsoft SQL Server 2016 Data Warehouse Fast Track

FlashStack 23TB Solution with Cisco UCS and Pure Storage FlashArray

CISCO UCS C240 M4 RACK SERVER

January 2017

FlashStack

TABLE OF CONTENTS

MODERN DATA WAREHOUSE

Data has become the life blood of organzations, and the amount of data is ever increasing with the types of data sources. The challenge is turning this data into meaningful information quickly to make informed business decisions. Business units can stress hardware infrastructures with long running queries, large volume data loads, and data aggregation.

Microsoft® SQL Server® 2016 introduces several improvements to accommodate the challenges of storing and processing these large amounts of data. SQL Server 2016 provides up to 10X compression of data utilizing column-store index tehnologies. This column-store technology provides the ability to scan only the columns needed while reducing IO requirements and memory required for a given number of rows form the source data warehouse. Efficient single-row lookup is improved by using additional (B-Tree) indices to columnstore-based tables.

Query optimization is also improved by Operator Pushdown which moves both filter and aggregation query operations closer to the data, reducing the volume of data which needs to be handled further on in query processing.

SQL Server 2016 provides Batch Mode Processing which processes 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.

RESOURCE UTILIZATION

The infrastructure that runs your SQL Server 2016 Data Warehouse needs to be balanced with regard to performance. Installing the fastest processors without considering the configuration as a whole, storage, memory, switch, etc. can lead to disappointing long-term results, as you need to scale. This testing on your own can be costly and time consuming.

The SQL Server 2016 Data Warehouse Fast Track program is a reference architecture designed to take the guessing out of building your data warehouse infrastructure. These reference architectures are already tested using bandwidth demanding workloads to meet specific query performance and scale in size requirements designated by the Microsoft SQL Server Performance team.

The FlashStack™ solution, an integrated infrastructure from Cisco® and Pure Storage®, is designed with the philosophy to deliver meaningful performance with a focus on optimization of resources, not just a single benchmark score or metric. Nobody wants their expensive processors waiting around for work. Using the Cisco UCS® C240 M4 Rack Server for compu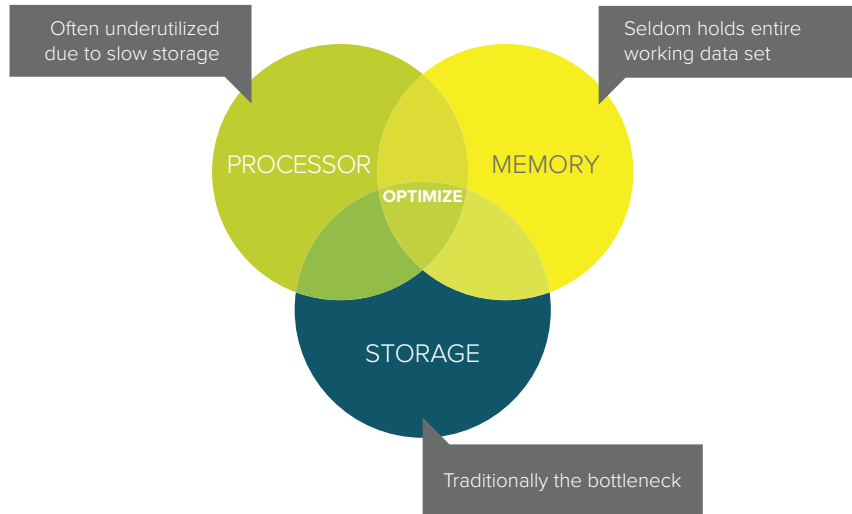te and the Pure Storage FlashArray/M product with solid state flash storage helps eliminate storage as the bottleneck and achieve a high utilization of all resources.

Cisco and Pure Storage have a history of collaborating to offer solutions that scale, are simple to manage, and maintain high performance. Pure Storage and Cisco have introduced Cisco Validated Designs on FlashStack Data Center offerings. Pure Storage has also introduced a number FlashStack solutions covering key data center workloads including Microsoft SQL Server, and Microsoft Exchange.

SOLUTION SUMMARY

Problem

Architectures based on a "build-your-own" method are time consuming and risky often resulting in poor resource utilization, system imbalance, and scaling problems.

Solution

The Pure Storage FlashArray//M and Cisco® Data Warehouse Fast Track reference architecture. Tested with SQL Server 2016 for Row Store and Columnstore workloads.

Results

A balanced system with efficient query processing predictable performance. 22TB rated user capacity. Row store IO throughput 3,530MB/s. Columnstore throughput 1,710 queries/Hr/TB.

Resource Utilization Database Environment

Often underutilized due to slow storage

Seldom holds entire working data set

PROCESSOR

MEMORY

OPTIMIZE

STORAGE

Traditionally the bottleneck

WORKLOAD OPTIMIZATION

Pure Storage FlashArray//M20

The Pure Storage FlashArray//M performance can make your business smarter by unleashing the power of real-time analytics, driving customer loyalty, and creating new, innovative customer experiences that simply weren't possible with disk-based storage. All by transforming your storage with the FlashArray//M.

With the FlashArray//M, organizations can dramatically reduce the complexity of storage to make IT more agile and efficient.

Cisco UCS C240 M4

Cisco MDS 9148S

The Cisco UCS C240 M4 Rack Server used in this reference architecture is our 2-socket, 2-rack-unit (2RU) rack server It offers outstanding performance and expandability for a wide range of storage and I/O-intensive infrastructure workloads, from big data to collaboration.

The enterprise-class Cisco UCS C240 M4 server extends the capabilities of the Cisco Unified Computing System™ (Cisco UCS) with the addition of the Intel® Xeon® processor E5-2600 v4 and v3 product family.

The Cisco MDS 9148S 16G Multilayer Fabric Switch is the next generation of the highly reliable Cisco MDS 9100 Series Switches. It includes up to 48 auto-sensing line-rate 16-Gbps.

In all, the Cisco MDS 9148S is a powerful and flexible switch that delivers high performance and comprehensive Enterprise-class features.

TEST CONFIGURATION

COMPUTE	DESCRIPTION
Vendor	Cisco
Model	Cisco UCS C240 M4
Processor	Qty.(2) Intel Xeon processor E5-2630 v4 20 Core
DRAM	512GB* Tested with max 118GB for scaling estimate purposes
Qlogic QLE2672	Qty.(2) 16Gb FC HBA
BIOS Power	High Throughput Enabled

TEST CONFIGURATION

STORAGE	INFORMATION/PARAMETER
Model	FA-m20 Qty.(4) 8Gb/s ports
Operating Environment	Purity 4.7.5
Total Array Capacity	20 TB
STORAGE	**INFORMATION/PARAMETER**
OS Boot Volume	500GB (operating system)
Mount Point Anchor	20GB (mount point for volumes)
System	50GB (volume for system database files.)
Log	1TB (volume for database log files.)
Data and tempdb	8TB (volume for database data and tempdb files.)
Backup	6TB (volume for database backup files.)
NETWORK	**INFORMATION/PARAMETER**
Fabric Switch	Qty.(2) Cisco MDS 9148S
WORKLOAD	**INFORMATION/PARAMETER**
OLAP Type Workload	Read/Write Ratio 80/20

TEST CONFIGURATION

APPLICATION	INFORMATION/PARAMETER
Microsoft SQL Server Enterprise	2016 13.0.1708.0 (X64)
Trace Flags	-E, -T1117, -T834
Min/Max Memory	118(MB)/118(GB)
OPERATING SYSTEM	**INFORMATION/PARAMETER**
Windows Server 2012 R2 w/Updates	Build 6.3.9600
Power Management	High Performance Enabled
Local Security Policy	Lock Pages in Memory Enabled
Perform Volume Maintenance Tasks	Instant File Initialization Enabled

RATED USER CAPACITY: 23TB

ROW STORE THROUGHPUT: 3,530 MB/s

Avg. CPU Utilization   89%

Columnstore Throughput: 1,710 /Queries/Hr/TB

Avg. CPU Utilization   100%
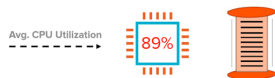
RESULTS

Data Warehouse workloads can range from loading large amounts of data to complex analytical processing of data for consumption. The I/O generated typically consists of concurrent streams of read operations. The SQL Server 2016 DWFT reference architecture using Cisco for compute and networking, and Pure Storage for fast solid state flash storage establishes an architecture ready to accommodate the most demanding complex read queries and scale requirements.

FlashStack

SUMMARY

Pure Storage greatly values its partnerships with Microsoft and Cisco. This collaborative approach creates a reference architecture solution that meets the performance and scaling needs of today's modern data warehouse.

Working together this decreases the time and implementation complexity for customers to create a balanced, reliable long-term solution.

DWFT REFERENCE ARCHITECTURE

DWFT Certification: #2016-009	Pure Storage FlashArray//M20 Cisco UCS C240 M4		Report Date: 9/23/2016
DWFT Rev. 5.4			

SYSTEM PROVIDER	SYSTEM NAME	PROCESSOR TYPE	MEMORY
CISCO	Cisco UCS C240 M4	Intel Xeon Processor E5-2630 V4 2.2 GHz (2/20/40)	384 GB

OPERATING SYSTEM	SQL SERVER EDITION
Windows Server 2012 R2	SQL Server 2016 Enterprise Edition

SYSTEM PROVIDER	STORAGE INFORMATION
PURESTORAGE®	FlashArray//M20 20TB Aggregated This Provisioned 8TB for Data & Tempdb, 1TB fr LOG, 500GB for OS boot

PRIMARY METRICS			
Rated User Data Capacity[1] (TB)	Row Store Relative Throughput[2]	Columnstore Relative Throughput[3]	Maximum User Data Capacity[1] (TB)
23	120	240	26

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.) (%)
120	146	3,078	3,983	3,530	89

COLUMNSTORE					
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.) (%)
240	1,710	2,357	N/A	N/A	100

The reference configuration is a 2 socket system rated for 25TB using SQL Server 2014 and the DWFT V4 methodology.

1 Assumes data compression ratio of 5:1

2 Percent ratio of the throughput to the row store throughput of the reference configuration.

3 Percent ratio of the throughput to the Columnstore throughput of the reference configuration.

*Reported metrics are based on the qualification which specifies database size SQL Server memory.

FlashStack

FLASHSTACK@PURESTORAGE.COM
WWW.PURESTORAGE.COM/CISCO
WWW.PURESTORAGE.COM/FLASHSTACK
WWW.CISCO.COM/GO/FLASHSTACK