Cisco UCS C240 M4 Rack Server with Lightning Ascend™ Gen. II SAS SSDs from SanDisk® 36TB Data Warehouse Fast Track Reference Architecture

Based on the SQL Server 2016 Data Warehouse Fast Track (DWFT) Reference Architecture

SOLUTION ID: SQL2016-Data-Warehouse-Fast-Track-36TB-SD

Table of Contents

Executive Summary

This guide details the server, storage, and software configurations for the Cisco UCS® C240 M4 Rack Server with Lightning Ascend Gen. II 12Gb/s SAS SSDs (1.6TB) from SanDisk which Microsoft® has validated for a Fast Track-rated user capacity of 36TB.

This document is for individuals (BI Architects, DBAs, Report-Developers, and IT Directors) involved in decision making who are looking for guidance when designing enterprise, business-intelligence applications.

The Microsoft SQL Server® Data Warehouse Fast Track (DWFT) reference architecture is designed to eliminate the complexity of properly sizing hardware, which helps reduce unnecessary scale-out of storage and servers. The sizing techniques used in SQL server DWFT will properly size servers, based on I/O and CPU consumption. This consumption-based approach ensures your data warehouse can fully take advantage of your hardware investment.

About the Cisco UCS C240 M4 Rack Server

Storage and I/O-Optimized Cisco UCS C-Series Rack Servers

The Cisco UCS C240 M4 Rack Server is a 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) portfolio in a 2RU form factor with the addition of the Intel® Xeon® processor E5-2600 v4 and v3 product family, which delivers a superb combination of performance, flexibility, and efficiency.

Product Overview

The Cisco UCS C240 M4 offers outstanding levels of internal memory and storage expandability with exceptional performance. It has the following features:

- Up to 24 DDR4 DIMMs at speeds of up to 2400 MHz, for improved performance and lower power consumption

- Up to 6 PCI Express (PCIe) 3.0 slots (4 full-height, full-length)

- Up to 24 small form factor (SFF) drives or 12 large form factor (LFF) drives, plus two (optional) internal SATA boot drives

- Support for 12-Gbps SAS drives

- A modular LAN-on-motherboard (mLOM) slot for installing a next-generation Cisco virtual interface card (VIC) or third-party network interface card (NIC) without consuming a PCIe slot

- 2 x 1 Gigabit Ethernet embedded LOM ports

- Support for up to two double-wide NVIDIA graphics processing units (GPUs), providing a graphics-rich experience to more virtual users

- Excellent reliability, availability, and serviceability (RAS) features with tool-free CPU insertion, an easy-to-use latching lid, hot-swappable and hot-pluggable components, and redundant Cisco® Flexible Flash (FlexFlash) SD cards

For additional details, refer to the following datasheet:

http://www.cisco.com/c/en/us/products/collateral/servers-unified-computing/ucs-c240-m4-rack-server/datasheet-c78-732455.html

Lightning® Product Family of SAS SSDs from SanDisk

The Lightning Ascend Gen. II 12Gb/s SAS SSD doubles interface speed, providing the highest performance for mission-critical hyperscale, cloud, and virtualized data center application workloads. This product offers a feature-rich robust design, including the following:

- SanDisk's innovative parallel processing architecture

- Full data path protection (T10-DIF support)

- Temperature-based power control

- SED instant secure erase

- MTBF of 2.5 million hours

- Power fail protection

The SSD is backward-compatible with 6Gb/s SAS and offers a single firmware binary platform for seamless integration and reduced qualification times. A summary of features is shown below:

Lightning Ascend™ Gen. II 12Gb/s SAS 2.5" SSDs	
	200GB – 1.6TB
Performance/IOPS	
Data Throughput (Sequential Read/Write)[1]/*	Up to 980/600 MB/s
Total IOPS* (Random Read) 4KB – (single port)	Up to 195,000 IOPS
Total IOPS* (Random Write) 4KB – (single port)	Up to 100,000 IOPS
Interface, Capacity, and Memory Technology	
Host Interface	12Gb/s SAS
Interface Ports	Dual (x2 wide) full-duplex
Form Factor	2.5 inch
19nm eMLC User Capacities[2]	200GB, 400GB, 800GB, 1.6TB
Sector Sizes Supported	512, 520, 4K bytes/sector T10-DIF 1, DIF 2 Support
Reliability	
Unrecovered Bit Error Rate (UBER)	<1 sector in 10^{18} bits read**
End-to-End Data Protection	No performance impact
Patrol Read (automatic)	No performance impact
Power-on Hours	8760 hours per year
MTBF[3]	2.5 million hours
Warranty[4]	Lesser of 5 years or maximum endurance used
Endurance	
DWPD (Random Read)	10 DWPD

For more information, visit

https://www.sandisk.com/business/datacenter/products/flash-devices/ssds/sas-ssd/lightning-gen2

About the Data Warehouse Fast Track Reference Architecture

The SQL Server Data Warehouse Fast Track reference architecture provides a scalable framework based on balancing I/O to achieve maximum performance from SMP-based servers. SQL Server Data Warehouse Fast Track eliminates the complexity of sizing servers with data warehouses by providing a set of data consumption rates that properly balances performance between the disk subsystem, CPU, and memory.

More information on SQL Server DWFT can be found here:

http://www.microsoft.com/en-us/server-cloud/data-warehouse-fast-track.aspx

Microsoft SQL Server 2016: Data Warehousing with Improved Column Store Technology

Microsoft SQL Server 2016 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, but also significantly increased performance, due to the dramatically reduced IO requirements given by the compression, 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 2016 adds the capability to add additional (B-Tree) indices to column store-based tables, which enables efficient single-row lookup.

In addition to these architectural features, we have 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 which needs to be handled further on in query processing.

Batch Mode Processing
SQL Server 2016 includes enhancements in 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.

Reference Architecture

The following table shows the configuration details for the Cisco UCS C240 M4 Rack Server with Lightning Ascend Gen. II SAS SSDs, using the Fast Track Data Warehouse architecture.

Server	Cisco UCS C240 M4 Rack Server SFF
Operating System	Microsoft Windows Server 2012 R2 Standard Edition
CPU	Intel Xeon E5-2697 v3 @ 2.6 GHz (2S/28C/56T)
PCI-E Slots	Up to six PCIe 3.0 slots
Drives	Up 24 SFF drives with 12Gps SAS drive support
Storage Controller	Single UCSC-MRAID12G with SAS 24-drive backplane
Drives	2 x 1.2TB SAS (OS)
RAM	512GB

Server UEFI Configuration

- Hyper-Threading was enabled.
- Operating mode was changed to "Maximum Performance".
- Fan Offset was set to "Increased Cooling".

General Settings

Operating System Settings

This Fast Track Data Warehouse test used the Microsoft Windows Server 2012 R2 Standard Edition OS. Standard installation steps were used to install the operating system with default values, followed by service packs and update patches.

Windows Configuration – Power Settings

The High Performance plan was chosen to reduce CPU throttling.

Storage Configuration

The Windows 2012R2 OS was configured on a pair of 1.2TB 6G SAS 10K SFF HDDs (mirrored).

The Lightning Ascend Gen. II SAS SSDs were configured using Storage Spaces. This technology allows Windows Server to virtualize storage by grouping industry-standard disks into storage pools and then creating virtual disks ("storage spaces") from the available capacity in those storage pools.

Storage Spaces enable cost-effective, highly available, scalable and flexible storage solutions for business-critical deployments.

This server is available with a single Cisco 12G SAS Modular RAID controller, configured to pass-through mode.

Physical Disk Storage Pool

Total Drives	Slots	BUS	Physical Disks	Presented Capacity per drive	Total presented capacity	Storage Pool ID
12	3-14	SAS	Lightning Ascend Gen. II SAS SSDs	1.45TB	17.5TB	SQL

SQL Storage Pool - Virtual Disks

Name	Layout	Provisioning	Capacity	Path
SQL Data	Parity	Fixed	12,000GB	C:\DB\DATA
SQL Logs	Mirror	Fixed	600GB	C:\DB\LOGS

SQL Server Settings

Database Configuration

A 1TB data warehouse schema was created for benchmarking using the Fast Track toolkit. The schema used a master filegroup with 4 additional filegroups.

TempDB Configuration

In total, eight 10GB tempdb files were created and stored on the volume designated for data files (C:\DB\DATA). The tempdb transaction log file was stored on the volume designated for log files (C:\DB\LOGS).

Memory Allocation

SQL Server was allocated 118GB of the available server memory (512GB). This amount is allocated as part of the Fast Track test criteria to drive backend disk activity during the Row Store query runs.

Local Security Policy

The SQL Server maintenance account was granted the following privileges:

- Enable Lock Pages in Memory - prevents SQL Server buffer pool pages from paging out.

- Perform Volume Maintenance Tasks - enables Instant File Initialization.

SQL Server 2016 Configuration Parameters

Parameter	Setting	Description
Memory Allocation	118GB	This is the Fast Track-required value for a 2-socket, 1TB database. Memory is deliberately constrained to enforce I/O pressure on the subsystem.
Max Degree of Parallelism, Row Store	28	When SQL Server runs on a computer with more than one microprocessor or CPU, it detects the best degree of parallelism (the number of processors that are used in the execution of a parallel plan).
Max Degree of Parallelism, Column Store	56	
Resource Governor Memory Allocation: Row Store, Column Store	12%	Default is 25%. This is reduced to 12% for both Row Store and Column Store to reduce the maximum memory consumed per query.
Fast Track - Required Start-up Parameters	-T1117	This trace flag ensures even growth of all files in a file group in case `autogrow` is enabled. The standard FTDW recommendation for database file growth is to pre-allocate rather than use `autogrow`(with the exception of tempdb).
Optional Startup Parameters	-T1118	Helps alleviate allocation bit map contention in tempdb by switching allocations to full extents (8 physically contiguous pages (64KB)).
Start-up Parameters - Row Store only	-T834	When this flag is set, SQL Server uses Windows large-page memory allocations for the buffer pool. This trace flag can improve throughput rates for many data warehousing workloads. This value is disabled for Column Store runs.

Measured Performance

During Fast Track Database Validation, Microsoft's Reference Point tool drives multiple concurrent query workloads designed to identify bottlenecks. The tool establishes the key performance metrics in the table below.

Scan Rate Type	Scan rate	Description
Rated User Capacity (TB)	36	Represents the optimal Fast Track-certified data capacity of the configuration.
		Allows for 5:1 compression with 10% recommended free space. 25% capacity is reserved for TempDB, while some memory and throughput based limits are also applied.
Row Store Relative Throughput	188	Percentage throughput of this configuration in comparison to the FTRA configuration. This result almost matched the reference configuration, using only a single socket. The reference architecture is a 25TB dual-socket configuration.
Column Store Relative Throughput	364	Percentage throughput of this configuration in comparison to the FTRA configuration.
Maximum User Data Capacity (TB)	41	Calculated, based on total disk capacity. Allows 5:1 compression. It factors recommended free space but ignores the throughput limits that are applied to the Rated User Capacity.
RS Measured Throughput (Q/Hr/TB)	244	Number of Row Store combined benchmark queries completed during the measurement interval. This is normalized to the 1TB database.
CS Measured Throughput (Q/Hr/TB)	2365	Number of Column Store combined benchmark queries completed during the measurement interval. This is normalized to the 1TB database.
Row Store Measured I/O Throughput (MB/S)	5547	Average of the Physical and Logical scan rate, which demonstrates the throughput capability of the drives.
		During maximum load, the physical read throughput peaked at 6000 MB/sec. This server was configured with a single Cisco 12G SAS Modular RAID controller.

SQL Server Data Warehouse Fast Track Certification

DWFT Certification #2016-005	Cisco C240-M4 with Lightning Ascend from SanDisk 36TB reference architecture for Microsoft SQL Server 2016 **DWFT Reference Architecture**		Report Date: 6/13/2016
DWFT Rev. 5.4			
System Provider	**System Name**	**Processor Type**	**Memory**
CISCO	Cisco C240-M4	Intel Xeon E5-2697 v3 @2.6GHz (2/28/56)	512 GB
Operating System		**SQL Server Edition**	
Windows Server 2012 R2		SQL Server 2016 Enterprise Edition	
Storage Provider	**Storage Information**		
SanDisk a Western Digital brand	12 x 1.6TB SanDisk Lightning Ascend - presented via Storage Spaces 12000GB allocated to Data and TempDB (Parity) 600GB allocated to LOG (Mirror)		

Primary Metrics

Rated User Data Capacity[1] (TB)	Row Store Relative Throughput[2]	Column Store Relative Throughput[3]	Maximum User Data Capacity[1] (TB)
36	188	364	41

Row Store

Relative Throughput[2] (Queries/Hr/TB)	Measured Throughput	Measured Scan Rate Physical (MB/Sec)	Measured Scan Rate Logical (MB/Sec)	Measured I/O Throughput (MB/Sec)	Measured CPU (Avg.) (%)
188	244	4,849	6,246	5,547	72

Column Store

Relative Throughput[2] (Queries/Hr/TB)	Measured Throughput	Measured Scan Rate Physical (MB/Sec)	Measured Scan Rate Logical (MB/Sec)	Measured I/O Throughput (MB/Sec)	Measured CPU (Avg.) (%)
364	2,365	2,182	N/A	N/A	98

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

[1] Assumes a 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 column store throughput of the reference configuration.

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

Summary

Together, Cisco and SanDisk dedicated hundreds of hours of testing to engineer the SQL Server DWFT solution to provide the most optimal reliability and performance. These series of tests pushed the Cisco UCS C240 M4 to peak performance without hardware failure. The reliability and performance experienced during testing is what can be expected in production environments.

The same configuration meets the need of both Row Store and Column Store configurations, delivering high measured I/O throughput in the Row Store configuration at 5.5GB/s, and high query rates in the Column Store configuration at 2365 Q/Hr/TB.

Bill of Materials

Cisco UCS C240 M4 – Non High Availability Option

Qty	SKU	Description
1	UCSC-C240-M4SX	UCS C240 M4 SFF 24 HD w/o CPU, mem, HD, PCIe, PS ,railkt w/expndr
10	N20-BBLKD	UCS 2.5-inch HDD blanking panel
2	UCSC-HS-C240M4	Heat sink for UCS C240 M4 rack servers
1	UCSC-SCCBL240	Supercap cable, 250mm
2	UCS-CPU-E52697D	2.60GHz E5-2697 v3/145W 14C/35 MB Cache/DDR4 2133 MHz
16	UCS-ML-1X324RU-A	32GB DDR4-2133-MHz LRDIMM/PC4-17000/quad rank/x4/1.2v
2	UCS-HD12T10KS2-E	1.2TB 6G SAS 10K rpm SFF HDD
1	UCSC-MRAID12G	Cisco 12G SAS Modular Raid Controller
12	UCS-SD16TB12S4-EP	1.6TB 2.5-inch Enterprise Performance 12G SAS SSD (10X endurance)
1	UCSC-MRAID12G-4GB	Cisco 12Gbps SAS 4GB FBWC Cache module (RAID 0/1/5/6)
2	UCSC-PSU2-1400W	1400W AC power supply for 2U & 4U C Series servers
2	CAB-C13-C14-2M	Power Cord Jumper, C13-C14 connectors, 2-meter length
1	UCSC-CMA-M4	Reversible CMA for C240 M4 Ball Bearing Rail Kit
1	UCSC-RAILB-M4	Ball Bearing Rail Kit for C220 M4 and C240 M4 rack servers
1	N20-BKVM	KVM local I/O cable for UCS server console port
1	UCSC-PCI-1B-240M4	Right PCIe Riser Board (Riser 1) (3x8) for 6 PCI slots
1	UCSC-PCI-2-C240M4	Left PCIe Riser Board (Riser 2) for C240 M4
1	CON-SNTP-C240M4SX	SMARTNET 24X7X4 UCS C240 M4 SFF 24 HD w/o CPU, mem
1	UCSC-MLOM-CSC-02	Cisco UCS VIC1227 VIC MLOM - Dual Port 10Gb SFP+
1	C1UCS-OPT-OUT	Cisco ONE Data Center Compute Opt Out Option