White Paper
November 2012

Cisco Reference Configurations for Microsoft SQL Server 2012 Fast Track Data Warehouse 4.0 with EMC VNX5500 Series Storage Systems

Microsoft SQL Server 2012 Fast Track Data Warehouse

Contents

Cisco Reference Configurations for Microsoft SQL Server 2012 Fast Track Data Warehouse 4.0 with EMC VNX5500 Series Storage Systems

White Paper
November 2012

What You Will Learn

Data warehouse solutions that support high capacity and high performance are increasingly essential to business success and competitive advantage. As organizations realize the importance of mining their accumulated data for important business insights, enterprise data warehouses have been growing rapidly, often doubling or tripling in size in a short time. This massive data growth is often accompanied by increased complexity, both in the management of large volumes of data and in making sure that the data is available for high-performance queries that support essential business decisions.

The reference configurations presented in this document address the common business challenges that IT functions face today, including:

- New application deployments to address business growth and opportunities along with growing demands
- Migration from existing proprietary systems to industry-standard hardware and software
- Hardware refresh demands fueled by growing complexity and response-time requirements
- Consolidation of applications from multiple older servers to fewer, more powerful systems to simplify operations and reduce operating expenses

With Microsoft SQL Server 2012, Microsoft continues to be the value and volume leader, offering comprehensive capabilities to support data warehouse applications. Microsoft has introduced the Microsoft SQL Server Fast Track Data Warehouse program to enable organizations to rapidly deploy predictable enterprise data warehousing infrastructure. Microsoft SQL Server Fast Track 4.0—the fourth generation of this initiative—is designed to let organizations quickly and reliably deploy a total system solution, including server, storage, and connectivity for data warehouses based on Microsoft SQL Server 2012. This approach helps dramatically reduce the burden on end users in selecting, sizing, and testing the

right combination of products and technologies while dramatically reducing the risk of failure.

The Microsoft SQL Server Fast Track Data Warehouse 4.0 reference configurations presented in this document are designed, tested, and validated by Cisco, EMC, and Microsoft to provide:

- Architectural guidance for customers, partners, and resellers who are evaluating, planning, or deploying Microsoft SQL Server–based data warehouse solutions
- Performance and capacity guidance for selecting server, storage, and connectivity solutions in which out-of-the box performance and rapid deployment are important
- Lower cost of ownership by using industry-standard hardware and software, eliminating the costs associated with custom performance benchmark testing and proofs of concept

This document introduces a medium-sized enterprise reference configuration and a large enterprise reference configuration, which together are designed to meet a broad range of data warehouse requirements. As of this writing, the reference configurations scale from 8 to 40 terabytes (TB) of recommended user capacity using the compression capabilities available in Microsoft SQL Server 2012.

The reference configurations presented in this document are based on:

- **Cisco UCS® C-Series Rack Servers:** With the intelligent Intel® Xeon® processor E7 family, these rack servers are designed to operate both in standalone environments and as part of the Cisco Unified Computing System™ (Cisco UCS).
- **EMC® VNX5500™ Series storage systems:** With industry-leading performance, EMC VNX5500 Series storage systems offer unified support for block and file usage protocols, a simple and intuitive management interface, five-nines reliability, and built-in features for supporting replication, disaster protection, and automatic storage tiering.
- **Cisco Nexus® 5000 Series Switches:** The Cisco Nexus 5000 Series provides a standards-based unified fabric that carries all I/O modalities, including Ethernet and Fibre Channel, over a single network, significantly reducing infrastructure cost.

Cisco and EMC Reference Configurations for Microsoft SQL Server Fast Track Data Warehouse

As a part of the Microsoft SQL Server Fast Track Data Warehouse program, Cisco and EMC partnered with Microsoft and worked jointly to complete a comprehensive qualification process. This approach helps ensure that the complete solution conforms to Microsoft's guidelines for the Microsoft SQL Server Fast Track Data Warehouse program.

Microsoft SQL Server Fast Track Data Warehouse Program

Optimized for common data warehouse, decision support, and business analytics deployments based on Microsoft SQL Server, the Microsoft SQL Server Fast Track Data Warehouse program allows organizations to deploy tested configurations using

proven technologies that offer balanced solution stacks. By using these reference configurations, new projects can be implemented quickly and confidently, avoiding much of the guesswork and many of the planning choices and difficult decisions involved in traditional approaches. Goals of the program include:

- Accelerated data warehouse projects with pretested and validated hardware and software configurations
- Reduced hardware and maintenance costs as a result of purchasing the right balanced hardware solution that is optimized for data warehouse workloads
- Reduced planning and setup costs as well as less risk as a result of choosing the right scale model, configuring the system correctly, and taking advantage of tuning guidance

Microsoft SQL Server Fast Track Data Warehouse reference configurations help optimize sequential I/O—critical to data warehouse performance—and provide a balanced approach to hardware configuration. Most system designs look at memory, I/O bandwidth, and storage capacity separately when planning a system. Instead, the Microsoft SQL Server Fast Track Data Warehouse reference architecture attempts to balance the overall system so that throughput is matched across the various hardware and software components specifically for data warehouse performance. Main considerations for balanced data warehouse application performance include:

- Computing power from available CPU cores
- Main memory for the use of the Microsoft SQL Server cache
- I/O channel throughput between the server and the storage system, including the combined throughput of the host bus adapters (HBAs) and switching hardware
- Storage system throughput for sequential read operations, caching capabilities, data storage capacity, and redundancy
- Column store index (CSI) technology for greater efficiency in page storage and indexing than is available from row-store approaches

The goal of a Microsoft SQL Server Fast Track Data Warehouse reference configuration is to achieve a cost-effective balance between Microsoft SQL Server data processing capabilities and the capabilities of the underlying hardware components.

Reference Configuration Components

Two reference configurations were designed and tested by Cisco and EMC—a medium-sized enterprise reference configuration and a large enterprise reference configuration—both based on the core components depicted in Figure 1. These reference configurations were intended to balance computing power, storage performance, and capacity while emphasizing the core values of Cisco and EMC. The reference configurations can be deployed as is, or they can be used as starting points for building larger systems. The main components of the reference configurations are described in detail in the sections that follow.

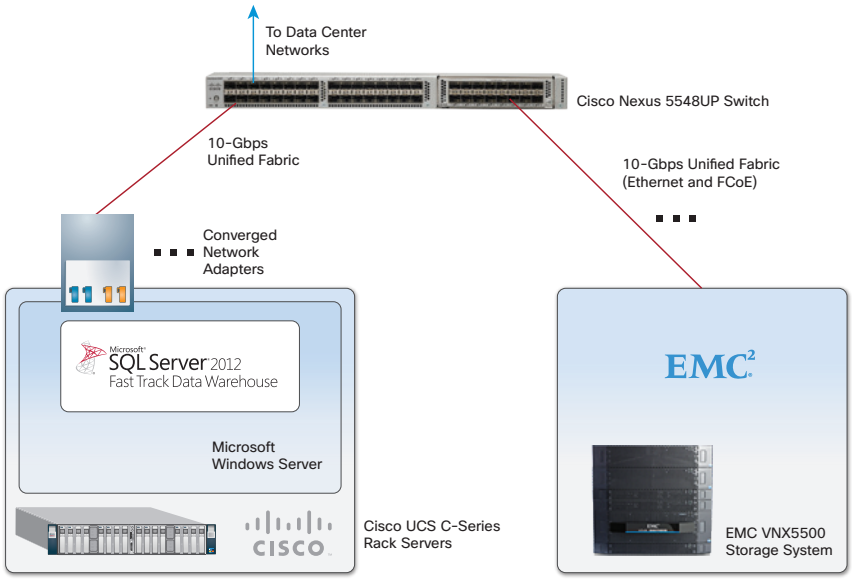

Figure 1. The Reference Configurations Consist of Cisco UCS C-Series Rack-Mount Servers, Cisco Nexus 5548UP Switches, and EMC VNX5500 Storage Systems

Cisco UCS C-Series Rack Servers with Intelligent Intel Xeon Processors

Both reference configurations are powered by Cisco UCS C-Series Rack Servers based on the Intel Xeon processor E7 family. Intel Xeon processors at the core of these systems deliver new levels of scalability and performance to support new business intelligence capabilities in Microsoft SQL Server 2012. New features, such as multibit error correction, provide advanced reliability, availability, and serviceability for enterprise database applications. In addition, the Intel Xeon processor E7 family delivers performance that is well suited for data-demanding workloads by delivering improved scalability and increased memory and I/O capacity. These processors help organizations adapt to rapidly changing short-term business demands, while addressing requirements for longer-term business growth. Advanced reliability and security features work to maintain data integrity, accelerate encrypted transactions, and increase the availability of mission-critical applications.

Designed to operate both in standalone environments and as part of Cisco UCS, Cisco UCS C-Series Rack Servers enable organizations to deploy systems incrementally—using as many or as few servers as needed—on a schedule that best suits the organization's timing and budget. Many organizations prefer rack servers because of the wide range of I/O options available in the form of PCI adapters. Cisco UCS C-Series servers support a broad range of I/O options, including interfaces supported by Cisco as well as adapters from third parties.

When deployed as part of Cisco UCS, these servers can become part of a single converged system whose entire configuration is programmable through Cisco UCS Manager and an XML API. An integral component of Cisco UCS Manager, Cisco UCS service profiles embody the entire system configuration, from firmware revisions and BIOS settings to network profiles. This automation of system

configuration dramatically simplifies and accelerates deployment of enterprise applications such as Microsoft SQL Server and makes rehosting of the software on larger or smaller servers a point-and-click operation.

Moving applications between servers to accommodate larger or smaller workloads is as straightforward as applying a Cisco UCS service profile to an alternative server and rebooting. Similar to standalone deployments, Cisco UCS uses a 10-Gbps unified fabric to support both networking and storage access. In contrast to standalone deployments, integration with Cisco UCS enables end-host mode, allowing the data center network to interact with the system as a single server rather than as an additional hierarchy of components.

Cisco Nexus 5000 Series Switches

As a part of the unified fabric component of the Cisco® Data Center Business Advantage (DCBA) architectural framework, Cisco Nexus 5000 Series Switches deliver a single network technology that supports all I/O in the system. The unified fabric simplifies the infrastructure by using converged network adapters (CNAs) that each present multiple Ethernet network interface cards (NICs) and Fibre Channel HBAs to the server's enhanced PCI (PCIe) bus. Use of CNAs makes the unified fabric's existence completely transparent to the operating system. Storage administrators can continue to use the same management tools and best practices while storage traffic is physically carried over the system's unified fabric.

A unified fabric that supports 10 Gigabit Ethernet and Fibre Channel over Ethernet (FCoE) offers many advantages:

- Fewer interface cards need to be purchased, installed, configured, powered, cooled, and managed.
- No Fibre Channel switches are needed to support the fabric at the rack level.
- The cost of cabling and transceivers is significantly reduced: fiber transceivers are costly, and fiber cabling is fragile. The 10 Gigabit direct-attached copper cabling supported by the Cisco Nexus 5000 Series is inexpensive and reliable.

In the reference configurations described in this document, the unified fabric extends directly to the EMC VNX5500 storage systems, eliminating the need for any Fibre Channel switching or cabling, further reducing costs. FCoE traffic is carried end to end over the Cisco Nexus 5000 Series low-latency 10 Gigabit Ethernet and FCoE network. To faithfully support the Fibre Channel model, all FCoE traffic is passed using a lossless priority on the Ethernet network so that no packets containing storage traffic can be lost due to network congestion. The use of FCoE does more than reduce cost: it helps increase performance as a result of the 10-Gbps speed of the underlying Ethernet network.

EMC VNX5500 Storage Systems

EMC VNX5500 storage systems are powerful, affordable, and simple, offering reliability, manageability, modular scalability, and performance to address sustained data access bandwidth for demanding data warehouse applications. Up to 250 drives are supported, including SAS, NL-SAS, and flash drives. Multiple communications protocols are supported, including Network File System (NFS), Common Internet File System (CIFS), Multiprotocol File System (MPFS), parallel NFS

(pNFS), Fibre Channel, Small Computer System Interface over IP (iSCSI), and FCoE. The system can deliver up to 50 percent more throughput by adding 4-port, 6-Gb SAS UltraFlex I/O modules. The total storage system, including all the enclosures for the disk drives, occupies nine rack units (9RU) of space.

The systems also provide leading array software features to support various types of disk drives, data placement, and autotiering. Other software capabilities include virtual logical unit number (LUN) migration and expansion, rapid point-in-time snapshots, stored data block cloning, and remote mirroring. Higher bandwidth results from the adoption of high-power multicore processors, fast memory chips, multichannel 6-GBps SAS back-end buses, and new generation-2 PCI and PCIe buses.

Whether accessed as block- or file-based storage objects, all data can be managed uniformly in a single EMC Unisphere™ web-based window. As described in the previous section, EMC VNX5500 storage systems are FCoE enabled, so storage traffic is carried across the entire network using FCoE rather than native Fibre Channel, resulting in significant cost savings.

Reference Configuration Overview and Performance

As a part of the Cisco and EMC collaboration, medium-sized enterprise and large enterprise configurations were defined to meet a range of data warehouse requirements, as shown in Table 1. Both configurations connect through a Cisco Nexus 5548UP Switch using different numbers of CNAs as appropriate to the requirements of each reference configuration. Cisco Nexus switches also can be used to connect other IT infrastructure as required in addition to the reference configuration components. Additional details of the configurations are provided in the sections that follow.

Table 1. Medium-Sized Enterprise and Large Enterprise Reference Configurations

Reference Configuration	Server	Processor	Memory	Settings
Medium-sized enterprise	Cisco UCS C260 M2 Rack Server	2 Intel Xeon processors E7-2870	256 GB	EMC VNX5500 storage system with 95 SAS disk drives
Large enterprise	Cisco UCS C460 M2 High-Performance Rack Server	4 Intel Xeon processors E7-4870	256 GB	2 EMC VNX5500 storage systems, each with 181 SAS disk drives

The Microsoft SQL Server Fast Track 4.0 workload consisted of a set of both simple and moderately complex queries representative of workloads in modern data warehouse environments. As part of the certification process, these queries ran from 5 to 40 concurrent query streams and ran from 30 minutes to 1 hour

each. From the results of these tests, both logical scan rates (how quickly data can be read from the buffer cache) and physical scan rates (how quickly data can be read from the physical disk) were determined. Table 2 lists the performance of both the medium-sized enterprise and large enterprise data warehouse reference configurations.

Table 2. Performance Testing Results for Medium-Sized and Large Enterprise Reference Configurations

Tested Criteria	Medium-Sized Enterprise Reference Configuration	Large Enterprise Reference Configuration
Maximum user data capacity	33 TB	66 TB
Recommended user data capacity	20 TB	40 TB
Benchmark scan rate: Logical	3583 MBps	7235 MBps
Benchmark scan rate: Physical	2226 MBps	4480 MBps
Average Microsoft SQL Server data processing rate	2904 MBps	5857 MBps
Maximum Microsoft SQL Server data processing rate	4960 MBps	7810 MBps

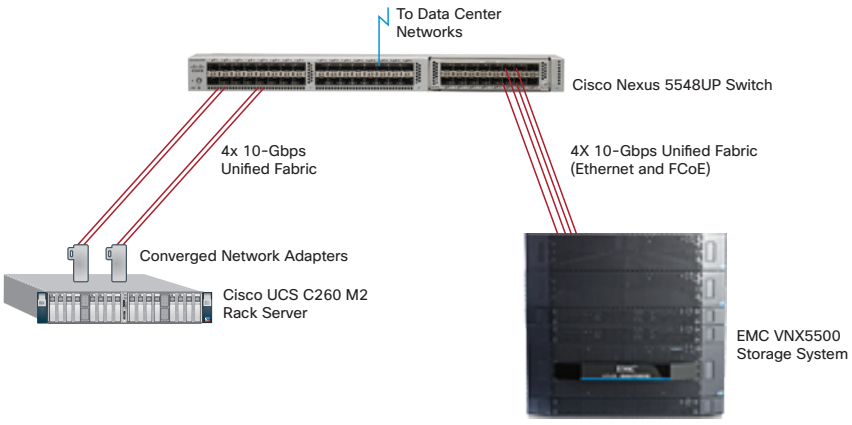

Figure 2. The Medium-Sized Enterprise Data Warehouse Configuration Uses an EMC VNX5500 Storage System Connected Through a Cisco Nexus 5548UP Switch

Medium-Sized Enterprise Data Warehouse Reference Configuration

The medium-sized enterprise configuration is intended to meet the performance, capacity, and price requirements of medium-sized to large data warehouse environments.

Configuration Details

Depicted in Figure 2 and detailed in Table 3, the medium-sized enterprise configuration offers considerable ease of use and scalability and provides a usable data capacity of 8 to 20 TB. Two QLogic QLE8152 dual-port 10-Gbps CNAs connect the Cisco UCS C260 rack server to the Cisco Nexus 5548UP with four 10-Gbps unified fabric links using FCoE.

Table 3. Medium-Sized Enterprise Configuration Details

Server	Cisco UCS C260 M2 Rack Server	Processor	lIntel Xeon processor E7-2870 (2.4 GHz, 30-MB cache, and 130 watts [W])
		Memory	256 GB (8 x 32-GB DDR3 DIMMs); up to 2048 GB supported
		Internal disk drives	2 Small Form-Factor (SFF) 600-GB 10K SAS disk drives
		CNA	4 QLogic QLE8152 dual-port 10-Gbps CNAs
Storage	EMC VNX5500	Disk drives	95 SFF 300-GB 10K SAS disk drives
		Data protection	RAID 5 for data and temporary database (tempdb) files; RAID 10 for log file
Connectivity	Switch	Cisco Nexus 5548UP Switch	
	Protocol	FCoE from server to switch, and Fibre Channel from switch to storage	

EMC VNX5500 Storage Layout

The EMC VNX5500 disk processing enclosure can support up to twenty-five 2.5-inch form-factor disk drives. Additional disk drives can be connected using disk-array enclosures (DAEs) that each provide 25 additional drive slots (up to 250 disk

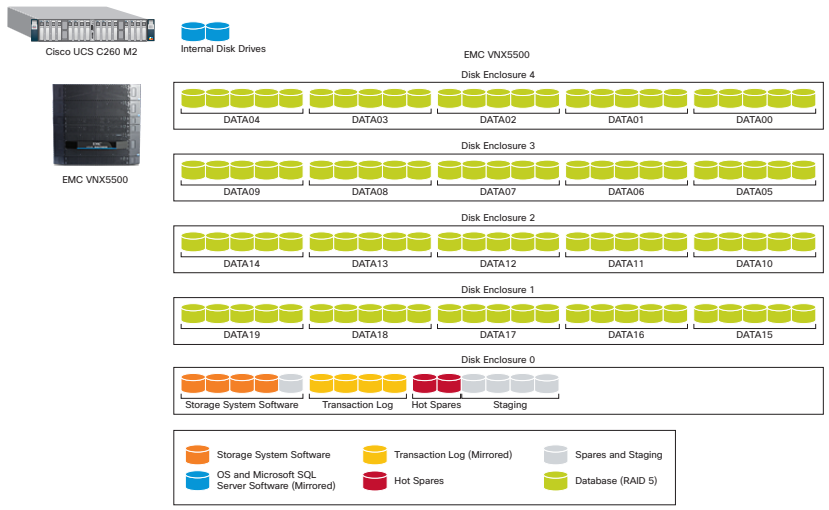

Figure 3. The EMC VNX5500 Storage System Is Configured with Four DAEs

drives per enclosure). One EMC VNX5500 storage system with four installed DAEs is used in this reference configuration. The total storage system, including all the enclosures for the disk drives, occupies 9RU of rack space. Figure 3 illustrates the data layout for the reference configuration.

Disk drives in the EMC VNX5500 storage system are configured according to their use in the data warehouse system:

- **Storage system:** Four drives are reserved for the EMC VNX storage system's operating environment, shown as dark orange in the illustration.
- **Transaction log:** Four drives in the form of a 2+2 RAID 10 group are used for the transaction log, shown as dark yellow in the illustration.
- **Data LUNs:** 80 drives are used for database files. These 80 drives are organized into 4+1 RAID 5 sets, known as RAID groups in EMC VNX terminology, and are colored green in the illustration.
- **Hot spares:** Two drives are reserved as hot spares (colored red in the diagram) and are ready to be automatically switched in to immediately rebuild any data drive in any RAID 5 group, reducing the chances of data loss from a drive failure in one of the data LUNs.
- **Staging area:** Four drives are used as a staging area, such as for data landing; these drives are colored gray in the illustration.

Design Considerations

The RAID 5 or parity RAID configuration consists of four data drives protected by one parity drive. In the RAID 5 configuration, each 256-KB chunk of database data stored on one of the data LUNs is striped in 64-KB increments across four of the five drives in each RAID set. On the fifth drive, 64 KB of parity information is computed and stored based on the data on the other four drives.

RAID 5 was chosen as the storage deployment layout for two main reasons:

- Lower cost of data protection compared to RAID 10, with RAID 5 requiring only 20 percent of the total usable capacity for storing parity data
- High read performance because all drives contribute to the delivery of data

EMC VNX5500 Storage System Cache Settings

In this reference configuration, 2000 MB of DRAM from the storage processors supports host write caching, and 500 MB supports read-ahead caching. Figure 4 provides a screenshot from the storage management browser window.

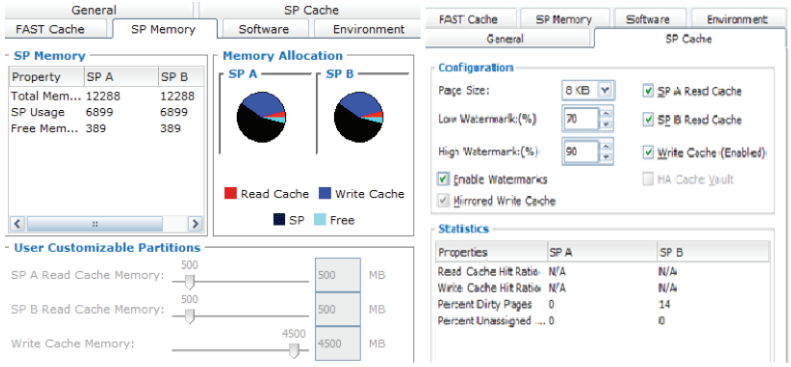

Figure 4. The EMC VNX5500 Storage System Can Be Configured to Use Its DRAM to Cache Both Read and Write Operations

Large Enterprise Data Warehouse Reference Configuration

The large enterprise data warehouse reference configuration is intended to meet the performance, capacity, and price requirements of enterprise-class data warehouse environments, while offering the ease of use of the EMC VNX5500 storage system.

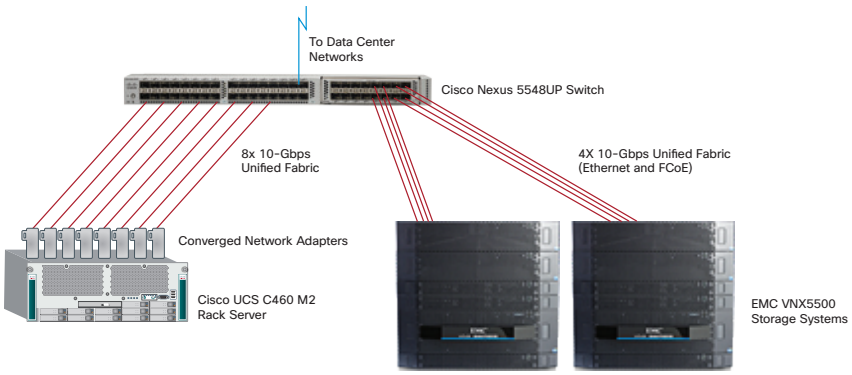

Figure 5. The Large Enterprise Data Warehouse Configuration Uses Two EMC VNX5500 Storage Systems for End-to-End FCoE Connectivity

Configuration Details

The large enterprise data warehouse configuration is depicted in Figure 5 and detailed in Table 4. The configuration consists of a Cisco UCS C460 M2 Rack Server equipped with four Intel Xeon processors E7-4870, 256 GB of memory, and eight QLogic QLE8152 dual-port 10-Gbps CNAs. The storage consists of two EMC VNX5500 storage systems connected through a Cisco Nexus 5548UP Switch.

Table 4. Large Data Warehouse Reference Configuration Details

Server	Cisco UCS C460 M2 Rack Server	Processor	4 Intel Xeon processors E7-4870 (2.40 GHz, 30-MB cache, and 130W)
		Memory	256 GB (32 x 64-GB DDR3 1333-MHz RDIMM); up to 2048 GB supported
		Internal disk drives	2 SFF 300-GB 10K SAS disk drives
		CNA	8 QLogic QLE8152 dual-port 10-Gbps CNAs
Storage	Two EMC VNX5500 storage systems	Disk drives	181 SFF 300-GB 10K SAS disk drives
		Data protection	RAID 5 for data and tempdb files; RAID 10 for log files
Connectivity	Switch	Cisco Nexus 5548UP	
	Protocol	FCoE from server to switch, and Fibre Channel from switch to storage	

Storage Configuration and Layout

Each of the EMC VNX5500 storage systems connected to the Cisco UCS C460 M2 high-performance rack server is configured with 175 drives. The design principles for the RAID groups and LUN layout are similar to those for the medium-sized enterprise reference configuration. This consistent approach allows a deployment that may have been initially sized for a midrange requirement to be expanded and upgraded more readily into a larger enterprise configuration. Figure 6 shows the configuration.

Figure 5. Four Disk-Array Enclosures Are Installed in Each of Two EMC VNX5500 Storage Systems in the Large Enterprise Reference Configuration

Bill of Materials

Tables 5 and 6 list the bills of materials for the medium-sized enterprise and large enterprise data warehouse reference configurations.

Table 5. Bill of Materials for Medium-Sized Enterprise Data Warehouse Reference Configuration

Server	Part Number	Quantity
UCS C260 M2 Rack Server (w/o CPU, MRB, PSU)	C260-BASE-2646	1
850 W Power Supply Unit for C-Series 250 M2	R250-PSU2-850W	1
Server	**Part Number**	**Quantity**
2.4 GHz Intel Xeon E7-2870 130W 10C/30M cache	UCS-CPU-E72870	2
2X8GB DDR3-1333-MHz RDIMM/PC3-10600/dual rank/ x2/1.35v	UCS-MR-2X082RX-C	8
LSI 6 G MegaRAID 9261-8i card	R2XX-PL003	1
300 GB 6 Gb SAS 10K rpm SFF HDD	A03-D300GA2	2
QLogic QLE8152 Dual Port 10 Gb CNA	N2XX-AQPCI01	4
Switch		
Nexus 5548 UP Chassis, 32 10 GbE Ports, 2 PS, 2 Fans	N5K-C5548UP-FA	1
10GBASE-CU SFP+ Cable 3 Meter	SFP-H10GB-CU3M	16
Storage	**Model Number**	**Quantity**
VNX5500 (including 4 FLARE drives)	VNX55D253010F	1
2-Port 10 GbE FCoE Module	VSPMXGFCOOPAS	2
25x2.5" Drive Disk-Array Enclosure	V2-DAE-R-25-A	4
300 GB 2.5" 10K drives	VX-2S10-300	95
Unisphere Block and VNX OE VNX5500	UNIB-V55	1
2nd Optional SPS for VNX5500	VNXSPSAS	1

Table 6. Bill of Materials for the Large Enterprise Data Warehouse Reference Configuration

Server	Part Number	Quantity
UCS C460 M2 Rack Server	UCS-BASE-M2-460	1
850W Power Supply Unit for C-Series C460	RC460-PSU2-850W	4
2.4GHz Xeon E7-4870 130W CPU/30MB cache	UCS-CPU-E74870	4
Memory Riser Board for C460 M2 Rack Server	UCSC-MRB-002-C460	8
2X8GB DDR3-1333-MHz RDIMM/PC3-10600/dual rank/x2/1.35v	UCS-MR-2X082RX-C	16
LSI Controller 9260-8i	RC460-PL001	1
600GB 6Gb SAS 10K RPM SFF HDD	A03-D600GA2	2
QLogic QLE8152 Dual Port 10 Gb CNA	N2XX-AQPCI01	8
Switch		
Nexus 5548 UP Chassis, 32 10GbE Ports, 2 PS, 2 Fans	N5K-C5548UP-FA	1
10GBASE-CU SFP+ Cable 3 Meter	SFP-H10GB-CU3M	16
Storage	**Model Number**	**Quantity**
VNX5500 (including 4 FLARE drives)	VNX55D253010F	2
25x2.5" Drive Disk-Array Enclosure	V2-DAE-R-25-A	4
300 GB 2.5" 10K drive	VX-2S10-300	181
Unisphere Block and VNX OE VNX5500	UNIB-V55	2
2nd Optional SPS for VNX5500	VNXSPSAS	2

Conclusion

The Microsoft SQL Server Data Warehouse 4.0 reference configurations presented in this document enable rapid deployment of Microsoft SQL Server 2012 enterprise data warehouse implementations. Designed and tested by Cisco, EMC, and Microsoft, these configurations offer proven performance, reliability, and lower total cost of ownership (TCO) for organizations deploying infrastructure for data warehouse applications. Perhaps most important, these configurations

help organizations quickly achieve known, validated, and predictable results rather than having to undertake the challenge of designing their own data warehouse implementations.

The combination of intelligent Intel Xeon processors, Cisco UCS C-Series Rack Servers, Cisco Nexus 5548UP unified fabric switches, and innovative EMC VNX5500 storage systems results in configurations that can be scaled easily to accommodate a range of data warehouse workloads. The reduction in the number of switches, the reduced physical size of the server, and the much smaller storage footprint of the 2.5-inch SFF disk drives all contribute to the physical compactness of the complete system. With fewer components, higher effective storage density, and lower energy costs, these reference configurations address the growing concerns of many IT professionals who support growing data warehouses.

For More Information

To learn more about the Microsoft SQL Server Fast Track Data Warehouse solution, visit http://www.cisco.com/go/microsoft.

Acknowledgments

This document was made possible through the efforts of Raghunath Nambiar and the EMC Partner Engineering team with thanks to Mike Bassett, Girish Kulkarni, and Rex Backman.

Americas Headquarters
Cisco Systems, Inc.
San Jose, CA

Asia Pacific Headquarters
Cisco Systems (USA) Pte. Ltd.
Singapore

Europe Headquarters
Cisco Systems International BV Amsterdam,
The Netherlands

Cisco has more than 200 offices worldwide. Addresses, phone numbers, and fax numbers are listed on the Cisco Website at **www.cisco.com/go/offices**.

LE-36905-00 11/12