EMC Solutions for Tiered Storage for Microsoft SQL Server 2008

Enabled by EMC CLARiiON CX4 Series iSCSI and Windows 2008

Reference Architecture

EMC Global Solutions

Published January, 2010

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

Benchmark results are highly dependent upon workload, specific application requirements, and system design and implementation. Relative system performance will vary as a result of these and other factors. Therefore, this workload should not be used as a substitute for a specific customer application benchmark when critical capacity planning and/or product evaluation decisions are contemplated.

All performance data contained in this report was obtained in a rigorously controlled environment. Results obtained in other operating environments may vary significantly.

EMC Corporation does not warrant or represent that a user can or will achieve similar performance expressed in transactions per minute.

No warranty of system performance or price/performance is expressed or implied in this document. Use, copying, and distribution of any EMC software described in this publication requires an applicable software license.

For the most up-to-date listing of EMC product names, see EMC Corporation Trademarks on EMC.com.

All other trademarks used herein are the property of their respective owners.

Part number: H6016.1

Contents

About this Solution

About this Solution

Purpose

The purpose of this document is to illustrate a Microsoft SQL Server 2008 solution on a EMC® CLARiiON® CX4-240 using iSCSI to leverage IP SAN connectivity.

The business challenge

With increasing demands and limited resources, today's midsize enterprises face challenges similar to their larger counterparts. Apart from managing expanding IT costs and reducing the risk of business disruption, some of the challenges midsize enterprises face include:

- Consolidating multiple database applications scattered throughout the enterprise
- Ensuring information access, availability, and continuity
- Maximizing server and storage utilization to deliver optimal system performance
- Managing upgrades and migrations
- Reducing the demands on limited IT resources and budgets
- Decreasing the complexity of technology choices

The technology solution

Centralization of information and consolidation of resources create immense economic advantages for any business. This solution offers to:

Maintain service levels — Keep critical and revenue-generating Microsoft applications available and running at peak performance.

Reduce costs — Minimize the cost of database and information management.

Reduce risk — Provide a proven reference architecture with tested and validated configurations.

Accelerate implementations — Let EMC Professional Services and ASN-certified EMC Partners provide rapid assessment and efficient implementation.

Solution details

The following section provides high-level information on the key components used in this solution. Refer to "Environment profile" on page 6 for more detailed information.

EMC CLARiiON CX4-240 — The EMC CLARiiON CX4-240 delivers best-in-class performance for midrange networked storage that scales seamlessly up to 231 TB of capacity. The CX4 model 240 supports up to 256 highly available, dual-connected hosts, and scales from five to 240 disks. All CX4 models come pre-configured with Fibre Channel (FC) and iSCSI connectivity, allowing customers to choose the best connectivity for their specific applications.

Microsoft SQL Server 2008 — SQL Server 2008 delivers on Microsoft's Data Platform vision by helping organizations to manage any data, any place, any time. Store data from structured, semi-structured, and unstructured documents, such as images and rich media, directly within the database. SQL Server 2008 delivers a rich set of integrated services that enables organizations to do more with their data, such as query, search, synchronize, report, and analyze.[1]

1. Microsoft website

Environment profile

Figure 3 on page 13 shows the overall architecture of the solution, which consists of a consolidation module that comprises the following components:

- Client (or public) LAN and iSCSI SAN networks
- Microsoft SQL Server 2008 connected to the client LAN and iSCSI SAN networks
- Clients connected to Microsoft SQL Server 2008 through the client LAN

The solution architecture has the following general characteristics:

- The Microsoft SQL Server 2008 resides on a two-node active/passive cluster to provide local high availability.
- The Microsoft SQL Server 2008 Enterprise Edition (64-bit) runs on Windows 2008 Enterprise Edition 64-bit.
- EMC PowerPath® (multipath software) runs on each server in the Microsoft SQL Server 2008 cluster to provide HBA load-balancing and fail-over capabilities.
- The database datafiles, logfiles, TempDB files, and system database files reside on separate RAID groups (physical RAID arrays). The LUNs for the database files and database logfiles are designed specifically to sustain TPC-C (simulated OLTP) type I/O loads.
- IP (Gigabit Ethernet and VLAN) configuration is designed to balance and distribute the network traffic.
- All database files are stored on the CLARiiON CX4-240 storage system, thereby decreasing server replacement complexity.
- The management and client networks shown in Figure 3 include IP switches. These switches support Gigabit Ethernet (GbE) connections.

Storage architecture

There are two storage processors (SPs) per CX4-240, each with a base host connectivity of:

- Two 4 Gb/s FC optical ports: FCP SCSI-3 protocol, FC-AL and FC-SW support
- Two 1 Gb/s Ethernet copper iSCSI ports

The system that was tested for this solution had an extra two 1 Gb/s iSCSI ports per SP. Additionally, 1 Gb/s iSCSI and 4 Gb/s FCP ports can be added, and in future releases, 10 Gb/s iSCSI and 8 Gb/s FCP ports can be added.

Figure 1 illustrates the EMC CLARiiON CX4-240 storage processor ports.

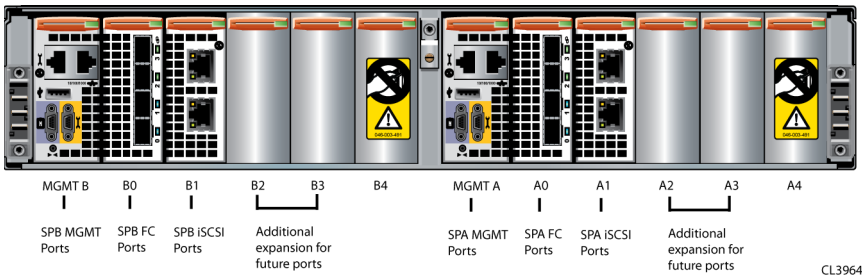

Figure 1 EMC CLARiiON CX4-240 storage processor ports

Disk configuration

The CLARiiON storage system consists of a series of disk-array enclosures (DAEs). Each DAE can hold up to 15 disks. The CX4-240 can scale up to 16 DAEs with a total of 240 disks. The disks can be logically combined across DAEs to form RAID groups. The RAID groups can be segmented into logical unit numbers, or LUNs.

The LUNs can be exposed to the Microsoft Windows operating system. When additional storage is required, disks can be added to DAEs that are not yet full, or new DAEs can be installed and combined into additional RAID groups. The LUNs can be expanded to form metaLUNs. A metaLUN retains the logical unit number of the original LUN. The metaLUNs can also be exposed to the Microsoft Windows operating system.

Table 1 details the disk configuration that was used as a part of this solution.

Table 1 Disk configuration

Configuration	Number of disks				
	System	TempDB	Data	Log	Backup
1	2	2	12	4	9

There was a single disk configuration tested that had 12 disks for the TPC-C database data and four database log disks.

Note: EMC CLARiiON FLARE® two-disk RAID 1/0 is similar to RAID 1 (FLARE is the CLARiiON operating environment). Both employ one-to-one disk mirroring. The significant difference between the two is that RAID 1/0 supports RAID group expansion (assigning additional disks to a RAID group) whereas RAID 1 does not.

Figure 2 illustrates the distribution of the application files across the disks.

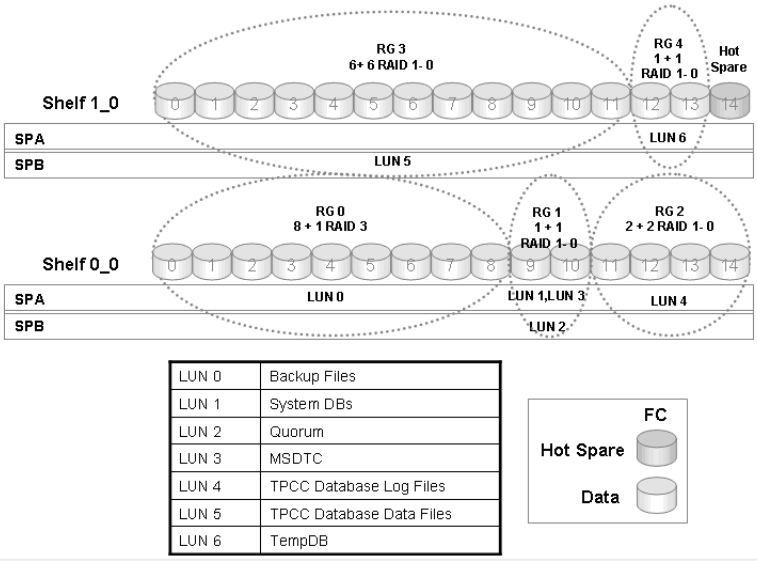

Figure 2 Distribution of application files

In the configuration, the TPC-C database datafiles were stored in two datafiles, each in its own LUN, and both LUNs were taken from the RAID 1/0 group of 12 disks. Each LUN of 200 GB (total of 400 GB) contained a single database datafile of 175 GB (350 GB total). Therefore, the total size of the database datafiles was 350 GB, of which 240 GB was used at the beginning of the test iteration set.

The TPC-C database logfile was contained on a single LUN, which was presented from the RAID 1/0 group of four disks. The single database log LUN was 100 GB, and contained a 50 GB database log file.

The two disks allocated to the system area contained a LUN that had all of the system database files and logfiles such as msdb, master, model, and so on. If the change rate of any of the system databases is high, or if the more recent data is critical, then it is recommended that the system database logfiles be stored on a

LUN that is on a different RAID group, thereby physically isolating the logs from the database to maximize recoverability. In the solution that was tested, as the system databases sustained few or no changes at regular intervals, this was not an issue.

The quorum LUN for a cluster can also be located on the same disks as the system databases. In most scenarios, this should not be an issue. However, if the RAID group were to become busy, there is a possibility that the cluster might see the quorum LUN as unavailable and cause a needless cluster failover. Therefore, it is important that the quorum LUN be located on a RAID group that will not become excessively busy. For the current solution, as the system databases are never put under much stress, it is acceptable to have the quorum LUN in the same RAID group as the system databases.

Tempdb was isolated to its own two-disk RAID 1/0 group, as the isolation of tempdb is a recommended best practice from Microsoft. However, these disks were seldom used as the TPC-C workload puts effectively zero workload on the tempdb database.

Nine disks were set up in a RAID 3 array to be used for backup and restore purposes. RAID 3 was used because it provides an excellent throughput for serial I/O and the backups are exclusively serial.

Table 2 lists each file/activity type and indicates where it resides or is recorded.

Table 2 Application file or activity types and locations

File or activity type	Location	RAID level	Total number of physical disks
Database binary files	Host local disk	RAID 1	N/A
TPC-C database logfiles	CX4-240	RAID 1/0	4
TempDB files	CX4-240	RAID 1/0	2
TPC-C database datafiles	CX4-240	RAID 1/0	12
Cluster quorum disk	CX4-240	RAID 1/0	2
MSDTC[a]	CX4-240	RAID 1/0	2
System database files	CX4-240	RAID 1/0	2

a. Microsoft Distributed Transaction Coordinator. This software enables a single transaction to include data from multiple source locations. The software then coordinates the process of committing the transaction across the various locations.

Disk alignment

Unlike Windows 2003, Windows 2008 Disk Manager automatically creates aligned partitions, so there is no need to use diskpart or other realignment methods. For further information on disk alignment refer to the following documents:

- *White Paper: Using diskpar and diskpart to Align Partitions on Windows Basic and Dynamic Disks*
- *Aligning GPT Basic and Dynamic Disks for Microsoft Windows 2003 Technical Note*

Note: These documents are available on EMC Powerlink, if you do not have access please contact your EMC representative.

Microsoft SQL Server 2008 architecture

Each of the two servers in the Microsoft SQL Server 2008 cluster has six network interfaces. Four of the network interfaces connect to the SAN, one network interface connects to the other server in the cluster (to enable the heartbeat through a crossover), and one interface connects to the client (public) network. Table 3 lists each interface and describes its use.

Table 3 Application server network interface configuration

Interface port ID	Description
Eth0	Public network
Eth1	Open for future use
Eth2	SAN connection
Eth3	SAN connection
Eth4	SAN connection
Eth5	SAN connection

High availability and failover

The validated solution provides protection at the storage and host layers.

Storage layer

The CLARiiON CX4-240 has two SPs to provide high availability and load balancing. In the validated solution, the two SPs provide seamless failover

capabilities for the storage. This minimizes disruption during maintenance or possible component failure. The RAID disk configuration on the CLARiiON back end provides protection against hard disk failures.

Host layer

The application hosts have redundant power supplies, iSCSI connections, and network connections to reduce the impact of host hardware failure.

EMC PowerPath (multipath software) was used to load balance I/O across the iSCSI paths and provide path failover in the event of the loss of a path.

Microsoft Failover Clustering

Microsoft Failover Clustering, formerly known as Microsoft Cluster Services (MSCS), is enabled on each of the two clustered application servers. The servers operate in active/passive mode to provide local protection against the potential hardware failure of the active server.

MSCS shared storage, which includes quorum disk files, database files, and all system database files, resides on CLARiiON storage.

Application architecture

No special configuration settings were used for Microsoft SQL Server 2008 during the solution testing. All default settings were used for memory, processors, network, threads, and so on. The only exception was that it was verified that the account under which the SQL Server service ran had the Lock Pages In Memory and the Perform Volume Maintenance Tasks user rights to allow the SQL Server to prevent pages from being swapped, and to enable fast file initialization.

Hardware resources

The hardware resources used in the solution are listed in Table 4.

Table 4 Hardware

Equipment	Quantity	Configuration	Notes
EMC CLARiiON CX4-240	1	• 2 x SPs • 4 x 1 Gb/s iSCSI connections per SP • 3 x FC shelves (DAE) • 15 x 300 GB (15k) FC disks per DAE	FLARE 3.28
4U server	2	• 4 x 2.8 GHz AMD Opteron dual-core processors • 32 GB of memory • 2 x 146 GB 15k internal SCSI disks • 2 x On-board 10/100/1000 MB Ethernet NICs • 4 x 10/100/1000 MB Ethernet NICs	One active/passive cluster
1U server	3	• 2 x quad-core Intel 5300 series Xeon • 16 GB of memory	VMware ESX Server 3.5 with Benchmark Factory Client VMs
Switch	2	• 48 (copper) 1 Gb/s ports	

Software resources

The software resources used in the solution are listed in Table 5.

Table 5 Software

Software	Number of licenses	Configuration
Microsoft Windows 2008 Server Enterprise Edition, 64-bit	2	SQL Application Server One license per node in the cluster Using EMC PowerPath 5.0.0.94
Microsoft SQL Server 2008 Enterprise, 64-bit	1	One license per cluster

Physical architecture

Figure 3 illustrates the overall physical architecture of the solution.

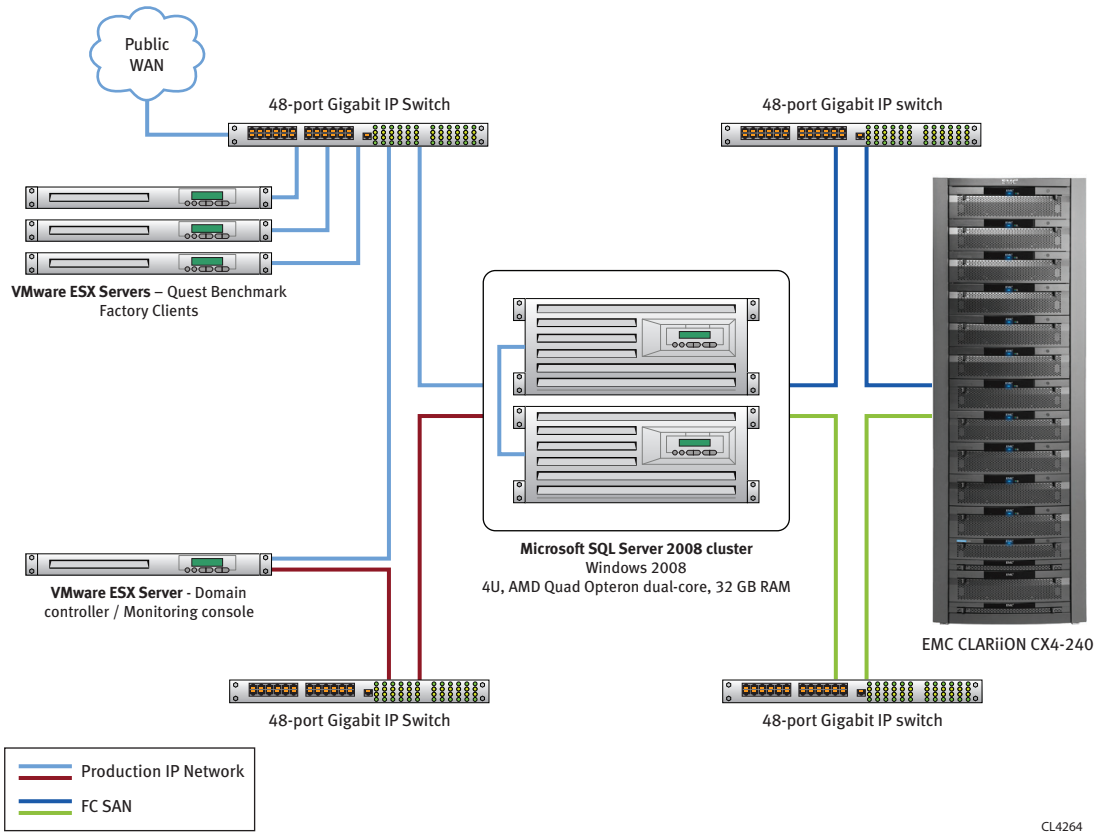

Figure 3 SQL Server 2008 with EMC CLARiiON CX4 series iSCSI enabled by Microsoft Windows 2008

Conclusion

This document depicts a solution that utilizes a CX4-240 that is hosting the storage for SQL Server 2008 on Windows 2008. The Windows 2008 server accesses the CX4-240 LUNs via iSCSI.

This solution was tested with an OLTP-type workload (TPC-C), which consists primarily of smaller random I/Os. Therefore, the stress placed on the physical iSCSI connection is more about IO/s, or IOPS, than MB/s. iSCSI is currently limited to 1 Gb/s (soon to be 10 Gb/s), which translates to around 125 MB/s, although an estimated conservative throughput of 100 MB/s is almost always used for planning purposes.

The configuration as tested had four iSCSI connections from the server to the CX4-240. Therefore, an aggregate throughput of approximately 400 MB/s would be achievable. This is approximately equal to a single 4 Gb/s Fibre Channel connection. If a specific implementation needs more than 400 MB/s, then the following two options are available.

iSCSI — More iSCSI connections could be added to increase the aggregate throughput, but it is important to understand that more than one server will likely be using the CX4-240 simultaneously, so all of the available bandwidth should not be planned to be available to a single server. Therefore, it is important to plan out the I/O needs of all servers that will be connected to the CX4-240 to be sure that iSCSI is a good fit for all servers.

Fibre Channel — Another option is to use FCP either for all of the servers, or only for the servers that require the greatest bandwidth. Since the CX4-240 allows for multiple FCP and iSCSI ports to be available and active simultaneously, a hybrid approach is usually the best answer when only a few (or one) servers need high MB/s throughput.

All implementations require proper analysis and planning to achieve maximum benefit and ROI from the implementation.

To learn more about this and other solutions contact an EMC representative or visit *www.EMC.com/solutions/microsoft*.