

EMC[®] Celerra[®] Solutions for Microsoft SQL Server 2005 Always On

Technical Note

P/N 300-006-726

REV A01

April 28, 2008

This technical note contains information on these topics:

◆ Overview	2
◆ Disclaimer	2
◆ Features	2
◆ EMC compliance	2
◆ Core 1.00: Windows Logo certification	3
◆ Core 1.01: Windows API support	3
◆ Core 1.02: Stable media	3
◆ Core 1.03: Forced Unit Access and Forced Write-through.....	4
◆ Core 1.04: Asynchronous capabilities	5
◆ Core 1.05: Write ordering.....	5
◆ Core 1.06: Torn page protection.....	5
◆ Core 1.07: NTFS support.....	6
◆ Recommended 1.08: Partner enhanced escalation.....	6
◆ Appendix A: Microsoft Always On specification version 1.02.....	7
◆ Appendix B: References	12

Overview

This tech note provides information on EMC® Celerra® solutions based on the Microsoft SQL Server Always On Storage Solution Review program.

Information regarding the Microsoft SQL Server 2005 Always On program is available at <http://www.microsoft.com/sql/alwayson>.

Disclaimer

This document is produced independently of Microsoft Corporation. Microsoft Corporation expressly disclaims responsibility for and makes no warranty—express or implied—with respect to the accuracy of the contents of this document.

The information contained in this document represents the current view of EMC on the issues discussed as of the publication date. Because of changing market conditions, it should not be interpreted as a commitment on the part of EMC. Also, EMC cannot guarantee the accuracy of any information presented after the publication date.

Features

The Microsoft SQL Server Always On program provides formal compliance documentation for solutions provided by vendors. EMC provides a number of highly available solutions, which are covered by this framework. This document provides coverage of those storage-based solutions, which are compliant under the definition of SQL Server Always On.

EMC compliance

The subsequent sections document EMC compliance for the requirements as provided by SQL Server Always On documentation. The specifications that apply to this Always On submission are listed in [Appendix A: Microsoft Always On specification version 1.02](#).

Under the program, EMC has provided compliance with the Celerra NS Series products.

NOTE: Celerra NS Series products use EMC CLARiiON® or Symmetrix® as the backend storage. The iSCSI protocol is used to provision the Celerra storage to the SQL Server to store the database and log files. Unless otherwise stated explicitly, iSCSI is the default Celerra protocol for the SQL Server.

Core 1.00: Windows Logo certification

All EMC Celerra NS series products are logo certified for iSCSI protocols. Windows logo certification requirements and platform listings are available in the Windows Catalog under the storage array category.

The following URL provides a link to the WindowsServer Catalog. Search **EMC** under the **Storage** category.

<http://www.microsoft.com/windows/catalog/server/>

EMC additionally provides solutions for clustering with Celerra. Microsoft also provides a listing within the Windows Catalog for logo certification under **Hardware > Cluster Solutions**.

Core 1.01: Windows API support

All EMC Celerra products covered under this compliance statement fully support the core Windows API. EMC Celerra supports all native Windows file system calls when accessed via iSCSI.

Write operations to supported EMC Celerra storage platforms guarantee delivery to stable media as defined in subsequent sections of this document. Cache write operations to Celerra storage are protected by battery backup systems and other cache protection mechanisms, such as cache write mirroring and cache destaging to "cache vault."

Core 1.02: Stable media

EMC Celerra guarantees stable storage.

All EMC Celerra products covered under this compliance statement fully adhere to SQL Server Write Ahead Logging (WAL) protocols and meet ACID (Atomicity, Consistency, Isolation, and Durability) requirements as defined in the SQL Server 2000 I/O Basics documentation.

EMC Celerra also provides highly available RAID-protected storage and n+1 power and cooling redundancy, and is designed for data-in-place upgrades.

Fail-Safe Networking (FSN)

FSN groups together two network connections on the Celerra system as an active/standby pair for high availability. This configuration provides protection from port and switch failures. FSN is important for iSCSI since the TCP connection is transferred to the surviving interface with no service interruption for the iSCSI host initiator.

The Celerra contains two IP storage blades. These blades can operate independently, or they can operate in active/passive mode, with the passive blade serving as a failover device for the active blade. The Celerra can have multiple blades to provide high availability and load balancing. In the solution for SQL Server, primary and standby blades provide seamless failover capabilities for the Celerra storage. This minimizes end-user disruption during routine Celerra maintenance such as upgrading the DART OS.

EMC Celerra supports Windows MC/S (Multiple connections per Session) provided by the Microsoft iSCSI Initiator, which allows multiple TCP/IP connections between Microsoft SQL Server and EMC Celerra storage during the same iSCSI session, either on the same or a different physical link.

EMC Celerra and replication products ensure that log predecessor writes are honored. These solutions utilize EMC consistency technology.

Core 1.03: Forced Unit Access and Forced Write-through

All EMC Celerra products covered under this compliance statement adhere to Forced Unit Access and Forced Write-through requirements.

The Celerra is a dedicated network server that is optimized for file access to data and delivers high-end features in a midtier package. The EMC Celerra NS Series leverages highly available and RAID-protected back-end storage and industry-leading availability, performance, and ease of management.

EMC storage arrays are integrated cache disk arrays (ICDA). These systems provide onboard caching mechanisms to optimize I/O operations for connected servers and associated applications. Write operations specifically benefit from the speed of write operations to cache. Cache I/O operations are typically orders of magnitude faster than write operations to the physical disk media. All EMC Celerra storage utilizes a protection mechanism to ensure the durability and persistence of updated (write) data stored within the cache. Specifically, for Celerra products included within the Always On program, a number of mechanisms are provided.

Battery backup

EMC Celerra storage includes battery backup components. These battery backup components are tested and certified to support the required operations in the event of a failure in the primary power supply.

Generally for EMC storage arrays under primary power failure, cache memory is written to persistent durable media in a designated cache vault located on specific disks within the array. When primary power is restored, the cache vault is reloaded into memory, and the pending updates are submitted to the relevant logical units. In no case are partial I/O operations propagated to the logical unit.

Cache vault areas are themselves implemented in a RAID configuration. Thus, the vault area is protected against disk failures.

For certain EMC Storage arrays, battery backup power maintains power to all disk resources and ensures that all I/O operations complete to disk.

Mirrored write cache

EMC Celerra storage implements write cache mirroring protection. Under this scheme, updated cache areas are implemented in a RAID 1 convention. As a result of this implementation, updates are fully redundant and are protected against a single point of failure, such as a memory board fault.

Core 1.04: Asynchronous capabilities

All EMC Celerra products under this compliance statement adhere to this requirement. EMC Celerra will not transition asynchronous I/O operations from a host into synchronous operations.

Core 1.05: Write ordering

All EMC Celerra under this compliance statement adhere to and can enforce write ordering. All writes to the EMC Celerra are written in the order in which they are received from the SQL Server instance.

Core 1.06: Torn page protection

All EMC Celerra products under this compliance statement support this core requirement.

Core 1.07: NTFS support

All EMC Celerra configurations covered under this compliance statement provide full support for all NTFS capabilities.

Recommended 1.08: Partner enhanced escalation

EMC is a Microsoft Gold Certified Support partner and both companies have signed a Cooperative Support Agreement (CSA). The agreement formally commits EMC and Microsoft to work together for mutual customers with valid configurations and support contracts. Additionally, the agreement defines call escalation and transfer processes that make support resources available to each company 24x7x365 on a worldwide basis.

Appendix A: Microsoft Always On specification version 1.02

Requirements

This section contains the core requirements for the Always On Storage Solution Review program as provided by Microsoft when this compliance document is submitted. Storage system capabilities are divided into two categories: Required and Recommended.

Type	Definition
Required	A capability or property the subsystem <i>must</i> provide to pass the requirements of the Always On Storage Solution Review Program.
Recommended	A capability or property the subsystem <i>should</i> provide for optimal compatibility with SQL Server.

An Always On solution white paper must document the product, feature or features, and specific product configurations that support each of these core requirements before you submit the paper for review to the Always On Storage Solution Review program.

List of core requirements

Core 1.00	Windows Logo certification	Required
Microsoft Windows logo certification helps ensure the safety of Microsoft SQL Server data by testing various aspects of the products. To be compliant with the Always On Storage Solution Review program, solutions must pass the Certified for Windows logo testing.		

Core 1.01	Core Windows API support	Required
Microsoft SQL Server uses several APIs to accomplish secure data storage. A storage solution must ensure that a system supports specific API properties throughout the various layers and implementations of the I/O solution.		

Core 1.02 Stable media	Required
<p>Microsoft SQL Server relies on the Write-Ahead Logging (WAL) protocol to maintain the Atomicity, Consistency, Isolation, and Durability (ACID) properties of the database and to guarantee data integrity. WAL relies on stable media capabilities. A solution must comply with this stable media guarantee.</p> <p>For detailed information, consult the "Power Outage Testing – Pull The Plug" section in Chapter 2 of the <i>Microsoft SQL Server I/O Basics</i>.</p>	
Core 1.03 Forced Unit Access (FUA) and Write-Through	Required
<p>To support Write-Ahead Logging (WAL), Microsoft SQL Server uses FILE_FLAG_WRITETHROUGH and FlushFileBuffers to open files. Both of these options must be supported by storage solutions.</p> <p>All components in a solution must honor the write-to-stable media intent. This includes, but is not limited to, caching components.</p> <p>It is not enough to only honor WAL for Microsoft SQL Server log files. Data files and backup streams also depend on WAL behavior.</p> <p>Some storage products include battery-backed caching mechanisms as part of the write-through guarantee. If these caching mechanisms are present in the solution, the Always On solution white paper should document the practical limits of the write-through guarantee for a production environment.</p> <p>For more information, consult the links listed in the Appendix B: References section at the end of this paper, and the following Microsoft Knowledge Base article: KB917043 - Key factors to consider when you evaluate third-party file cache systems with Microsoft SQL Server.</p>	
Core 1.04 Asynchronous capabilities	Required
<p>Microsoft SQL Server performs most of its I/O using asynchronous capabilities. If a request specifies asynchronous operation, no API call should cause a synchronous condition. Synchronous I/O can cause unexpected scheduler and concurrency issues. Therefore, an Always On solution must provide asynchronous I/O capabilities.</p> <p>For more information, consult the white paper, <i>How To Diagnose and Correct Errors 17883, 17884, 17887, and 17888</i> at: http://www.microsoft.com/technet/prodtechnol/sql/2005/diagand</p>	

[correcterrs.mspix](#)

Core 1.05 Write ordering	Required
<p>A tenet of the WAL protocol is write ordering or order preservation. An Always On solution must provide write ordering capabilities. The write-ordering requirement applies to both local and remote I/O destinations. If a database is split among physical paths, all paths must honor the ordering across all files related to the database. To satisfy this ordering requirement, products sometimes use a user-defined consistency group that encapsulates all the database files. An Always On solution white paper must include information about the configuration requirements needed for the solution to meet the write ordering requirement. For example, a solution requiring a consistency group might specify this configuration requirement as: "All files associated with a database must be configured in a single consistency group."</p> <p>Example Configuration</p> <p>A solution has the following configuration:</p> <ul style="list-style-type: none"> • Data File A LUN —Subsystem #1 • Log File B LUN—Subsystem #2 <p>If these subsystems use separate physical paths with different caching, Microsoft SQL Server is unable to support this configuration because the caching mechanisms may not present a coherent cache. The subsystems would require a third element to maintain cache coherency across the disparate caches.</p> <p>The same caching problem described in the example configuration can also occur across network boundaries. If a database backup is written to a UNC path but FlushFileBuffers only guarantees that the local system file cache is flushed, Microsoft SQL Server is exposed to data loss.</p> <p>For more information about write ordering requirements, consult the "Write Ordering," "FlushFileBuffers," "Backup Hardening," and "Remote Mirroring" sections of the white paper, <i>Microsoft SQL Server 2000 I/O Basics</i>.</p>	

Core 1.06 Torn I/O protection	Required
<p>An Always On solution must provide sector alignment and sizing in a way that prevents torn I/O. This includes splitting I/Os across various I/O entities in the I/O path.</p> <p>Additionally, a solution must accurately report sector size to Windows I/O APIs. Accurately reporting sector size helps prevent sector size mismatches and avoid torn writes. For example, a drive that does 4 KB writes reports 512 bytes while the drive performs a read/write of the 4 KB sectors. This inaccuracy in reporting sector size can create a condition where data is lost and exposed as torn writes. An Always On solution must document configurations in such a way that use sector sizes from the sector size list that is supported by Microsoft SQL Server: 512, 1024, 2048, and 4096 bytes.</p> <p>To indicate when a torn-write situation occurs, it is recommended that the solution log appropriate warning events.</p> <p>The Always On solution white paper must include information about the configuration requirements needed for the solution to meet the torn I/O requirements.</p> <p>For more information, consult the “Torn I/O,” “Log Parity,” and “Sector Size” sections located in Chapter 2 of the white paper, <i>Microsoft SQL Server I/O Basics</i>.</p>	

Core 1.07 NTFS support	Required
<p>You must support NTFS capabilities. This includes, but is not limited to, the following:</p> <ul style="list-style-type: none"> • Sparse Files • File Streams • Encryption • Compression • All Security Properties <p>You must support sparse files on NTFS-based file systems. Microsoft SQL Server 2005 uses sparse files in support of DBCC CHECK* commands and snapshot databases.</p> <p>Common copy and compression utilities may not honor sparse file metadata, but instead copy all bytes, ignoring the sparse allocations, and requiring full storage space. Vendors may choose to provide utilities to copy or move sparse files without destroying the sparse file intent.</p> <p>Support must be provided for streams on NTFS-based files. Microsoft SQL Server 2005 uses sparse file streams files in support of DBCC CHECK* commands.</p>	

Core 1.07 Partner Enhanced Escalation Process membership	Recommended
<p>The Partner Enhanced Escalation Process (PEEP) is a joint escalation process established between Microsoft and partners. The non-binding memorandum (MOU) establishes direct, cross company joint escalation assistance and issue management.</p>	

Appendix B: References

General reference

Microsoft SQL Server 2000 I/O Basics (applies to SQL Server 7.0, SQL Server 2000, and SQL Server 2005)

EMC Celerra, CLARiiON, and Symmetrix storage provide solutions for Microsoft SQL Server Always On Technologies. For details refer to:

http://www.emc.com/solutions/microsoft/sql_server/on_program.jsp

EMC customer documentation

EMC provides documentation to existing customers through the Powerlink[®] site (<http://Powerlink.EMC.com/>), a password-protected customer- and partner-only extranet.

For more information on EMC Celerra and integration with Microsoft SQL Server consult:

- ♦ *EMC Solutions for Microsoft SQL Server 2005 EMC Celerra NS Series iSCSI – Applied Best Practices Guide*

Copyright © 2008 EMC Corporation. All Rights Reserved.

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

THE INFORMATION IN THIS PUBLICATION IS PROVIDED "AS IS." EMC CORPORATION 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 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.