EMC MISSION CRITICAL INFRASTRUCTURE
FOR MICROSOFT SQL SERVER 2012

EMC Symmetrix VMAX 10K, EMC FAST VP, SQL Server AlwaysOn Availability Groups, VMware vSphere

- Storage optimization with Symmetrix VMAX 10K FAST VP
- AlwaysOn Availability Groups—Readable secondary replicas
- VMware vSphere server virtualization

EMC Solutions Group

Abstract

This white paper describes how to deploy and implement SQL Server 2012 Availability Groups on the EMC® Symmetrix VMAX® 10K platform. It demonstrates the performance improvements of intelligent EMC Fully Automated Storage Tiering for Virtual Pools (FAST VP) technology on EMC Symmetrix VMAX 10K, which automatically promotes hot data to a high-performance Flash virtual pool and demotes cold data to a high-capacity SATA virtual pool. It also shows the integration of Symmetrix VMAX 10K with VMware® vSphere 5 and the ease of maintenance of storage and applications in a virtualized environment.

July 2012
# Table of contents

**Executive summary** ............................................................................................................................... 5  
- Business case ....................................................................................................................................... 5  
- Solution overview ............................................................................................................................... 5  
- Key results ........................................................................................................................................ 6  

**Introduction** ...................................................................................................................................... 7  
- Purpose ................................................................................................................................................ 7  
- Scope ............................................................................................................................................... 7  
- Audience ......................................................................................................................................... 7  
- Terminology ..................................................................................................................................... 7  

**Technology overview** .......................................................................................................................... 9  
- Introduction to the key components .................................................................................................. 9  
- EMC Symmetrix VMAX 10K storage array ....................................................................................... 9  
- FAST VP on Symmetrix VMAX 10K ................................................................................................. 10  
- EMC Virtual Storage Integrator ....................................................................................................... 10  
- EMC PowerPath/VE .......................................................................................................................... 10  
- VMware vSphere 5 ............................................................................................................................ 11  
- Microsoft SQL Server 2012 ............................................................................................................... 11  
  - AlwaysOn ....................................................................................................................................... 11  
  - AlwaysOn Availability Groups ...................................................................................................... 11  
  - Availability replicas and roles ........................................................................................................ 12  
  - Availability modes .......................................................................................................................... 13  
  - Availability group listener ............................................................................................................. 14  
  - Indirect checkpoints ...................................................................................................................... 14  

**Solution configuration** .......................................................................................................................... 15  
- Solution overview .............................................................................................................................. 15  
- Solution architecture ......................................................................................................................... 15  
- SQL Server 2012 user profile and design .......................................................................................... 17  
- Hardware resources ............................................................................................................................ 20  
- Software resource .............................................................................................................................. 20  

**AlwaysOn Availability Groups configuration** ..................................................................................... 21  
- Overview .......................................................................................................................................... 21  
- Prerequisites .................................................................................................................................... 21  
- WSFC setup ..................................................................................................................................... 21  
- WSFC quorum modes and voting configuration .............................................................................. 21  
- Enabling AlwaysOn .......................................................................................................................... 22  
- AlwaysOn Availability Groups design ............................................................................................ 23
Executive summary

Business case
Microsoft has released SQL Server 2012 with an array of new and improved capabilities with emphasis on reliability, availability, serviceability, and performance, all of which are flexible and ready to use.

A critical success factor for any SQL Server deployment is to deliver a high-performance, resilient, and an efficient storage solution that makes the most of the storage hardware and drives the best return on investment (ROI) for the customer.

Working together, Microsoft and EMC supply the critical technical necessities to deliver high-performance enterprise-class availability solutions for Microsoft SQL Server 2012 environments. By combining the power of Microsoft SQL Server 2012 AlwaysOn Availability Groups with EMC® Symmetrix VMAX® 10K storage in a virtualized environment, this solution not only introduces new methods of reduced downtime in a failure, but it also enables faster and easier data accessibility that improves business intelligence and analytics within the business.

Symmetrix VMAX 10K combines simplicity with automatic performance self-tuning in an advanced enterprise-class storage system at a very cost-effective price. Especially when enabled with EMC Fully Automated Storage Tiering for Virtual Pools (FAST VP), which transparently automates storage tiering to ensure optimal performance for critical workloads, VMAX 10K can help meet the storage demands of SQL Server 2012 with reduced cost and simplified management.

Solution overview
The purpose of this solution is to illustrate the performance and function of enterprise-class Microsoft SQL Server instances in a virtualized environment on the EMC Symmetrix VMAX 10K storage platform running SQL Server 2012 with the AlwaysOn feature.

This solution uses a component of the new SQL Server 2012 AlwaysOn transactional replication technology known as Availability Groups to provide readable secondary replicas of production databases.

This solution's SQL Server dataset is composed of four mission-critical, highly active online transaction processing (OLTP) databases, totaling 1.8 TB of data, which are replicated to a readable secondary copy using the SQL Server 2012 AlwaysOn Availability Groups feature.

This solution demonstrates Symmetrix VMAX 10K as a simple, efficient, and high-availability storage platform for enterprise-class SQL Server 2012 infrastructures in a virtualized environment. The integration of Symmetrix VMAX 10K with virtualization provides simple storage management for a virtualized SQL Server 2012 deployment, enhanced by both Symmetrix VMAX 10K-specific tools, such as Symmetrix Management Console (SMC), and free-of-charge integration tools for VMware, such as EMC Virtual Storage Integrator (VSI).

This solution also shows that EMC FAST VP provides the ability to automatically move the highly active data (hot data) to a Flash drive tier while moving the less-accessed data to a SATA drive tier, ensuring that data is always “in the right place at the right time.”
This solution demonstrates the following results:

- EMC Symmetrix VMAX 10K provides validated and trusted storage environments for Microsoft SQL Server 2012.
- Symmetrix VMAX 10K can easily service multiple SQL instances with different workloads leveraging AlwaysOn readable secondary databases.
- EMC FAST VP allows Symmetrix VMAX 10K storage arrays to maximize storage efficiency by providing auto-tiering of SQL Server database storage.
- Symmetrix VMAX 10K integration with VMware provides simple and effective management and configuration of a SQL Server 2012 environment.
- With the well-designed storage layout and the flexibility and power of Symmetrix VMAX 10K, it is now possible to utilize additional features of AlwaysOn Availability Groups in providing accessibly to production data.
**Introduction**

**Purpose**
This white paper showcases the ability of EMC Symmetrix VMAX 10K storage to easily support heavy SQL Server OLTP workloads. An AlwaysOn Availability Group secondary replica provides access to live data without impacting production database performance.

Using EMC FAST VP auto-tiering technology ensures that SQL Server 2012 can sustain required performance for the primary database while providing read-only access to the secondary replica of the SQL Server 2012 AlwaysOn Availability Groups.

**Scope**
This white paper:

- Showcases the performance and function of enterprise-class SQL Server instances in a virtualized VMware environment running SQL Server 2012 with the AlwaysOn feature.
- Demonstrates storage performance optimization with FAST VP.
- Shows Symmetrix VMAX 10K capabilities and integration with a virtualized SQL Server environment.

**Audience**
This white paper is intended for SQL Server 2012 database administrators and storage architects involved in planning, architecting, or administering an environment with Symmetrix VMAX 10K.

**Terminology**
This paper includes the following terminology.

**Table 1. Terminology**

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Availability group</td>
<td>A high availability and disaster recovery solution that provides an enterprise-level alternative to database mirroring. Introduced in SQL Server 2012, the AlwaysOn Availability Groups feature maximizes the availability of a set of user databases.</td>
</tr>
<tr>
<td>Availability replica</td>
<td>An instantiation of an availability group that is hosted by a specific instance of SQL Server and maintains a local copy of each availability database that belongs to the availability group. Two types of availability replicas exist—a single primary replica (refer Primary replica) and up to four secondary replicas (refer to Readable secondary replica).</td>
</tr>
<tr>
<td>Data synchronization</td>
<td>Process by which changes to a primary database are reproduced on a secondary database.</td>
</tr>
<tr>
<td>Term</td>
<td>Definition</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>FAST VP</td>
<td>Fully Automated Storage Tiering for Virtual Pools. A feature of Symmetrix storage arrays that automates the identification of data volumes for the purpose of allocating or reallocating business application data across different performance and capacity tiers within the storage array.</td>
</tr>
<tr>
<td>Indirect checkpoints</td>
<td>New for Microsoft SQL Server 2012, indirect checkpoints provide a configurable database-level alternative to automatic checkpoints.</td>
</tr>
<tr>
<td>OLTP</td>
<td>Online transaction processing. Typical applications of OLTP include data entry and retrieval transaction processing.</td>
</tr>
<tr>
<td>Primary replica</td>
<td>The availability replica that makes the primary databases available for read/write connections from clients and sends transaction log records for each primary database to every secondary replica.</td>
</tr>
<tr>
<td>Readable secondary replica</td>
<td>Secondary replica databases configured to allow read-only client connections.</td>
</tr>
<tr>
<td>Reseeding</td>
<td>Process of copying a database from a primary replica to corresponding secondary replicas.</td>
</tr>
<tr>
<td>RDM</td>
<td>Raw device mapping.</td>
</tr>
<tr>
<td>Symmetrix VMAX 10K</td>
<td>A solution for providing the highest levels of reliability and availability in a scalable storage for medium-sized enterprises and larger organizations. With support for multiple disk technologies, EMC Virtual Provisioning™ storage allocation, and support for Fully Automated Storage Tiering (FAST), Symmetrix VMAX 10K provides an industry-leading solution for the market.</td>
</tr>
<tr>
<td>TPS</td>
<td>Transactions per second.</td>
</tr>
</tbody>
</table>
Technology overview

Introduction to the key components

This section provides an overview of the technologies that are used in this solution:

- EMC Symmetrix VMAX 10K storage array
- FAST VP on Symmetrix VMAX 10K
- EMC Virtual Storage Integrator
- EMC PowerPath®/VE
- VMware® vSphere®5
- Microsoft SQL Server 2012
- AlwaysOn Availability Groups

EMC Symmetrix VMAX 10K storage array

Symmetrix VMAX 10K is a new enterprise storage platform built to provide leading high-end virtual storage capabilities to a growing number of IT organizations and service providers with demanding storage requirements and limited resources.

Symmetrix VMAX 10K is designed for easy installation, setup, and use. It is ideal for customers who need increased failure-mode performance and is an ideal entry into a Symmetrix-based storage infrastructure.

Leveraging the Virtual Matrix Architecture, Symmetrix VMAX 10K provides enterprise-level reliability, availability, and serviceability.

Symmetrix VMAX 10K includes preconfiguration for easy setup on the same day the array is received, with an installation that takes less than four hours.

Symmetrix VMAX 10K is a 100 percent virtually provisioned system. Virtual provisioning gives a host, application, or file system the view that it has more storage than is physically provided. Physical storage is allocated only when the data is written, rather than when the application is initially configured. This eliminates manual calculations and can also reduce power and cooling costs by decreasing the amount of idle storage capacity in the array.

Adding FAST VP for fully automated tiered storage makes the installation and daily operation easier for IT organizations with limited IT resources and staff.

Building on EMC’s industry-leading VMware integration, Symmetrix VMAX 10K now offers even more efficient enterprise storage because of new integration with the VMware vSphere 5 cloud infrastructure platform. EMC Virtual Storage Integrator (VSI) for VMware simplifies the process of integrating EMC storage into a virtualized environment.
**FAST VP on Symmetrix VMAX 10K**

Symmetrix VMAX 10K integrates simple self-tuning into an advanced enterprise-class storage system at a very cost-effective price. By combining EMC subvolume, autotiering, and Virtual Provisioning technology, FAST VP enables storage administrators to implement automated policy-driven plans. These plans dynamically perform nondisruptive changes to tiers of different applications by ensuring that high-performance drives serve the hot spots of a volume or LUN and less expensive drives serve the inactive data.

With FAST VP, customers can achieve:

- Better performance with lower cost, fewer drives, less power and cooling, and a smaller footprint
- Maximum utilization of Flash drives for high-performance workloads
- Lower cost of storage by placing the less accessed data on SATA drives
- Radically simplified automated management in a tiered environment

**EMC Virtual Storage Integrator**

EMC VSI is a VMware vCenter™ plug-in available to all VMware users with EMC storage in their environment.

VSI simplifies the process of mapping VMware vSphere™ data stores to LUNs and NFS shares on EMC storage. It also helps pinpoint the location of virtual machines and raw device mapping files on the array. The VMware administrator's visibility into the storage layer can help troubleshoot storage performance issues and simplify communication among server, storage, and virtualization teams.

**EMC PowerPath/VE**

In this solution, EMC PowerPath/VE software was used on a VMware vSphere host in a VMware high availability (HA) cluster. PowerPath allows the host to connect to a LUN through more than one storage processor port; this is known as multipathing. PowerPath optimizes multipathed LUNs through load-balancing algorithms. Port-load balancing equalizes the I/O workload across all available channels. Hosts connected to Symmetrix VMAX 10K benefit from multipathing.

The advantages of multipathing are:

- Failover from port to port on the same storage processor, maintaining an even system load and minimizing LUN trespassing
- Port-load balancing across storage processor ports and host bus adapters (HBAs)
- Higher bandwidth attachment from host to storage system
VMware vSphere 5

VMware vSphere 5 is the latest VMware virtual data center operating system. It continues to transform IT infrastructure into the most efficient, shared, and on-demand utility, with built-in availability, scalability, and security services for all applications and simple, proactive automated management.

vSphere 5 has the following scalability and performance enhancements, which enable a virtual machine to leverage more resources from the hypervisor:

- 32-way virtual Symmetrical Multiprocessing (SMP).
  ESXi 5.0 supports virtual machines with up to 32 virtual CPUs, which lets you run larger CPU-intensive workloads such as SQL Server 2012 on the VMware ESXi operating system.
- 1 TB virtual machine RAM. You can assign up to 1 TB of RAM to ESXi 5.0 virtual machines.
- Up to 1,000,000 IOPS for each VMware vSphere server.
- More than 36 GB/s network bandwidth.

Building on EMC’s industry-leading VMware integration, Symmetrix VMAX 10K now offers even more efficient enterprise storage through new integration with VMware vSphere 5 cloud infrastructure.

Microsoft SQL Server 2012

Microsoft SQL Server 2012 is the latest version of Microsoft database management and analysis system for e-commerce, line-of-business, and data warehousing solutions. This solution depicts one of the latest replication features of SQL Server 2012, AlwaysOn (specifically, AlwaysOn Availability Groups).

AlwaysOn

SQL Server AlwaysOn refers to the new comprehensive HA and disaster recovery solution for SQL Server 2012. AlwaysOn presents new and enhanced capabilities for both specific databases and entire instances, providing flexibility to support various high availability configurations through:

- AlwaysOn failover cluster instances
- AlwaysOn Availability Groups

This solution explores AlwaysOn Availability Groups, with a focus on the transaction-level replication feature to provide access to near-live readable secondary replicas of production databases.

AlwaysOn Availability Groups

AlwaysOn Availability Groups is an HA and disaster recovery solution introduced in SQL Server 2012, enabling administrators to maximize availability for one or more user databases. SQL Server instances are configured so that a single primary database or a group of primary databases can have up to four secondary database copies residing on Windows Server failover cluster (WSFC) nodes.
Availability replicas and roles

Availability groups consist of a set of two or more failover partners referred to as availability replicas. Each availability replica is hosted on a separate instance of SQL Server which, in turn, resides on a separate node of a WSFC. Each of the SQL Server instances is a SQL Server failover cluster instance (FCI), or a stand-alone instance with AlwaysOn Availability Groups enabled, as shown in Figure 1.

Figure 1. SQL Server AlwaysOn Availability Groups

Each availability replica hosts a copy of the availability databases in the availability group. Each availability replica is assigned an initial role (primary or secondary):

- **Primary replica**
  Holds the primary role and there can be only one. A primary replica hosts the read-write databases that are known as the primary databases.

- **Secondary replica**
  There can be up to four replicas, each holding the secondary role and hosting the read-only databases.

Any secondary replica can become the primary replica as the result of failover.

**Readable secondary replicas**

Secondary replicas can be configured so that, while in the secondary role, they accept read-only client connections to local databases. These secondary databases are referred to as readable secondary replicas. The data on secondary replicas is near realtime. Secondary databases are not set to read-only. Unlike a read-only database, which is static, a secondary replica is dynamic and continuously changing as the corresponding primary database changes.
Directing read-only connections to readable secondary replicas provides the following benefits:

- Offloads your secondary read-only workloads from your primary replica, which conserves its resources for your mission-critical workloads
  
  **Note** If you have a mission-critical read-only workload or a workload that cannot tolerate latency, it is best to run it on the primary replica.

- Improves your return on investment (ROI) for the systems that host readable secondary replicas

In addition, readable secondary replicas provide robust support for read-only operations, as follows:

- Temporary statistics on readable secondary database optimize read-only queries.

- Read-only workloads use row versioning to remove blocking contention on the secondary databases. All queries that run against the secondary databases are automatically mapped to the snapshot isolation transaction level, even when other transaction isolation levels are explicitly set. Also, all locking hints are ignored. This eliminates read/write contention.

**Availability modes**

Each availability group has an availability mode setting, which determines whether the primary replica has to wait for a transaction to be committed to a database before the corresponding secondary replica has written the transaction log to disk (hardening the log).

**AlwaysOn Availability Groups** support two modes:

- **Asynchronous-commit mode**
  
  In asynchronous-commit mode, the primary replica commits a transaction without acknowledgement that an asynchronous-commit replica has hardened the log. Asynchronous-commit mode minimizes transaction latency, allowing the secondary databases to lag behind the primary. This makes data loss possible.

- **Synchronous-commit mode**
  
  In synchronous-commit mode, the primary replica waits for acknowledgment that a synchronous-commit secondary replica has hardened the log before committing a transaction. Synchronous-commit mode increases transaction latency but protects against data loss, meaning that as long as the secondary databases are in a synchronized state with the primary database, committed transactions are fully protected.
**Availability group listener**

After creating an availability group listener, client connectivity can be provided to a database of the availability group. This unique DNS name serves as a virtual network name (VNN) and provides a set of resources that are attached to an availability group to direct client connections to the appropriate availability replica.

For a detailed description of availability group listeners, refer to Microsoft Developer Network Platforms.

**Indirect checkpoints**

Indirect checkpoints, new in SQL Server 2012, provide a configurable database-level alternative to automatic checkpoints. In the event of a system crash, indirect checkpoints provide potentially faster, more predictable recovery time than automatic checkpoints.

Indirect checkpoints offer the following advantages:

- Indirect checkpoints can reduce overall database recovery time.
- Indirect checkpoints enable you to control database recovery time by factoring in the cost of random I/O during Redo. This enables a server instance to stay within an upper bound for recovery times for a given database (except when a long-running transaction causes excessive Undo times).
- Indirect checkpoints reduce checkpoint-related I/O spiking by continually writing dirty pages to disk in the background.

In testing, by implementing indirect checkpoints, I/O spiking was significantly reduced. Especially on the secondary copy, checkpoint spikes were almost undetectable.

**Note**  
Online transactional workloads on databases configured with indirect checkpoints might experience performance degradation. This is because the background writer for indirect checkpoints sometimes increases the total write load for server instances.

For a detailed description of indirect checkpoints, refer to Microsoft Developer Network Platforms.
Solution configuration

Solution overview
This solution includes a two-engine Symmetrix VMAX 10K storage array with three tiers of storage managed and controlled by FAST VP policies. The storage array supports two SQL Server instances in a virtualized environment running SQL Server 2012 with AlwaysOn Availability Groups.

The solution features an AlwaysOn Availability Group readable copy at the secondary site. The availability group dedicates the primary replica to mission-critical OLTP workloads and segregates the demands of online analytical processing/business intelligence (OLAP/BI) on the secondary replica.

Solution architecture
The solution design included the following physical components:

- Two vSphere ESXi servers, each hosting one SQL Server virtual machine
- Symmetrix VMAX 10K SAN storage
- Three FAST VP-enabled storage tiers
Figure 2 shows an overview of the solution architecture.

Figure 2. Solution architecture
The environment consists of two SQL Server 2012 virtual machines on VMware vSphere 5 and Symmetrix VMAX 10K. Designed with 5 to 10 percent overhead, this virtualized SQL Server 2012 environment provides excellent performance and throughput with a reduced footprint. This results in greater server resource efficiency and reduces power and cooling costs.

vSphere 5 provides a new feature that supports up to 32 processors for virtual machines. This enables enterprise-class SQL Server 2012 to handle more transactions by providing more powerful processing capabilities.

Table 2 shows the profile of the SQL Server configuration.

**Table 2. SQL Server profile**

<table>
<thead>
<tr>
<th>Profile</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>SQL database capacity</td>
<td>4 user databases – 1.8 TB in total</td>
</tr>
<tr>
<td></td>
<td>1x50 GB (5,000 user)</td>
</tr>
<tr>
<td></td>
<td>1x250 GB (25,000 user)</td>
</tr>
<tr>
<td></td>
<td>1x500 GB (50,000 user)</td>
</tr>
<tr>
<td></td>
<td>1x1 TB (100,000 user)</td>
</tr>
<tr>
<td>Number of SQL Server instances</td>
<td>2</td>
</tr>
<tr>
<td>Number of user databases for each virtual machine</td>
<td>4</td>
</tr>
<tr>
<td>Number of virtual machines</td>
<td>2</td>
</tr>
<tr>
<td>SQL Server virtual machine config</td>
<td>Primary server: 24 Intel Xeon X5670 CPUs (2.27 GHz/4x8 cores)/64 GB memory</td>
</tr>
<tr>
<td></td>
<td>Secondary server: 16 Intel Xeon X5670 CPUs (2.27 GHz/4x8 cores)/64 GB memory</td>
</tr>
<tr>
<td>Concurrent users</td>
<td>Mixed to simulate hot, warm, and cold workloads across the databases</td>
</tr>
</tbody>
</table>
Table 3 and Figure 3 show the SQL Server LUN design. The design follows these best practices:

- Distribute the data and log to different LUNs on the shared thin pools
- Use separate file groups for large databases (tpce2, tpce3, and tpce4)
- Ensure configuration consistencies for the primary and secondary servers (data, log, and tempDB).

### Table 3. SQL Server LUN design

<table>
<thead>
<tr>
<th>Item</th>
<th>Component</th>
<th>LUN capacity (GB)</th>
<th>Quantity</th>
<th>Total capacity (GB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SQL 01</td>
<td>TempDB</td>
<td>100</td>
<td>4</td>
<td>400</td>
</tr>
<tr>
<td>(Primary</td>
<td>TempDB log</td>
<td>100</td>
<td>1</td>
<td>100</td>
</tr>
<tr>
<td>server)</td>
<td>database1</td>
<td>300</td>
<td>1</td>
<td>300</td>
</tr>
<tr>
<td></td>
<td>database 2</td>
<td>100</td>
<td>6</td>
<td>600</td>
</tr>
<tr>
<td></td>
<td>database 3</td>
<td>150</td>
<td>6</td>
<td>900</td>
</tr>
<tr>
<td></td>
<td>database 4</td>
<td>450</td>
<td>3</td>
<td>1,350</td>
</tr>
<tr>
<td></td>
<td>database_log1</td>
<td>120</td>
<td>1</td>
<td>120</td>
</tr>
<tr>
<td></td>
<td>database_log2</td>
<td>175</td>
<td>1</td>
<td>175</td>
</tr>
<tr>
<td></td>
<td>database_log3</td>
<td>200</td>
<td>1</td>
<td>200</td>
</tr>
<tr>
<td></td>
<td>database_log4</td>
<td>400</td>
<td>1</td>
<td>400</td>
</tr>
<tr>
<td>SQL 02</td>
<td>TempDB</td>
<td>100</td>
<td>4</td>
<td>400</td>
</tr>
<tr>
<td>(Secondary</td>
<td>TempDB log</td>
<td>100</td>
<td>1</td>
<td>100</td>
</tr>
<tr>
<td>server)</td>
<td>database 1 (replica)</td>
<td>300</td>
<td>1</td>
<td>300</td>
</tr>
<tr>
<td></td>
<td>database 2 (replica)</td>
<td>100</td>
<td>6</td>
<td>600</td>
</tr>
<tr>
<td></td>
<td>database 3 (replica)</td>
<td>150</td>
<td>6</td>
<td>900</td>
</tr>
<tr>
<td></td>
<td>database 4 (replica)</td>
<td>450</td>
<td>3</td>
<td>1,350</td>
</tr>
<tr>
<td></td>
<td>database_log1 (replica)</td>
<td>120</td>
<td>1</td>
<td>120</td>
</tr>
<tr>
<td></td>
<td>database_log2 (replica)</td>
<td>175</td>
<td>1</td>
<td>175</td>
</tr>
<tr>
<td></td>
<td>database_log3 (replica)</td>
<td>200</td>
<td>1</td>
<td>200</td>
</tr>
<tr>
<td></td>
<td>database_log4 (replica)</td>
<td>400</td>
<td>1</td>
<td>400</td>
</tr>
</tbody>
</table>
The following list shows the Windows and SQL Server 2012 configuration of each virtual machine. Default values were used for all other settings:

- **Lock Pages in Memory** privilege was granted to the SQL startup account. Refer to [Pre-Configuration Database Optimizations](#) for more information.

- An NTFS allocation unit size of 64 KB was specified for the user data device. Refer to [SQL Server Best Practices Article](#) for more information.

- Indirect checkpoint feature was used to reduce I/O spikes for SQL Server 2012.
Table 4 shows the hardware resources used to validate this solution.

<table>
<thead>
<tr>
<th>Equipment</th>
<th>Quantity</th>
<th>Configuration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symmetrix VMAX 10K</td>
<td>1</td>
<td>VMAX 10K 7.2K 2000 GB (SATA)—Quantity: 64</td>
</tr>
<tr>
<td></td>
<td></td>
<td>VMAX 10K 15K 450 GB (FC)—Quantity: 120</td>
</tr>
<tr>
<td></td>
<td></td>
<td>VMAX 10K Flash 200 GB (SSD)—Quantity: 20</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Engines: 2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cache: 96 GB (Mirrored)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FAST VP enabled</td>
</tr>
<tr>
<td>32-core Intel Nehalem-EX</td>
<td>2</td>
<td>Primary replica SQL Server virtual machine</td>
</tr>
<tr>
<td>servers</td>
<td></td>
<td>24-core 192 GB RAM</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Secondary replica SQL Server virtual machine</td>
</tr>
<tr>
<td>Fibre Channel switches</td>
<td>2</td>
<td>8 GB FC switches</td>
</tr>
<tr>
<td>GbE network switch</td>
<td>1</td>
<td>Network switch—24 ports</td>
</tr>
</tbody>
</table>

Table 5 shows the software resources used in this solution.

<table>
<thead>
<tr>
<th>Software</th>
<th>Quantity</th>
<th>Version</th>
<th>Purpose</th>
</tr>
</thead>
</table>
| Windows Server 2008 R2       | 4        | 2008 R2 x64 SP1 | Source and target virtual machines  
|                              |          |          | Load virtual machines  
|                              |          |          | Domain controller  
|                              |          |          | VMware vCenter server                                             |
| VMware ESXi 5                | 2        | 5.0.0    | Two-node ESXi cluster to host virtual machines                         |
| VMware vCenter 5             | 1        | 5.0.0    | Virtual center management server                                          |
| SQL Server 2012              | 2        | RC1      | Database software                                                |
| EMC PowerPath                | 2        | 5.5      | Advanced multipathing for SQL production and secondary host HBAs         |
| VSI                          | 1        | 5.1      | Provisioning and viewing storage from vCenter to simplify storage management |
AlwaysOn Availability Groups configuration

Overview
The following sections outline the process for implementing the AlwaysOn Availability Groups across two SQL Server instances. The solution team connected both SQL servers to the same Symmetrix VMAX 10K storage array.

Prerequisites
Refer to the AlwaysOn prerequisites in the Appendix.

WSFC setup
As part of the configuration for SQL Server 2012 availability group creation, we configured a two-node cluster on Windows Server 2008 R2. You can find the steps required in the creation of Windows failover cluster in the Appendix.

WSFC quorum modes and voting configuration
SQL Server AlwaysOn Availability Groups take advantage of Windows Server failover Cluster (WSFC) technology. WSFC uses a quorum-based approach to monitoring overall cluster health and provide maximum node-level fault tolerance.

This solution used the quorum configuration of **Node and File Share Majority**. We configured a file share witness on a host outside of the cluster, created a share called **Witness** on a server named SMC, and gave the cluster named **spsqlcluster** read/write permissions to the share at both the share level and NTFS level, as shown in Figure 4.

After configuring the shared folder with the appropriate permissions, we then changed the quorum type. We did this through the **Failover Cluster Manager** by following these steps:

1. Right-click the cluster, select More Actions and Configure Cluster Quorum Settings.
2. Select **Node and File Share Majority** as the quorum configuration (Figure 4).
3. Follow the wizard to completion, entering the path to the created file share called **Witness**.

![Configure Cluster Quorum wizard](image)

**Figure 4. Configure Cluster Quorum wizard**
After completing the remaining steps in the wizard, the Quorum Configuration showed **Node and File Share Majority** in the Failover Cluster Manager, as shown in Figure 5.

![Figure 5. Failover Cluster Manager—Quorum configuration](image)

One of the major considerations for node and file share majority quorum is where to place the file share witness. In this solution, we placed the file share majority quorum on the secondary site host.

To enable AlwaysOn HA in SQL Server Configuration Manager, follow these steps:

1. Open **SQL Server Configuration Manager**.
2. Right-click **SQL Server Engine Services**.
3. Click **AlwaysOn High Availability** and select the **Enable AlwaysOn Availability Groups** checkbox, as shown in Figure 6. You must do this for both instances on the two nodes of cluster.

**Note:** Enabling AlwaysOn requires SQL Server services to be restarted on both the nodes and instances.

![Figure 6. Enabling AlwaysOn High Availability in SQL Server Configuration Manager](image)
In this solution, to meet the requirement to define a set of user databases as a single unit for failover, the solution team put four databases into the same availability group.

On the primary replica, we created a fileshare folder accessible to the secondary replica. We backed up databases and logs to the share folder locally, and then we restored the databases to a remote secondary server on a 10 Gbps network. We wrote each database’s backup to four files and ran the four database backups concurrently. There was a total of 1.8 TB data allocated in databases of 1 TB, 500 GB, 250 GB, and 50 GB within one availability group.

To emphasize best possible Recovery Point Objective (RPO) over performance, we configured the secondary replica as a synchronous replica as shown in Figure 7. In the synchronous mode, there is no data loss on both sites after failover.

![Figure 7. Availability status dash board](image)
To allow read-only connections to the secondary replica, set **Readable secondary** property to **Yes** for the secondary replica as shown in Figure 8.

![Figure 8. Setting readable secondary replica property](image)

We used indirect checkpoints to reduce the performance impact caused by running auto/manual checkpoints. The **Target Recovery Time (Seconds)** for each database is set to **450**, as shown in Figure 9.

![Figure 9. Setting target recovery time](image)
As can be seen in the average disk latency checkpoints and the disk transfers per second checkpoints in Figure 10, by using auto/manual checkpoints, high I/O spiking appeared and lasted for more than 1 minute at the checkpoints.

Figure 10. I/O spiking when using auto/manual checkpoints
When configuring indirect checkpoints with databases, there was no visible I/O spiking, and the average latency was quite low, as shown in Figure 11.

![Without I/O spike with indirect checkpoint](image)

**Figure 11.** No visible I/O spiking when using indirect checkpoints
Storage design

Overview
Symmetrix Virtual Provisioning technology was used for storage provisioning. Thin data pool devices were pre-configured and bound to the appropriate thin pool. Because the storage resources were allocated across all data devices in the thin pool, I/Os were widely striped. As a result, application I/O workload against the thin devices was evenly distributed across all resources.

Symmetrix VMAX 10K design guidelines
The following list provides general design guidelines for running OLTP applications on Symmetrix VMAX 10K:

- The back end and front end share the same processors. Therefore, leverage as many front-end ports as possible to distribute the workload on these CPUs.
- The layout of the physical drives is important. Therefore, use default configurations if possible.

FAST VP design for SQL Server 2012 Availability Groups
Outlined below are some general guidelines for sizing a consolidated workload with FAST VP:

- Flash drives provide the best performance for workloads with higher numbers of random reads. Because of the sequential write nature of the database log LUNs, enable FAST VP only for data LUNs to effectively utilize Flash drives.
- For large thin devices, use eight-way striped metadevices for better performance.
- Perform sizing according to FAST VP policy requirements. Guidelines recommend the following selections:
  - RAID 5 Flash drives for best total cost of ownership (TCO)
  - RAID 6 SATA drives for best data protection
  - RAID 1 FC drives for best performance
- Use the same FAST VP policy for both the primary and secondary sites to ensure consistent performance in case of a failover. After failover, FAST VP continues to gather and analyze statistics and performs the appropriate tier allocation changes to meet performance and storage tier allocation requirements.

Table 6 shows the FAST VP tiers and policy used in this solution.

<table>
<thead>
<tr>
<th>Tier name</th>
<th>Drive technology</th>
<th>FAST VP settings</th>
</tr>
</thead>
<tbody>
<tr>
<td>MS_Flash</td>
<td>200 GB Flash drives</td>
<td>30%</td>
</tr>
<tr>
<td>MS_FiberChannel</td>
<td>450 GB 15K FC drives</td>
<td>60%</td>
</tr>
<tr>
<td>MS_SATA</td>
<td>2 TB 7.2K SATA drives</td>
<td>10%</td>
</tr>
</tbody>
</table>
The storage design objective was to achieve the following host performance:

- Data LUN latency of less than 20 ms
- Total IOPS of around 30,000 for both sites

The disk group configuration is shown in Table 7.

### Table 7. Disk group configuration

<table>
<thead>
<tr>
<th>Drive type</th>
<th>RAID type</th>
<th>Number of drives required</th>
<th>Recommended maximum IOPS</th>
<th>Supported maximum IOPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flash</td>
<td>R5</td>
<td>16</td>
<td>1,800</td>
<td>28,800</td>
</tr>
<tr>
<td>FC</td>
<td>R1</td>
<td>112</td>
<td>150</td>
<td>16,800</td>
</tr>
<tr>
<td>SATA</td>
<td>R6</td>
<td>60</td>
<td>50</td>
<td>3,000</td>
</tr>
</tbody>
</table>

The solution team placed two FAST VP thin pools (primary and secondary) across all three tiers (three virtual pools). These two thin pools served the SQL Server tempDB and log, database, and log LUNs on each site of the availability group:

- Primary pool: For LUNs on the primary SQL Server
- Secondary pool: For LUNs on the secondary SQL Server

Each SQL Server had its own datastore and virtual disk for its operating system. SQL Server OS/boot LUNs used a VMware virtual disk (VMDK) on its own datastore; all database LUNs used physical raw device mapping (pRDM) to allow direct guest operating system access to the hardware.

Because SQL Server 2012 can automatically detect Non-Uniform Memory Architecture (NUMA), and SQL Server processor and memory allocation can be optimized for NUMA, this solution implemented the following design best practices:

- Keep the number of physical cores and vCPUs in a 1:1 ratio. Ensure there are no overcommitted CPUs.
- Consider the NUMA node size when sizing virtual machines. For the primary site, assign 4x6 virtual CPUs to the SQL Server virtual machine. For the secondary site, assign 4x4 virtual CPUs to the SQL Server virtual machine.
- Fully reserve RAM for SQL Server virtual machines to avoid any memory ballooning.

### Improving the throughput of HBA ports in an ESXi environment

To improve the throughput of HBA cards on ESXi servers, we set the queue depth of the HBA to 64 (the default is 32). When setting the queue depth, consider the following factors:

- The number of virtual machines on the ESXi server
- The application I/O requirements
In the environment with only two or three virtual machines, a limited number of LUNs, and a high I/O-demanding application, changing the queue depth to 64 improved the I/O performance by over 30 percent in testing.

Notes

- The highest queue depth configurable for the HBA ports is 128, provided that the ESXi server has exclusive access to the LUNs configured with its ports.

- For ESXi servers with multiple virtual machines, the value of Disk.SchedNumReqOutstanding in VMware’s advanced options needs to match the queue depth.

For more information about ESXi server configuration, refer to the VMware knowledge base website and VMware ESX scalable storage performance manual.

Selecting SCSI driver type for data LUNs

VMware paravirtual SCSI (PVSCSI) adapters are high-performance storage drivers that can improve throughput and reduce CPU utilization. PVSCSI adapters are best suited for SAN environments, where hardware or applications drive high I/O throughput. There was up to a 25 percent improvement when using PVSCSI instead of the legacy LSI driver. Refer to the EMC white paper: EMC Storage Optimization and High Availability for Microsoft SQL Server 2008 R2 for more information.

As shown in Figure 12, the SCSI controller’s type changed to paravirtual to improve the driver efficiency (the default SCSI controller driver is LSI Logic SAS). LUNs were spread across all available SCSI drivers.

VSI supports Symmetrix VMAX and VMAX 10K, VNX and VNXe, CLARiiON CX4, and Celerra NS series for both block and file support. It is also an agentless architecture so that administrators do not need to install agents on the hosts. In this solution, VSI was installed on a VMware vSphere client host and was used to provide details of the storage for the SQL Server 2012 VMware hosts at both the primary and secondary sites. The vSphere client host had access to the Symmetrix VMAX 10K gatekeeper and was used to provide details of the SQL Server 2012 database storage. For more information about EMC VSI, refer to EMC VSI for VMware vSphere: Storage Viewer Product Guide 5.1.
Figure 13 shows the SQL Server hosts and the Symmetrix VMAX 10K storage mapping with details about the LUNs.

<table>
<thead>
<tr>
<th>LUN Identifier</th>
<th>Filename</th>
<th>Compatibility</th>
</tr>
</thead>
<tbody>
<tr>
<td>00000000000000000000000000000000</td>
<td>VSI_LUN_10K vmdk</td>
<td>prysl0103010201030405 vmdk</td>
</tr>
<tr>
<td>00000000000000000000000000000000</td>
<td>VSI_LUN_10K vmdk</td>
<td>prysl0103010201030405 vmdk</td>
</tr>
<tr>
<td>00000000000000000000000000000000</td>
<td>VSI_LUN_10K vmdk</td>
<td>prysl0103010201030405 vmdk</td>
</tr>
<tr>
<td>00000000000000000000000000000000</td>
<td>VSI_LUN_10K vmdk</td>
<td>prysl0103010201030405 vmdk</td>
</tr>
</tbody>
</table>

**Multipath design considerations**

Sixteen ports on Symmetrix VMAX 10K were zoned to the 8 GB SAN switch. Each ESXi server had two HBA ports connected to the SAN switch. The 16 ports were from two Symmetrix VMAX 10K engines. Each had eight ports zoned to one HBA port on the ESXi server. Thus, each device on the host had 16 paths. Through these paths, the workload could be evenly distributed between the two Symmetrix VMAX 10K engines.

EMC PowerPath/VE software optimized multipathing through load-balancing algorithms. Port load balancing equalizes the I/O workload across all available channels.

Symmetrix VMAX 10K has built-in integration with VMware. In this solution, we utilized this integration to simplify the configuration of storage for the virtualized SQL Server 2012 database environment. The tool used to leverage this integration is called Symmetrix Management Console (SMC). Combined with EMC VSI, this bidirectional "storage-to-server" and "server-to-storage" capability is unique to EMC.
Follow these steps to assign LUNs to the SQL Server virtual machines:

1. In the **Tasks** tab of the SMC wizard, select **Manage Virtual Server Storage**, as shown in Figure 14.

![Figure 14. SMC task for managing virtual server storage](image1)

2. From the **Manage Virtual Server Storage** wizard, select the ESX server and the virtual machine from the connected hosts as shown in Figure 15, and then click the **Add/Remove VM Storage** button.

![Figure 15. Managing virtual server storage](image2)
3. Assign LUNs to the virtual machine as shown in Figure 16. The LUNs are now ready for formatting.

Figure 16. Adding/removing virtual machine storage
Testing and validation

Test overview
The solution validates the performance and functionality of enterprise class SQL Server instances in a virtualized VMware environment running SQL Server 2012 with the AlwaysOn feature on the Symmetrix VMAX 10K storage platform.

About benchmark performance results
Benchmark results are highly dependent on 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.

Test objectives
The objectives of testing involved:

- Enabling FAST VP on the storage array to boost performance
- Comparing the performance with and without AlwaysOn Availability Groups created in synchronous mode
- Comparing the performance of pushing a read-only workload on the primary replica and the secondary readable replica.

Test scenarios
This solution tested the following scenarios:

- **Scenario 1: OLTP baseline without SQL Server 2012 AlwaysOn enabled**
  Running a baseline OLTP workload on the primary site without enabling SQL Server 2012 AlwaysOn. Enabling Symmetrix VMAX 10K using FAST VP to monitor how it automatically moves data to the most suitable tiers.

- **Scenario 2: AlwaysOn availability group creation**
  Creating an AlwaysOn availability group on SQL Server 2012 with the secondary replica in synchronous mode and with storage serviced by FAST VP.

- **Scenario 3: OLTP baseline with SQL Server 2012 AlwaysOn enabled**
  Running a baseline OLTP workload on the primary site with SQL Server 2012 AlwaysOn enabled on Symmetrix VMAX 10K using FAST VP.

- **Scenario 4: Adding OLAP workloads to the primary replica**
  Running OLAP workloads on the primary replica. The OLTP workload on the primary replica remained the same as the baseline.
- **Scenario 5: Directing OLAP workloads to the secondary replica**
  Running OLAP workloads on the secondary replica to relieve the workloads on the primary replica. The OLTP workload on the primary replica remained the same as the baseline.

**Test procedures**
The EMC solution team conducted a series of tests by running concurrent OLTP workloads against the target databases on the WSFC and stand-alone SQL Server instances.

The test procedure was as follows:

1. Enabled FAST VP with load running. When the performance was steady, we measured the baseline performance.
2. Created one availability group on all user databases and recorded the performance during this process.
3. Measured the performance with the OLTP workload on the primary-site databases.
4. Generated read-only workloads on the primary site and monitored the resulting performance impact.
5. Generated read-only workloads on the secondary site and monitored the resulting performance impact.

**Note** Workload profile parameters remained the same for all tests.

**Test results**

**Enabling Symmetrix VMAX 10K FAST VP**

Before FAST VP was enabled, all database LUNs were initially bound to the RAID 1 FC storage pool (112,450 GB 15K FC disks). After enabling the FAST VP policy, the highly active (hot) data was automatically moved to the FAST Flash tier, and the inactive (cold) data was moved to the SATA tier.

After enabling the FAST VP policy, the total IOPS increased from around 11,300 to 25,000 with better disk latency, more transactions from the client load, and no change in other performance counters. This represents a 121 percent improvement in storage performance, while no changes were required and no downtime was incurred on the SQL Server host.

Figure 17 shows disk throughput for baseline SQL Server load before and after enabling FAST VP. For this configuration, we configured 2-hour FAST VP sampling time and 2-hour analysis time. In total, it took 4 hours before data started to move based on the FAST VP policy. The data movement was stabilized after another 4 hours.

After enabling the FAST VP policy, the total transaction per second increased from around 1,100 to 2,300 with better disk latency, more disk IOPS, and no change in other performance counters.
Before enabling FAST VP, the database LUN was bound to the FC tier. All data was on the FC tier, as shown in Figure 18.

After enabling FAST VP, the storage reached a balanced state. The data was on different tiers, as shown in Figure 19.
**Availability group creation**

Before enabling the availability group, we created a database and transaction log backup of the databases in the availability group and then restored the databases to the secondary instance.

After the databases were restored, creation of the availability group took only few seconds. The secondary databases were synchronized with the primary. The primary database did not have heavy transactions during the availability group creation.

**Effects of availability group secondary copy**

After enabling the availability group, we replicated the production database transactions from the primary replica to the secondary replica. In this solution, we configured the secondary replica to be in synchronous-commit mode with read-only data access.

The secondary copy in synchronous-commit mode protected against data loss. In the test environment, with the same configuration, the production database (primary replica) IOPS decreased approximately 5 percent after enabling the availability group. The total user transactions per second from the OLTP workload decreased 7.5 percent. SQL Server database transactions decreased only 5 percent. The decrease would have been caused by any synchronous replication mechanism. The transactions needed to traverse the availability group link and become durable on the secondary replicas before the transactions were completed. This is relatively negligible in TPC-E testing. The storage array in this case performed the same as without availability group.

Table 8 and Figure 20 show more details.

<table>
<thead>
<tr>
<th>Table 8. Performance data comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tests</strong></td>
</tr>
<tr>
<td>Baseline</td>
</tr>
<tr>
<td>Baseline with availability group</td>
</tr>
</tbody>
</table>

1 TempDB transactions were included.
Impact of readable secondary replica

A readable secondary replica can offload the various types of read-only workload transactions (OLAP, SSRS, backup, and so on) from primary databases to the read-only copies at the secondary site. This can dedicate the primary databases in the availability group to mission-critical OLTP workloads.

This section describes test results of scenarios 3, 4, and 5 in Test scenarios. It compares effects of the readable secondary replica in these scenarios.

In the solution environment, when the read-only OLAP workload was allocated to the primary site, the total transactions per second from the OLTP workload at the primary site decreased approximately 30 percent while CPU usage increased 67 percent.

When we directed the workload to the secondary readable database copy, the primary site performance counters returned to their original levels. Compared with the read-only OLAP workload, the total transactions per second from the workloads for both OLTP and OLAP increased 36 percent.

The secondary site could handle many more read-only OLAP workload transactions than the primary site without any impact on the primary databases. In testing, the secondary site handled approximately 55 percent more transactions than the primary site handled with the same workload.

Running an OLAP workload on the primary site increased the I/O latency. In testing, the write latency on the primary site increased from 9 ms to 13 ms after we directed the OLAP workload to the primary site. Table 9 provides more details.
Table 9. Performance data with OLTP load

<table>
<thead>
<tr>
<th>Performance</th>
<th>OLTP load only</th>
<th>OLAP load on the primary site</th>
<th>OLAP load on the secondary site</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CPU</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary</td>
<td>45%</td>
<td>75%</td>
<td>54%</td>
</tr>
<tr>
<td>Secondary</td>
<td>3.3%</td>
<td>16.4%</td>
<td>35%</td>
</tr>
<tr>
<td><strong>Client TPS</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary</td>
<td>2,335 (OLTP)</td>
<td>2,360 (1,676 from OLTP and 684 from OLAP)</td>
<td>2,157 (OLTP)</td>
</tr>
<tr>
<td>Secondary</td>
<td>0</td>
<td>0</td>
<td>1,059</td>
</tr>
<tr>
<td><strong>SQL Server TPS</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary</td>
<td>9,322</td>
<td>13,500</td>
<td>12,500</td>
</tr>
<tr>
<td>Secondary</td>
<td>583</td>
<td>300</td>
<td>4,600</td>
</tr>
<tr>
<td><strong>IOPS</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary</td>
<td>23,800</td>
<td>25,300</td>
<td>23,300</td>
</tr>
<tr>
<td>Secondary</td>
<td>3,290</td>
<td>2,500</td>
<td>7,000</td>
</tr>
<tr>
<td><strong>Latency (ms) (read/write/transfer)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary</td>
<td>5/9/5</td>
<td>5/13/6</td>
<td>5/9/5</td>
</tr>
<tr>
<td>Secondary</td>
<td>5/2/3</td>
<td>4/1/2</td>
<td>5/1/5</td>
</tr>
</tbody>
</table>

Figure 21 compares the performance of the OLTP workloads, including the I/O latency on the primary site and the total IOPS on both sites.

Figure 22 shows the CPU usage and transactions per second (TPS) with an OLTP workload of the primary and secondary sites.
Figure 22.  CPU usage and TPS comparison for OLTP load

The CPU usage on the primary site decreased significantly when running a read-only load on the secondary site. Meanwhile, the CPU usage increased and was used more efficiently on the secondary site with the read-only load running on that site. Therefore, 24 percent more TPS were handled.

Figure 23 shows the client TPS comparison for OLTP and OLAP loads.

Figure 23.  Client TPS comparison for OLTP and OLAP loads
The TPS for OLTP decreased when adding OLAP load to the primary site. The OLAP load had little performance impact on the primary site after moving OLAP loads to the secondary site. The TPS for OLAP increased tremendously when directing OLAP loads to the secondary site.

**Reeseeding**

In test scenario 2 of Test scenarios, we backed up databases and logs locally, and then we restored the databases to a remote secondary server. We wrote each database’s backup to four files and ran the four database backups concurrently. All the backups and restores completed within 6 hours. The log backup and restore for each database completed within 1 minute. Table 10 shows the database reseeding time.

<table>
<thead>
<tr>
<th>Database</th>
<th>Database backup time (hh:mm:ss)</th>
<th>Database restore time</th>
<th>Total time</th>
</tr>
</thead>
<tbody>
<tr>
<td>TPCE1</td>
<td>0:08:12</td>
<td>0:55:04</td>
<td>1:03:16</td>
</tr>
<tr>
<td>TPCE2</td>
<td>0:25:50</td>
<td>2:17:32</td>
<td>2:43:22</td>
</tr>
<tr>
<td>TPCE3</td>
<td>0:43:51</td>
<td>3:21:42</td>
<td>4:05:33</td>
</tr>
<tr>
<td>TPCE4</td>
<td>1:14:43</td>
<td>4:44:59</td>
<td>5:59:42</td>
</tr>
</tbody>
</table>

**Overview**

From a storage perspective, the performance of the array and thin pools was excellent. Test scenario 5 in Test scenarios had the highest workload, in which the OLTP workload was on the primary replica and the read-only workload was on the secondary replica. The graphs in this section show the performance behavior of the highest workload.

After FAST VP stabilized the data movement, the disk pools reached less than 62 percent maximum utilization. The average utilization of disk pools was approximately 58 percent. The Flash tier utilization was less than 60 percent during this period. The SATA tier utilization was approximately 25 percent. Overall, the system can handle more IOPS and peaks as designed.

Figure 24 shows that both front-end and back-end CPU utilization of the Symmetrix VMAX 10K storage system was below 50 percent. Symmetrix VMAX 10K can handle the workload without any system stress.
Figure 24. VMAX 10K maximum utilization

Figure 25 shows sixteen ports on the Symmetrix VMAX 10K. With well-designed multipath load balancing enabled by PowerPath/VE, I/O on the ESXi servers hosting the primary and secondary virtual machines can be evenly distributed to the Symmetrix VMAX 10K front-end ports.

Figure 25. Storage tier utilization

With a well-configured back-end disk layout, I/O can be evenly distributed to both Symmetrix VMAX 10K front-end and back-end ports, as shown in Figure 26 and Figure 27.
Table 11 and Figure 28 show a performance comparison of the disk pool utilization, front-end and back-end CPU utilization, system IOPS for the baseline, and the read-only load on the primary and secondary replicas.

Table 11 indicates the following:

- CPU utilization had no pressure. The system had additional room to serve more workload.
- Back-end CPU utilization was more than twice as much as front-end CPU utilization.
• The additional back-end I/O was absorbed by the cache of the two-engine Symmetrix VMAX 10K.

Table 11. Symmetrix VMAX 10K performance

<table>
<thead>
<tr>
<th>Performance</th>
<th>Baseline</th>
<th>Availability group baseline</th>
<th>Baseline with read-only load on primary replica</th>
<th>Baseline with read-only load on secondary replica</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disk utilization</td>
<td>47.37%</td>
<td>53.54%</td>
<td>37.11%</td>
<td>44.14%</td>
</tr>
<tr>
<td>Front-end CPU utilization</td>
<td>23.01%</td>
<td>26.05%</td>
<td>23.26%</td>
<td>27.61%</td>
</tr>
<tr>
<td>Back-end CPU utilization</td>
<td>61.12%</td>
<td>67.89%</td>
<td>58.70%</td>
<td>67.92%</td>
</tr>
<tr>
<td>System IOPS</td>
<td>25,101</td>
<td>25,890</td>
<td>26,283</td>
<td>30,200</td>
</tr>
</tbody>
</table>

The same read-only load on the secondary SQL Server instance produced more IOPS while the CPU and disk utilization remained relatively flat as shown in Figure 28.

![Symmetrix VMAX 10K performance comparison](image)

From the VMware perspective, the ESXi servers, which hosted the SQL Server 2012 VMware hosts, consistently operated with sufficient performance. Test scenario 5 in Test scenarios had the highest workload, in which the OLTP workload was on the primary replica and the read-only workload was on the secondary replica.

Figure 29 shows the CPU usage of the ESXi server that served the primary copy of the SQL database virtual machine (SQL01). The CPU usage was approximately 60 percent during the testing.

VMware vSphere
ESXi server performance
Figure 29. CPU usage for the ESXi server serving the primary SQL virtual machine

Figure 30 shows the CPU usage of the ESXi server that served the secondary copy of the SQL database virtual machine (SQL02). The CPU usage was approximately 17 percent during the testing, which was much lower than for the primary copy.

Figure 30. CPU usage for the ESXi server serving the secondary SQL virtual machine
Conclusion

EMC Symmetrix VMAX 10K is proven to be highly scalable and reliable. It is capable of accommodating the storage performance and capacity requirements of SQL Server 2012. Symmetrix VMAX 10K can fully support a virtualized SQL Server 2012 environment with a readable secondary copy and the use of AlwaysOn Availability Groups.

- Symmetrix VMAX 10K can easily service more than 30,000 IOPS from SQL Server with an OLTP workload at the primary site while providing read-only access to near-live data through a readable secondary copy in SQL Server 2012 AlwaysOn Availability Groups.
- EMC FAST VP allows Symmetrix VMAX 10K storage arrays to maximize storage efficiency and provide autotiering of SQL Server database storage to the most suitable storage tier. It took approximately 4 hours for initial data movement to be stabilized. (Data sampling time and analysis time were configured to be 2 hours, respectively.)
- With Symmetrix VMAX 10K’s integration with VMware, the virtualized SQL Server 2012 environment is relatively simple to provision, manage, and monitor and ensures scalability and performance.

Findings

The key findings of testing of this proven solution are as follows:

- The Symmetrix VMAX 10K storage array, with its factory preconfigured settings, is simple and easy to incorporate into existing infrastructure. Built-in integration with VMware makes it a perfect fit for a virtualized environment, as demonstrated by this SQL Server 2012 solution.
- By moving an OLAP read-only workload from the primary replica to a secondary replica, the primary-site CPU and OLTP loads were significantly offloaded. More transactions per second and IOPS for both OLTP and OLAP loads were achieved.
- The process of creating availability groups requires enough storage to hold the compressed backups and completed within 6 hours.
- VMware paravirtual SCSI (PVSCSI) adapters are best suited for SAN, where hardware or applications drive a very high amount of I/O throughput. There is up to 25 percent improvement over legacy VMware virtual LSI Controllers.
- Storage performance can be improved by adjusting the maximum queue depth on HBA ports for the VMware vSphere ESXi server. In testing, there was more than a 30 percent improvement when the maximum queue depth for HBA ports were doubled from the default (32) to 64.
- Indirect checkpoints, a new feature in SQL Server 2012, provides a configurable database-level alternative to automatic checkpoints. In testing, I/O spiking was significantly reduced, especially on the secondary copy, where the checkpoint spikes were almost undetectable.
- Enabling availability groups in synchronous mode has very little impact on the OLTP load.
**References**

**White papers**
For additional information, refer to these white papers:

- *Provisioning EMC Symmetrix VMAX 10K Storage for VMware vSphere Environments*
- *Configuration Best Practices for Microsoft SQL Server and EMC Symmetrix VMAX 10K*
- *Implementing Fully Automated Storage Tiering for Virtual Pools (FAST VP) for EMC Symmetrix VMAX Series Arrays*
- *EMC Storage Optimization and High Availability for Microsoft SQL Server 2008 R2*

**Product documentation**
For additional information, refer to these product documents on the VMware website:

- *Changing the Queue Depth for QLogic and Emulex HBAs*
- *VMware ESX Scalable Storage Performance*

**Other documentation**
For additional information about Microsoft SQL Server 2012 and AlwaysOn, refer to these documents:

- *AlwaysOn Availability Groups*
- *Indirect Checkpoints*
- *Availability Group Listener*
- *Pre-Configuration Database Optimizations*
- *SQL Server Best Practices Article*
Appendix

AlwaysOn prerequisites

Before implementing AlwaysOn Availability Groups, always check the latest requirements. Some of the current prerequisite exit requirements are listed here.

Windows system

Windows system prerequisites are:

- All availability replicas for the given availability group must run on comparable systems that can handle identical workloads and have sufficient disk space for all databases in the availability group.
- The host computer must be a WSFC node.
- Microsoft Windows Enterprise is required because of the clustering features.
- Enable the .NET Framework 3.5.1 feature for all servers.

Note As primary databases grow, their corresponding secondary databases grow the same amount.

SQL Server 2012

SQL Server 2012 prerequisites are as follows:

- Each server instance must be running the Enterprise Edition of Microsoft SQL Server 2012.
- All server instances that host availability replicas for an availability group must use the same SQL Server collation.
- Enabling AlwaysOn Availability Groups requires membership in the administrator group on the local computer and full control on the WSFC cluster. AlwaysOn Availability Groups needs full control, and enabling AlwaysOn Availability Groups on an instance of SQL Server gives it full control of the WSFC cluster (through the Service SID).
- Enable the AlwaysOn Availability Groups feature on each server instance that will host an availability replica for any availability group.
- Each server instance requires a database mirroring endpoint. Note that this endpoint is shared by all availability replicas and database mirroring partners and witnesses on the server instance.
- Creating an availability group requires membership in the db_owner fixed database role and either CREATE AVAILABILITY GROUP server permission, CONTROL AVAILABILITY GROUP permission, ALTER ANY AVAILABILITY GROUP permission, or CONTROL SERVER permission.
To be eligible to be added to an availability group, a database must:

- Be a user database. System databases cannot belong to an availability group.
- Be a read/write database. You cannot add read-only databases to an availability group.
- Possess a full database backup and at least one log backup (unless you plan to use the New Availability Group wizard to perform full initial data synchronization).
- Be a multiuser database.

**Note** If you plan to use a SQL Server FCI to host an availability replica, ensure that you understand the FCI restrictions and that the FCI requirements are met.

**WSFC setup**

To set up WSFC, follow these steps:

1. Add the failover clustering role through the Add Features wizard.
2. Rename network connections to what they represent (public or private).
3. Specify the network connections advanced settings and place the public network at the start of the list.
4. Configure network card settings. For a private network, settings must contain only the IP and subnet mask; nodes must be able to communicate across the network.
5. Open the **Failover Cluster Manager** and click **Validate a Configuration**. Add the two servers in the cluster and click **Next** to continue.
6. In **Failover Cluster Manager**, click **Create a Cluster**.
7. Name the cluster and specify an IP address.