HITACHI
Inspire the Next

Microsoft GOLD CERTIFIED Partner

Virtualizing Microsoft® SQL Server 2008 Using VMware vSphere 4 on the Hitachi Adaptable Modular Storage 2000 Family

Best Practices Guide

By Eduardo Freitas and Bob Ingram

November 2009

Summary

Server virtualization is enjoying rapid adoption in the 21st-century data center as organizations continue to plan for, validate and realize the benefits this technology can provide. Maintaining acceptable user experience levels, meeting service level agreements and optimizing capital and operational costs are critical objectives for IT administrators. IT administrators often function in more than one role, charged with SQL Server, storage and server virtualization responsibilities. This means that businesses of all sizes need storage solutions for virtualized applications that are easy to deploy and maintain, are highly available, provide flexible scalability, deliver predictable performance, and introduce consolidation-related savings. The solution described in this paper addresses each of these issues for organizations that are challenged by managing underutilized, sprawling servers and storage silos that traditional SQL Server 2008 implementations can generate.

This white paper describes a building block architecture that helps large and enterprise deployments achieve these critical business objectives. It provides best practices required to successfully deploy Microsoft® SQL Server 2008 in a virtualized environment using VMware vSphere 4 configurations on the Hitachi Adaptable Modular Storage 2000 family.

For best results use Acrobat Reader 8.0.

Feedback

Hitachi Data Systems welcomes your feedback. Please share your thoughts by sending an e-mail message to SolutionLab@hds.com. Be sure to include the title of this white paper in your e-mail message

Table of Contents

Virtualizing SQL Server 2008 Using vSphere 4 on the Hitachi Adaptable Modular Storage 2000 Family

Best Practices Guide

Server virtualization is enjoying rapid adoption in the 21st-century data center as organizations continue to plan for, validate and realize the benefits this technology can provide. As organizations become more familiar, confident and capable with virtual machine deployments, they naturally investigate the feasibility of applying server virtualization to mission-critical applications in pursuit of similar or even enhanced benefits. Databases are arguably vital or nearly all organizations operating in today's global marketplace and Microsoft® SQL Server 2008 is the database server of choice for many of these enterprises.

Maintaining acceptable user experience levels, meeting service level agreements and optimizing capital and operational costs are critical objectives for IT administrators. IT administrators often function in more than one role, charged with SQL Server, storage and server virtualization responsibilities. This means that businesses of all sizes need storage solutions for virtualized applications that are easy to deploy and maintain, are highly available, provide flexible scalability, deliver predictable performance, and introduce consolidation-related savings. The solution described in this paper addresses each of these issues for organizations that are challenged by managing underutilized, sprawling servers and storage silos that traditional SQL Server 2008 implementations can generate.

This white paper describes a building block architecture that helps large and enterprise deployments achieve these critical business objectives. It provides best practices required to successfully deploy Microsoft® SQL Server 2008 in a virtualized environment using VMware vSphere 4 configurations on the Hitachi Adaptable Modular Storage 2000 family.

The solution is intended for use by IT administrators responsible for Microsoft SQL Server 2008, virtualization or storage. It assumes familiarity with vSphere 4 and vCenter Server, Hitachi Storage Navigator Modular 2 software, Windows 2008 and SQL Server 2008.

Although the documented solution in this paper used the Hitachi Adaptable Modular Storage 2300 for testing, any member of the Hitachi Adaptable Modular Storage 2000 family provides a reliable, flexible and cost-effective storage platform for supporting demanding applications like SQL Server 2008.

Solution Components

This section describes the key components of this solution.

Hitachi Adaptable Modular Storage 2000 Family

The Hitachi Adaptable Modular Storage 2000 family is the only midrange storage product with symmetric active-active controllers that provide integrated, automated hardware-based front-to-back-end I/O load balancing. Both controllers in a 2000 family storage system are able to dynamically and automatically assign the access paths from the back of the controller to the LU. All LUs are accessible regardless of the physical storage front-end port or the server from which the access is requested. Utilization rates of each controller are monitored so that a more even distribution of workload between the two controllers can be maintained. When coupled with VMware round-robin load balancing, the 2000 family eliminates many complex and time consuming path planning tasks that storage administrators typically face.

No other midrange storage product that scales beyond 100TB has a serial attached SCSI (SAS) drive interface. The point-to-point back end design virtually eliminates I/O transfer delays and contention associated with Fibre Channel arbitration and provides significantly higher bandwidth and I/O concurrency.

Although the Hitachi Adaptable Modular Storage 2300 was used in the testing of this solution, the information in this paper is relevant for and the building block approach can be applied to the other 2000 family members with the proper changes to account for capacity and performance differences. The 2000 family is an easy-to-use, scalable, cost effective storage system for mission-critical business applications like SQL Server 2008. It is also a top choice for tiered and standalone storage, consolidation, business continuity, data replication, backup and archiving. The Adaptable Modular Storage 2300 offers a rich set of features in a model that scales to 480 disk drives and delivers enterprise-class performance and capabilities at a modular price. Table 1 lists some of the 2300's specification options. For more information about the other member of the 2000 family, see the Hitachi Adaptable Modular Storage 2000 family Web site.

Table 1. Hitachi Adaptable Modular Storage 2300 Specification Options

Raw Capacity	236TB SATA 105TB SAS
Internal Disk Drives (SAS unless otherwise noted)	146GB (15K RPM) 300GB (15K RPM) 400GB (10K RPM) 450GB (15K RPM) 500GB SATA II (7200 RPM) 1TB SATA II (7200 RPM)
Disk Drive Interfaces	SAS and SATA
Host Interfaces	Fibre Channel: 8Gb/sec or 4Gb/sec iSCSI: GigE
Maximum Host Connections	8 Fibre Channel or 4 iSCSI
Maximum Attached Hosts Through Virtual Ports	2,048
SAS Links	16
Maximum Number of LUs	4,096
Maximum LU Size	60TB
Controller Cache (per system)	8GB to 16GB

Hitachi Dynamic Provisioning Software

On Hitachi Adaptable Modular Storage 2000 family systems, Hitachi Dynamic Provisioning software provides a dynamic provisioning feature with thin provisioning and wide striping components that provide virtual storage capacity to eliminate application service interruptions, reduce costs and simplify administration, as follows:

- Optimizes or "right-sizes" storage performance and capacity based on business or application requirements.
- Supports deferring storage capacity upgrades to align with actual business usage.
- Simplifies and adds agility to the storage administration process.
- Provides performance improvements through automatic optimized wide striping of data across all available disks in a storage pool.

For more information, see the Hitachi Dynamic Provisioning Software datasheet.

The wide-striping technology that is fundamental to Hitachi Data Provisioning software dramatically improves performance, capacity utilization and management of your environment. By deploying your SQL Server 2008 using V-VOLs from Hitachi Dynamic Provisioning storage pools on the 2000 family, you can expect the following benefits:

- An improved I/O buffer to burst into during peak usage times or intense maintenance activities like content indexing or database integrity checks
- A smoothing effect to the SQL Server workload that can eliminate hot spots across the different RAID groups, resulting in reduce data migration moves related to performance or capacity constraints
- Elimination of excess, unutilized capacity by leveraging the combined capabilities all disks comprising a storage pool
- Elimination of the need to manage the placement of heavy load databases

VMware vSphere 4

VMware vSphere 4 can help reduce hardware footprints and capital expenses dramatically through server consolidation. Utilizing VMware products and features such as ESX, vCenter Server, High Availability (HA), Distributed Resource Scheduler (DRS) and Fault Tolerance (FT), vSphere allows for a robust environment, centralized management and gives administrators control over key capabilities.

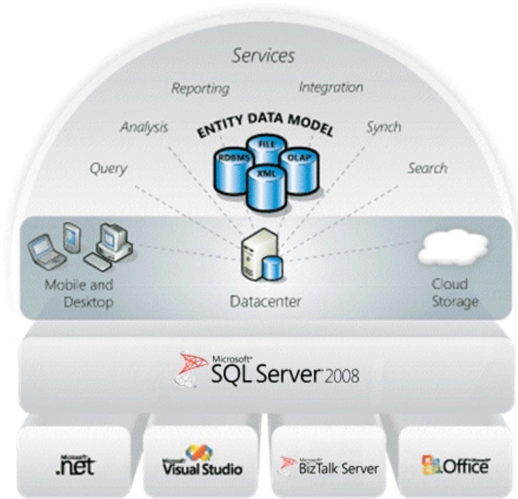
VMware provides product features that can help manage the entire infrastructure:

- **VMotion —** Allows for non-disruptive migration of both virtual machines and storage. Its performance graphs allow for monitoring of resources, virtual machines, resource pools and server utilization.
- **Distributed Resource Scheduler (DRS) —** Monitors resource utilization and intelligently allocates system resources as needed.
- **High Availability (HA) —** Monitors for hardware and OS failures and automatically restarts the virtual machine, providing cost effective failovers.
- **Fault Tolerance (FT) —** Provides continuous availability for applications by creating a live shadow instance of the virtual machine that stays in lockstep with the primary instance. If there is a hardware failure the shadow instance instantly takes over and eliminates even the smallest data loss.

SQL Server 2008

Microsoft SQL Server 2008 facilitates the management of any data, any place and any time. Together with the 2000 family, SQL Server 2008 provides a scalable, high-performance database engine for any midrange to enterprise level applications that require the highest levels of availability and security, while reducing the total cost of ownership through enhanced enterprise-class manageability for OLTP deployments at a midrange price. Figure 1 from Microsoft illustrates SQL Server 2008’s integration capabilities.

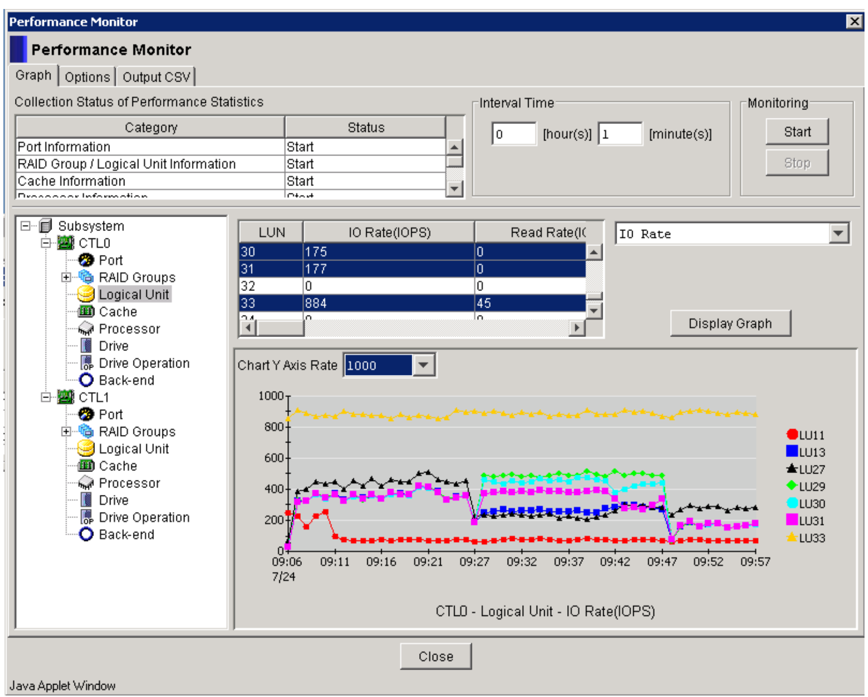
Figure 1. SQL Server 2008 Integration Capabilities



Microsoft Data Vision Platform diagram reprinted with permission from Microsoft Corporation.

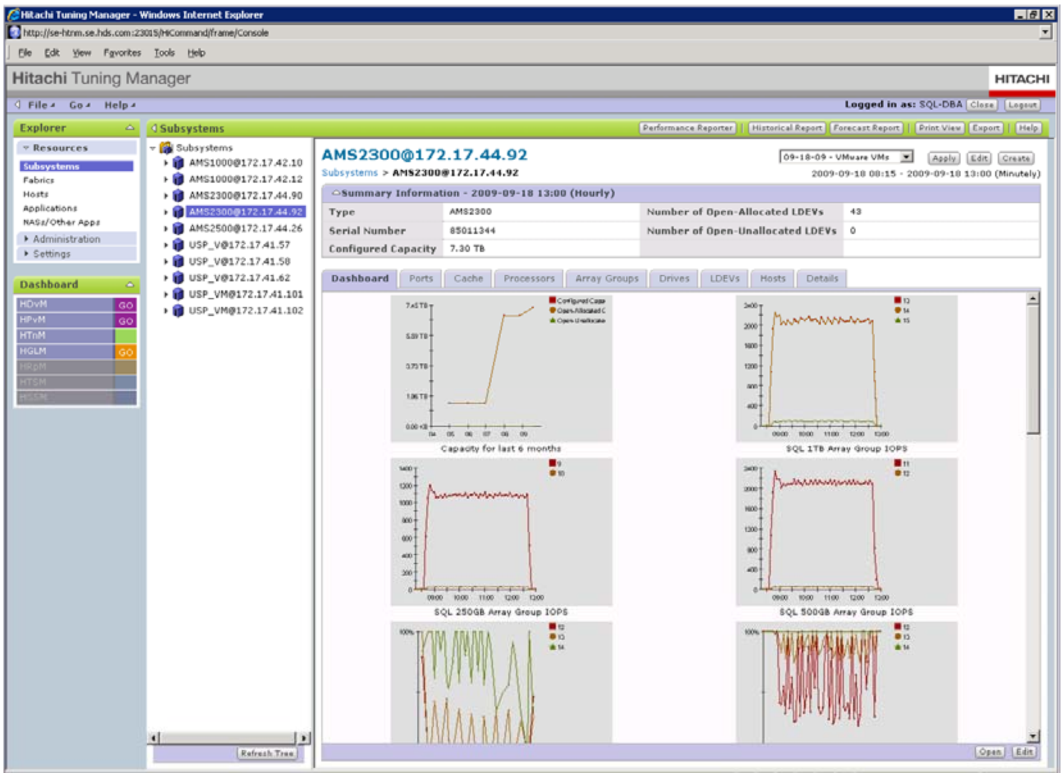
Hitachi Performance Monitor Feature

Hitachi Performance Monitor feature is part of the Storage Navigator Modular 2 software package. It acquires information on the performance of RAID groups, logical units (LUs) and other elements of the storage system while tracking the utilization rates of resources such as hard disk drives and processors. Information is displayed with line graphs in the Performance Monitor windows or saved in `.csv` files that you can analyze later. Use Performance Monitor to validate that the underlying storage system is not taxed by the SQL Server environment.



Hitachi Tuning Manager Software

Hitachi Tuning Manager software enables you to proactively monitor, manage and plan the performance and capacity for the Hitachi modular storage that is attached to your vSphere servers. Hitachi Tuning Manager software consolidates statistical performance data from the entire storage path. It collects performance and capacity data from the operating system, SQL Server instances, switch ports, and storage system information such as RAID groups and LUs IOPS and latency measures. Hitachi Tuning Manager provides the administrator a complete performance picture of the environment. It provides historical, current and forecast views of these metrics. For more information about Hitachi Tuning Manager, see the Hitachi Data Systems support portal.



Reference Architecture

For ease of management, scalability, and to provide predictable performance, this solution uses a building block approach, a familiar concept that Hitachi uses for current and previous generations of SQL Server and the Adaptable Modular Storage 2000 family. A single virtual machine (VM) running Windows Server 2008 and SQL Server 2008 with underlying storage from the Hitachi Adaptable Modular Storage 2300 make up a given building block for purposes of this paper. VMware recommends deploying VMs on shared storage like the 2000 family to take advantage of VMware features such as VMotion, HA and DRS.

LUs from the 2300 are allocated to the ESX hosts and formatted as virtual machine file system (VMFS) volumes from which virtual disks are created. The virtual disks (vDisks) are presented to the Windows Server 2008 guest operating system and can be partitioned and used in NTFS file systems for the SQL Server 2008 databases, `tempdb`, transaction logs, and possibly blob storage – although the later falls outside of this paper's scope. The solution building block supports databases starting at 250GB capacity and scale up to 1TB database capacity. Nevertheless, this is not a limitation of the building block scalability as more resources can be assigned to a given environment both from a storage as well as host perspective.

Table 2 describes an environment built using the building block architecture that supports three SQL Server instances. This reference architecture uses a 250GB database, a 500GB database and a 1TB database in a standard provisioned environment. The databases all use RAID-1+0 with different drive combinations depending on their capacity needs. The databases were also built by using a scalable approach from a virtual machine perspective.

Table 2. Reference Architecture Supporting Three Different Size Databases Using Standard LUs

Resource	Details for 250GB	Details for 500GB	Details for 1TB
Virtual machines	16GB memory 1 vCPU	32GB memory 2 vCPU	64GB memory 4 vCPU
Storage system - Hitachi Adaptable Modular Storage 2300	10 x 146GB 15K RPM SAS disks	16 x 146GB 15K RPM SAS disks	30 x 146GB 15K RPM SAS disks
RAID groups	1 x RAID-1+0 3D+3D 1 x RAID-1+0 2D+2D	1 x RAID-1+0 6D+6D 1 x RAID-1+0 2D+2D	2 x RAID-1+0 6D+6D 1 x RAID-1+0 3D+3D
LUs	1 x 300GB on RAID-1+0 3D+3D for database files	1 x 625GB on RAID-1+0 6D+6D for database files	2 x 625GB one on each RAID-1+0 6D+6D for database files
	1 x 30GB on RAID-1+0 3D+3D for `tempdb` files	1 x 62.5GB on RAID-1+0 6D+6D for `tempdb` files	2 x 62.5GB one on each RAID-1+0 3D+3D for `tempdb` files
	1 x 80GB on RAID-1+0 2D+2D for transaction logs*	1 x 160GB on RAID-1+0 2D+2D for transaction logs	1 x 250GB on RAID-1+0 3D+3D for transaction logs

* Although a 2D+2D RAID group was assigned for the transaction logs, the RAID group capacity might not be fully utilized until the user grows the existing database by another 250GB or adds a database. At that time an additional LU can be provisioned out of the RAID group and used for an additional transaction log file. For example, note how the 500GB architecture uses 160GB of the 2D+2D RAID group for transaction logs.

Additional storage was also provisioned to serve as the OS LUs for each of the SQL Server hosts. For that purpose, each SQL Server virtual server was provisioned on a 100GB LU from a common RAID-5 4+1 group.

When building the same type of environment while utilizing the Hitachi Dynamic Provisioning software, a single pool of 14 RAID-1+0 2+2 groups was created for the SQL Server database, transaction logs and `tempdb` volumes. Dedicated volumes within the pool simplify backup operations due to their granularity and also provide an easier way to monitor the capacity utilization of each of the volume types. Hitachi recommends using a RAID-1+0 2D+2D configuration as the building block for your Dynamic Provisioning pool as it provides good reliability, performance, as well as a good capacity increment for your pool.

Table 3 describes the same configuration described in Table 2 when deployed on a Hitachi Dynamic Provisioning software environment. This configuration supports the same database capacity describes on table 1 while providing the best performance by stripping the LUs across all disks in the single Dynamic Provisioning pool. Wide stripping reduces or eliminates hot spots caused by skewed workloads while ensuring the same level of RAID protection. In cases where log LUs are less busy than database LUs, I/Os are distributed across all disks and therefore more IOPS capability is available for a given SQL Server instance. RAID protection is provided to the Dynamic Provisioning pool through the use of multiple RAID-1+0 2D+2D groups.

Table 3. Reference Architecture Supporting Three Different Size Databases Using Hitachi Dynamic Provisioning

Resource	Details for 250GB	Details for 500GB	Details for 1TB
Virtual machines	1 16GB memory 1 vCPU	1 32GB memory 2 vCPU	1 64GB memory 4 vCPU
Storage system - Hitachi Adaptable Modular Storage 2300	11 x RAID-1+0 2D+2D, 44 x 146 15K RPM SAS disks for the database and `tempdb` Dynamic Provisioning pool 3 x RAID-1+0 2D+2D, 12 x 146 15K RPM SAS disks for the transaction logs Dynamic Provisioning pool		
LUs	1 x 300GB DP-VOL for database files 1 x 30GB DP-VOL for `tempdb` files 1 x 80GB DP-VOL for transaction logs	1 x 625GB DP-VOL for database files 1 x 62.5GB DP-VOL for `tempdb` files 1 x 160GB DP-VOL for transaction logs	1 x 1.2TB DP-VOL for database files 1 x 125GB DP-VOL for `tempdb` files 1 x 250GB DP-VOL for transaction logs

Notice that the same number of disks was utilized for both the standard (Table 2) as well as the Dynamic Provisioning configuration (Table 3. Also notice that the Dynamic Provisioning pool was created utilizing RAID-1+0 2+2 groups. While it is possible to create a pool with other RAID group types supported by the 2000 family, the RAID-1+0 2+2 provides the required protection level that Microsoft recommends for a SQL Server environment and a way to provide a small scalability increment for the pool. When increasing the amount of SQL Server virtual machines to an environment Hitachi Data Systems recommends that a separate pool be created.

The ESX hosts used in this test environment contain 128GB of memory and four quad-core AMD Opteron CPUs. Two ESX hosts are used at a minimum to achieve high availability. Additional SQL Server VMs can be added to one ESX host along with additional RAID groups on a 2000 family storage system to scale up to larger databases. Depending on the usage patterns, more VMs can be added to a single ESX host. To scale beyond the amount of resources a given ESX host can provide to an environment, additional ESX hosts must be added, ensuring that sufficient ESX hosts are employed to achieve redundancy or high availability.

This paper includes details regarding the test environment and methods used to validate the SQL Server 2008 instances and storage infrastructure. For more information about how to deploy and configure the environment, see the *Virtualizing SQL Server 2008 Using VMware vSphere 4 on the Hitachi Adaptable Modular Storage 2000 Family Deployment Guide*.

This solution focuses specifically on best practices for deploying SQL Server OLTP databases on the 2000 family. However, you can deploy other types of databases in this environment after conducting testing to ensure all components are provisioned appropriately to ensure performance.

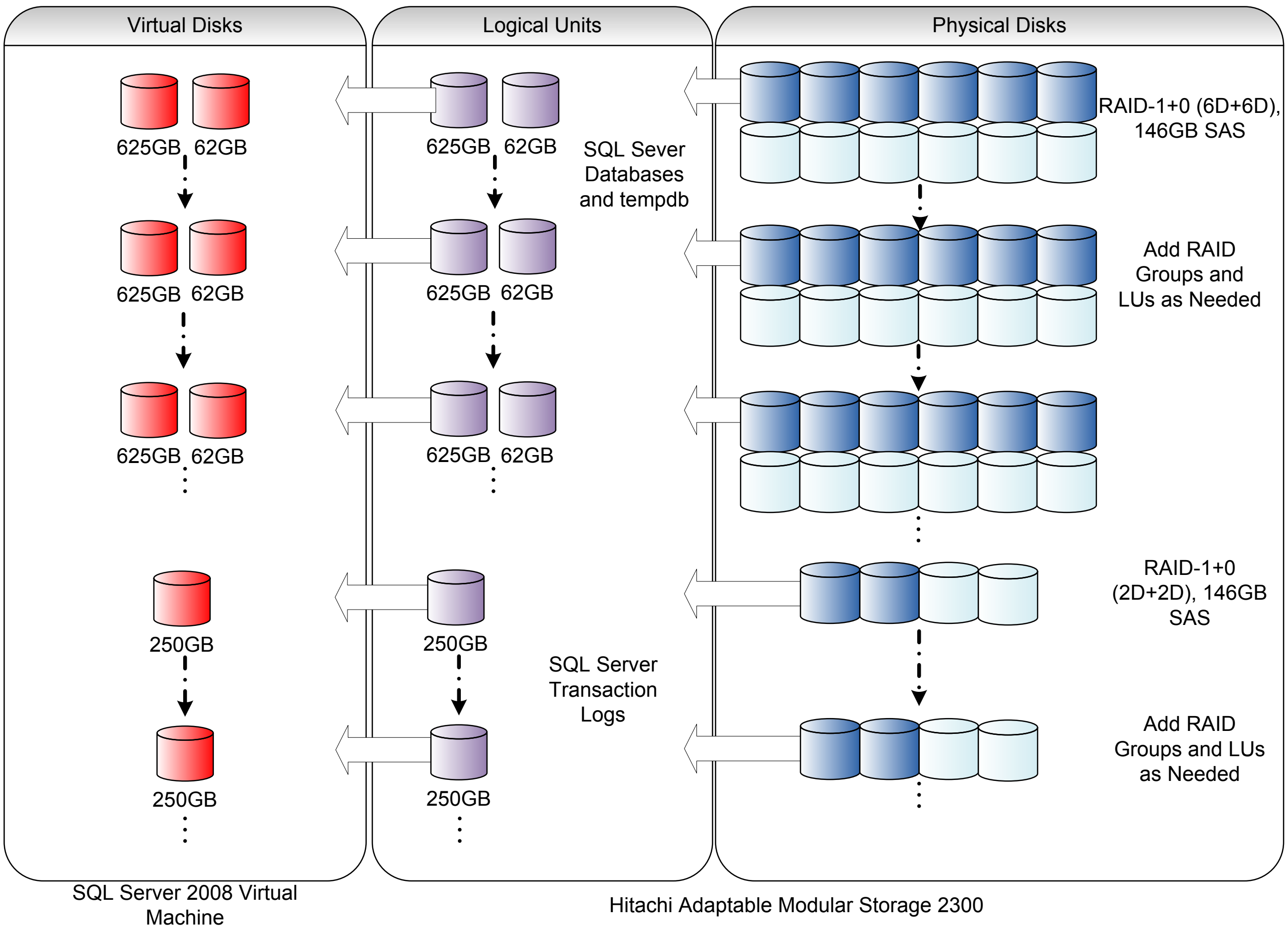
Storage Building Block

From a storage perspective, designing a SQL Server implementation on a virtual server environment is no different than performing the same activities on a non-virtualized environment. However, from a server perspective, you are required to perform a few additional tasks when deploying a virtual SQL Server environment compared to a non-virtualized SQL Server environment. Nevertheless, after the virtualized environment is set up, it is easy to manage using from VMware's vCenter. Deploying SQL Server 2008 using a building block approach allows you to easily manage and scale your environment. Additional VMs are easy to deploy using templates, and storage can be provisioned on a per RAID group basis. This paper provides a building block approach that can be utilized depending on your capacity needed for your environment. While the three reference architectures differ in capacity, they all use the same underlying building block. All the RAID groups and LUs for the environment were built by using the following guidelines:

- For standard provisioned environments, place database and log files on physically separate RAID groups.
- For standard provisioned environments tempdb files can be placed on the same RAID group as the database files. However, place them on a separate LU within the RAID group.
- For the best performance and availability, use RAID-1+0 should be utilized for the RAID groups for your SQL environment. Let your performance requirements for your environment determine the best combination of drives for a given RAID group, followed by the specific capacity.
- From a capacity perspective, always account for approximately 20 percent additional overhead for the database, `tempdb`, and transaction logs LUs.
- Ensure that any additional capacity added to the environment follows the same RAID-1+0 drive combination. This ensures that performance requirements continue to be met as additional capacity is needed.
- SQL Server virtual machines CPU and memory requirements are appropriately scaled for the expected database workloads.

With these guidelines in mind, you can build a scalable architecture that meets your performance and capacity requirements and ensures that you have a scalable and highly available environment. Figure 2 shows how an architecture built on an Adaptable Modular Storage 2300 that starts with a 500GB capacity can scale up while following the building block guidelines.

Figure 2. Physical to Virtual Disk Mapping of the 500GB Building Block Architecture



LUs are presented to the VMware ESX host and made into a VMFS. Note that when deploying environments that utilize Microsoft Cluster Services, you must use Raw Device Mapping on physical compatibility mode (pRDM) when mapping your LUs to a given virtual machine. For more information on best practices and deployment guidelines see the Microsoft Cluster Service (MSCS) support for ESX article.

Note that an additional 20 percent of capacity is accounted for on each of the LUs as per Microsoft's best practice recommendations. Additional capacity is also provided to ensure that the architecture continues to meet performance requirements for the environment. Also note that as capacity increases, it might be necessary to increase the amount of vCPUs and memory for a SQL Server virtual machine to maintain the same performance levels achieved for your initial building block.

The Hitachi Dynamic Provisioning software architecture uses 14 RAID-1+0, 2D+2D in two separate Dynamic Provisioning pools. One pool is used for the database and `tempdb` files, the other for the transaction log files. Each pool has independent LUs for each of the file types for the maximum protection level. Although a single Dynamic Provisioning pool can be created for database and transaction log files, a dual Dynamic Provisioning pool architecture ensures that the same level of protection you have on a standard provisioned environment at the RAID group level is also provided on a Hitachi Dynamic Provisioning software environment.

The Hitachi Dynamic Provisioning building block is also built following the guidelines described in the previous section.

Storage for the SQL Server databases is configured first for performance and capacity requirements for at the medium to high performance levels that a given SQL Server VM can handle. This is no different than what is done when sizing SQL Server for physical server deployments. The architectures described in this section require the following calculations be derived from the requirements of an existing or planned SQL Server 2008 environment:

- Overall capacity needed for databases and transaction logs including planned growth for SQL Server instances
- IOPS needed for SQL Server databases and transaction logs
- SQL Server instance host CPU and memory requirements for expected database workloads

Design Goals

This solution's building block architecture achieves the following design goals:

- Reach a minimum of 50 percent CPU utilization on the provisioned SQL Server 2008 VMs.
- Optimize storage configuration on the 2000 family storage system for best I/O throughput and ease of management for both standard and Hitachi Dynamic Provisioning configurations.
- Deliver sustainable and acceptable levels of IOPS for the within 20ms and 1ms response time for database and transaction logs transactions respectively as tested, measured through the use of an OLTP workload environment simulating a stock brokerage scenario and reported by the Windows Performance Monitor as well as SQL Server Management Studio accommodating burst and peak activities. SQL Server Dynamic Management Views (DMVs) were used to determine if the SQL Server database met Microsoft's performance recommendations.
- Deliver at least 80 percent disk capacity utilization for the database volumes.

Several other factors contributed to the considerations and trade-offs to arrive at the standard provisioned building block architecture, including these:

- For a very heavy user profile, you might need additional disks available for use to support the required IOPS in standard provisioned environments. As with any SQL Server deployment, make sure to test your environment and ensure the proper amount of disks are provisioned to the database from an IOPS perspective.
- Granularity of scale is determined by the database size along with the SQL Server virtual server provisioned hardware; additional storage can be provisioned adding RAID-1+0 groups for database and transaction logs. The specific drive combination for each RAID group depends on the initial building block environment that meets both performance and capacity requirements for your environment.

  For example, in a 500GB database configuration, a RAID-1+0 6+6 group is used to host both the database as well as the `tempdb` LUs while a RAID-1+0 2+2 group is used to host the transaction logs LU. To increase the capacity of the database by 500GB in a standard provisioned environment, create an additional RAID-1+0 6+6 group and provision LUs to host the additional database files and possibly additional `tempdb` volumes, if the number of `tempdb` files are not aligned with the number of CPU cores due to the increase of resources made available to a SQL Server VM. In this case, an additional RAID group for the transaction logs is not necessary due to the unused capacity still available from the original RAID group provisioned for this purpose. An additional increase of 500GB to the database capacity requires two additional RAID groups, one RAID-1+0 6+6 group for the database volumes and one RAID-1+0 2+2 group for the transaction logs.

  In a dynamic provisioned configuration, Dynamic Provisioning pool space can be allocated by adding one RAID-1+0 group to the Dynamic Provisioning pool. After that, the required LUs can be created and appropriately assigned to the ESX host.

- Appropriately devise a database growth plan, and keep in mind that VMware VMs support a total of 60 vDisks; while smaller capacity increments might support more granular scaling, you are limited by the number of increments per VM because you can have only 60 vDisks. Remember that each increment includes one LU for the database files, one LU for the transaction logs, and one LU for tempdb.
- ESX hosts support a total limit of 256 LUs and all LUs must be masked to all ESX hosts in a VMware cluster scenario.

Multipathing

To maintain a constant connection between the ESX hosts and storage, ESX supports multipathing. Multipathing allows multiple physical or logical connections from the host to the storage. To support multipathing, Hitachi Data Systems recommends that the physical host contain at least two HBAs that connect to at least one Fibre Channel port on each storage system controller.

In ESX, several types of multipathing policies are available through the VMware Native Multipathing Plugin (NMP):

- **Fixed (Default) —** Uses the designated preferred path, if it is configured. Otherwise, it uses the first working path discovered at system boot time. If the host cannot use the preferred path, it selects a random alternative available path. The host automatically reverts to the preferred path as soon as that path becomes available. This is the default setting on ESX and requires manual load balancing and path distribution across SAN fabrics and storage controllers.
- **Round-robin (Recommended) —** Uses a path selection algorithm that rotates through all available paths, enabling load distribution across the paths.

  Round-robin is the best choice for the Adaptable Modular Storage 2000 family of storage systems due to their symmetric active-active controller design. It ensures all resources within the storage system are utilized while maintaining the path failover capability of the environment and simplifying the ease of setup as all LUs are mapped to all ports assigned to an ESX host.
- **Most recently used (MRU) —** Selects the path the ESX host used most recently to access the given device. If this path becomes unavailable, the host switches to a different path and continues to use the new path while it is available.

Disk Management

ESX hosts can access LUs in two ways, through VMFS or raw device mapping (RDM). Virtual disks are files stored on a datastore, which is a logical container for the `.vmdk` files. The datastores are deployed on storage devices and make up the VMFS. VMFS is optimized for VMs and hides the specifics of the underlying storage.

VMFS can be accessed by several ESX hosts and the cluster feature allows for distributed file locking for the VMs. The VMFS can be extended while the client is running and can extend across multiple LUs. Using Storage VMotion, the `.vmdk` file can be moved to another datastore nondisruptively.

As an alternative, LUs can be mapped as a raw device mapping (RDM). This feature allows a LU to be mapped directly to a VM. RDMs are useful for command devices and any other device that requires direct communication to the storage processor. Hitachi Data Systems recommends against using RDM for SQL Server environments because the enhanced performance and features of the VMFS are not available. However, keep in mind that when using Microsoft Cluster Services (MSCS) you must use RDM in physical compatibility mode.

For more information about VMFS and RDM, see VMware's Fibre Channel SAN Configuration Guide.

Scaling SQL Server 2008

This section describes planning and deployment considerations to keep in mind when scaling your SQL Server environment.

Planning

Scaling SQL Server databases and instances requires planning and testing both from a server as well as a storage perspective. When scaling up, both capacity and performance are concerns and any production environment must be properly tested to ensure it satisfies end-user requirements. From an ESX perspective, you must calculate the maximum number of VMs that a single ESX host can run by noting the total number of CPUs and total memory requirements. Best practice is to not allocate to the virtual machine more resources than exist on the physical ESX host. Keep in mind that the ESX service console requires one CPU of at least 400MB of memory.

Scaling the Storage

When starting with a database in which the performance and capacity requirements fit within the described 250GB reference architecture, RAID-1+0 3+3 groups can be added to the environment as a scaling step. The same can be done from the reference architectures for the 500GB and 1TB. From a database perspective, Hitachi Data Systems lab tests show that each of the reference architectures described on this paper appropriately escalates the amount of IOPS from one configuration to the next while maintaining latency values at or below Microsoft best practice levels.

You can mix various drive combinations and capacity values with this architecture as long as you follow the storage block building guidelines described in the “Storage Building Block” section of this paper.

In a dynamic provisioned environment, additional LUs can be provisioned and assigned to a given port as with the standard provisioned environment. In cases where the overall Dynamic Provisioning pool might be nearing its limit, Hitachi Data Systems recommends creating a second Dynamic Provisioning pool to ensure full utilization of the wide striping capabilities of Hitachi Dynamic Provisioning software. Note that transaction log LUs also follow the same growth factor. You typically need to account for the sum of the transaction log LUs’ capacity to be between five and 10 percent of the space provisioned for database LUs.

Table 4 shows the resources utilized for the building block architecture with a 500GB capacity.

Table 4. Building Block Resources for 500GB SQL Server

Resource	***Details***
Virtual machines	1 32GB memory 2 x vCPU
Storage systems	Hitachi Adaptable Modular Storage 2300 23 x 146GB 15K RPM SAS disks Note: 2 disks are used as hot spares
RAID groups	1 x RAID-1+0 (6D+6D) 1 x RAID-1+0 (2D+2D) 1 RAID-5 (4D+1P)

LUs	1 x 625GB on RAID-1+0 (6D+6D) for database files 1 x 62.5GB on RAID-1+0 (6D+6D) for `tempdb` files 1 x 160GB on RAID-1+0 (2D+2D) for transaction logs 1 150GB on RAID-5 (4D+1P) for VM OS*

* The remainder of the space within the RAID-5 (4D+1P) group was used for additional virtual machines that were hosted on the same ESX server.

Like any other SQL Server deployment, always test before deploying your production environment. Remember that additional adjustments for things such as unanticipated growth, protection methods and service level agreements might become necessary. Hitachi Data Systems used industry standard OLTP workloads that pushed the virtual servers to 60 percent or more of their “% Processor time” on Windows Performance Monitor while maintaining latency levels at or below Microsoft best practice levels.

For more information about monitoring your SQL Server environment and Adaptable Modular Storage 2000 family, see the Performance Monitoring section of the Hitachi Adaptable Modular Storage 2000 Family Best Practices with Microsoft® SQL Server For OLTP Applications white paper.

A second 2300 might be needed for more complex environments in which a failover capability is necessary to meet your business and regulatory requirements. VMware vCenter Site Recovery Manager (SRM) and Hitachi Storage Replication Adapter (SRA) allow you to failover SQL Servers to a secondary ESX host and Hitachi Adaptable Modular Storage 2300 system.

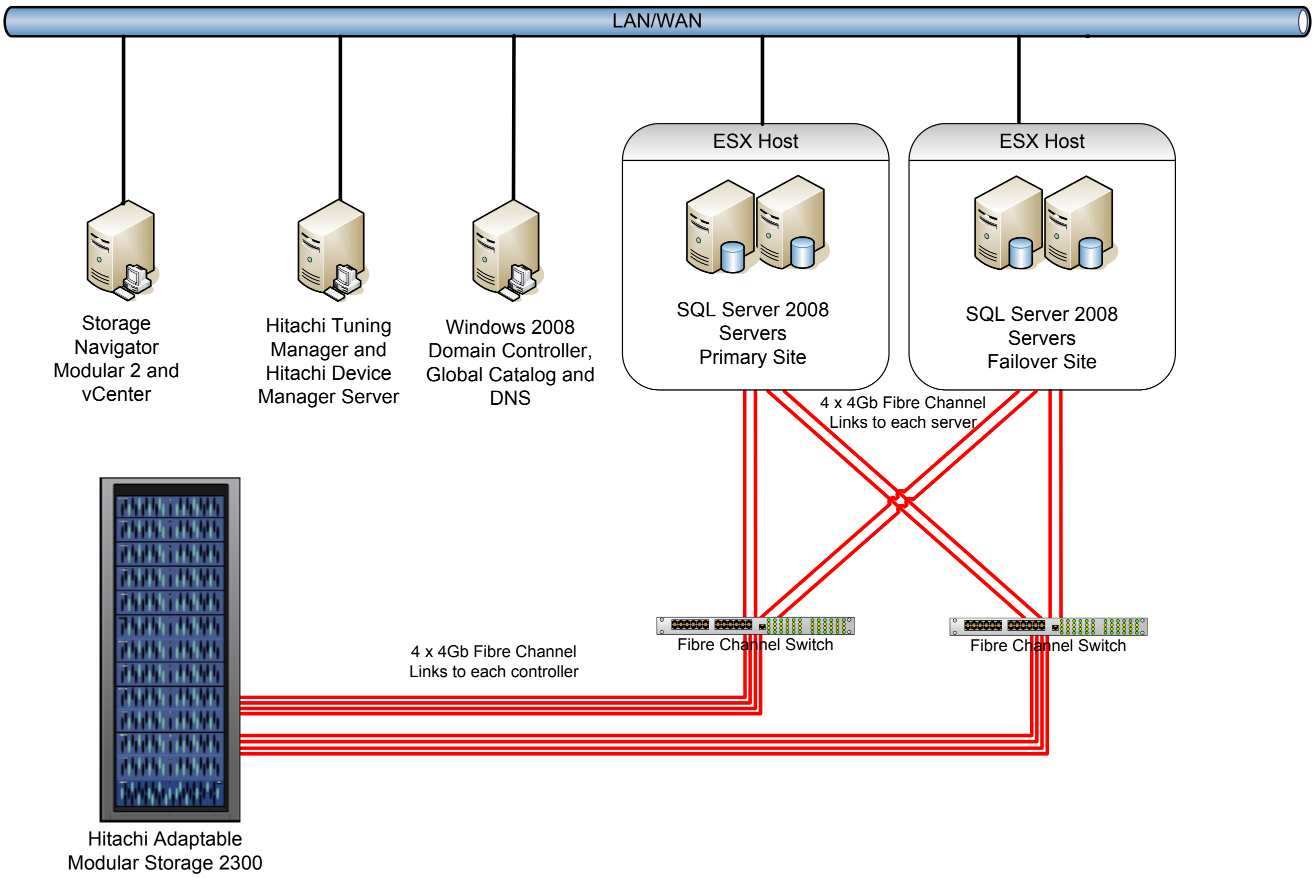
For more information, see the Hitachi Storage Replication Adapter Software - VMware vCenter Site Recovery Manager Deployment Guide.

VMware and Hitachi tools enable you to fully examine the resource utilization on your environment. On the vCenter **Performance** tab, you can examine the CPU and memory utilization of the ESX hosts and ensure the system is not over committed. On the storage side, Hitachi Tuning Manager provides a holistic view of all the Adaptable Modula Storage 2000 family performance related counters. This enables you to closely monitor the system to ensure that your environment is utilizing all available resources. You can also identify potential bottlenecks that might require you to add disks to the environment.

Tested Deployment

The following sections describe the hardware and software used in the deployment tested by Hitachi Data Systems. Figure 3 illustrates the test environment topology.

Figure 3. Test Environment Topology



Software

Table 5 describes the software required for the tested deployment described in this white paper.

Table 5. Software Resources

Software	Version
ESX	4.0.0, build 140815
vCenter	4.0 140742
Windows Server	2008 Enterprise edition, Service Pack 2
SQL Server 2008	2008 Enterprise edition, Service Pack 1
Hitachi Storage Navigator Modular 2	7.0
Hitachi Device Manager	6.2.0-00
Hitachi Tuning Manager	6.1.0-00
Hitachi Dynamic Provisioning License	Microcode dependent

In addition, for the tested deployment using Hitachi Dynamic Provisioning software, the license key was installed on the Hitachi Adaptable Modular 2300.

Hardware

The solution building block is designed to optimize resource utilization and throughput of the 2300 and the ESX host. Table 6 lists the hardware used in testing this architecture.

Table 6. Hardware Resources

Hardware	Configuration
Hitachi Adaptable Modular Storage 2300	Microcode 0870H 2 controllers 4 disk trays 15 146GB 15K RPM SAS disks per tray 16GB cache per controller
Two Brocade 51000 SAN switches	FOS 6.21* 16 4Gb Fibre Channel ports used
Two Dell R905 servers	BIOS firmware 4.0.3** 4 quad core AMD Opteron 1.9GHz processors 128GB memory 2 Emulex LPe11002 4Gb Fibre Channel HBAs

*Use the latest or the most recent Hitachi supported levels of microcode and FOS levels
**Use the latest available Hitachi and Microsoft supported BIOS firmware level.

Path Configuration

All LUs are masked and zoned to all four HBAs on each of the ESX hosts using four dedicated Fibre Channel ports on the 2300. Note that while four dedicated ports from the host and four from the storage system were utilized for each host to storage connection on the test, it is possible to use two ports from the host and two from the storage system without reaching the IOPS capability of a given port and therefore not creating a bottleneck in the environment.

Table 7 lists the connections between the ESX hosts and the storage system ports.

Table 7. Path Configuration

ESX Host	*Host HBA Number/ ESX Port Name*	*Switch*	*Director Zone Name*	*Hitachi Adaptable Modular 2500 Port*	*Hitachi Adaptable Modular Storage 2500 Host Group*
ESX 1	HBA 1 port 1 vmHBA 1	Brocade 5100-1	ESX_1_HBA1_AMS2300_0A	0A	ESX_primary_vmHBA_1
ESX 1	HBA 1 port 2 vmHBA 2	Brocade 5100-2	ESX_1_HBA1_AMS2300_1A	1A	ESX_primary_vmHBA_2
ESX 1	HBA 2 port 1 vmHBA 3	Brocade 5100-1	ESX_1_HBA3_AMS2300_0B	0B	ESX_primary_vmHBA_3
ESX 1	HBA 2 port 2 vmHBA 4	Brocade 5100-2	ESX_1_HBA4_AMS2300_1B	1B	ESX_primary_vmHBA_4
ESX 2	HBA 1 port 1 vmHBA 1	Brocade 5100-1	ESX_2_HBA1_AMS2300_0C	0C	ESX_secondary_vmHBA_1
ESX 2	HBA 1 port 2 vmHBA 2	Brocade 5100-2	ESX_2_HBA1_AMS2300_1C	1C	ESX_secondary_vmHBA_2
ESX 2	HBA 2 port 1 vmHBA 3	Brocade 5100-1	ESX_2_HBA3_AMS2300_0D	0D	ESX_secondary_vmHBA_3
ESX 2	HBA 2 port 2 vmHBA 4	Brocade 5100-2	ESX_2_HBA4_AMS2300_1D	1D	ESX_secondary_vmHBA_4

On each ESX host, the round robin multipath algorithm was used for all LUs presented to the hosts.

Storage Configuration Standard RAID Groups and LUs

This section describes how storage was configured for a single building block for the standard provisioned tested deployment.

Figure 4 shows the building block for the 500GB capacity environment.

Figure 4. Standard Provisioned 500GB Storage Building Block

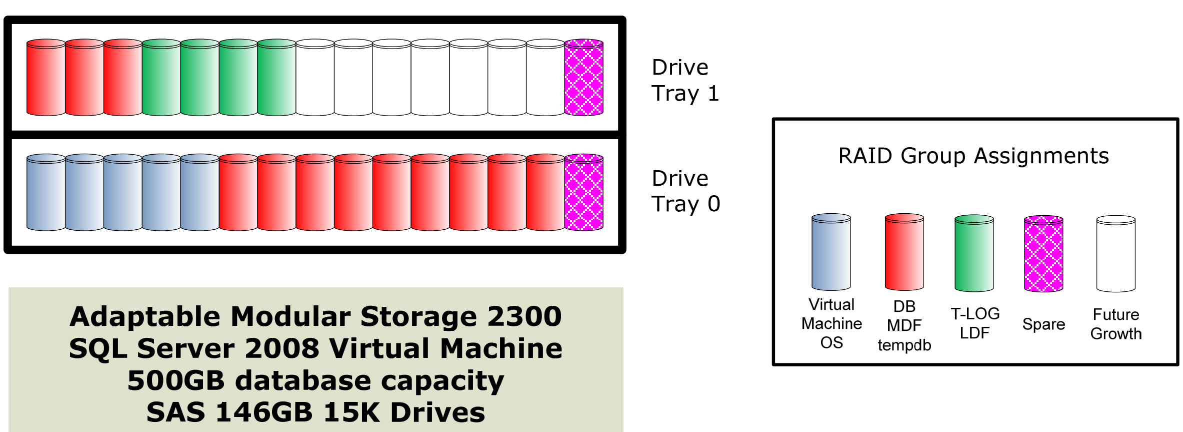


Table 8 shows the building block configuration for the 500GB environment.

Table 8. Detailed Storage Configuration for a Single 500GB Building Block

RAID Group	LUN	Size (GB)	RAID Level	RAID Type	Disk Spec	Description
1	1	100	RAID-5	4+1	146GB 15K	OS LU
2	1	625	RAID-1+0	6+6	146GB 15K	Database LU
2	2	62.5	RAID-1+0	6+6	146GB 15K	`tempdb` LU
3	3	160	RAID-1+0	2+2	146GB 15K	Transaction logs LU

Storage Configuration Using Hitachi Dynamic Provisioning Software

This virtualized SQL Server solution uses one Hitachi Adaptable Modular Storage 2300 with the Hitachi Dynamic Provisioning software configuration for three different size databases provisioned to three different SQL Server VMs. Each VM was also configured in a building block approach. The RAID group that hosts the VM's OS resides on a separate RAID-5 4+1 groups.

Using Hitachi Dynamic Provisioning software, the I/O workload is distributed across all drives in the Dynamic Provisioning pool, which consists of 14 RAID-1+0 2+2 groups. All LUs for the building block using Hitachi Dynamic Provisioning software are allocated from two Dynamic Provisioning pools, one for database and `tempdb` LUs and another for the transaction logs LUs.

Table 9 lists the configuration of RAID groups and LUs for a single building block using Hitachi Dynamic Provisioning software. Note that all storage is in Dynamic Provisioning pool 0.

Table 9. Detailed Storage Configuration for a Single Building Block using Hitachi Dynamic Provisioning Software

LU	Size	Description
4	300GB	Database LU for SQL Server 1
5	30GB	`tempdb` for SQL Server 1
6	80GB	Transaction logs for SQL Server 1
7	625GB	Database for SQL Server 2
8	62.5GB	`tempdb` for SQL Server 2
9	160GB	Transaction logs for SQL Server 2
10	1.2TB	Database for SQL Server 3
11	125GB	`tempdb` for SQL Server 3
12	250GB	Transaction logs for SQL Server 3

Hitachi Data Systems recommends that you create a Dynamic Provisioning pool for each building block.

Test Methodology

An industry standard OLTP workload that simulates a stock brokerage scenario was exercised on each of the building block architectures that and then additional SQL Server VMs and storage building blocks were added to scale up the configuration. Results were collected and the number of building blocks increased until the design goal was achieved. The results show that as each architecture was scaled up, the IOPS levels at a minimum doubled at each increment while maintaining the latency levels at or below Microsoft best practice recommendations.

Table 10 lists the OLTP workload parameters used as a baseline for the tests.

Table 10. OLTP Workload Test Parameters

Parameter	*Description*
Test scenario	OLTP workload for a stock brokerage firm
Virtual machine CPU utilization	≥50%
Minimum database LU capacity usage	80%
Minimum individual disk busy rate	50%
Number of active users	Variable, scaling with architecture size
Test type	Performance
Test duration	≥3 hours

Table 11 describes the success criteria established by Microsoft's best practice recommendations and monitored using Windows Performance Monitor.

Table 11. OLTP Workload Success Criteria and Test Results

Description	*Criteria*	*Results*
Average achieved database IOPS 250GB architecture 500GB architecture 1TB architecture	Varies	 1083IOPS 2100IOPS 4177IOPS
Database Avg. Disk sec/read Disk sec/Write	≤ 20ms ≤ 20ms	Average 10-20ms Maximum 20ms
Transaction Log Avg. Disk sec/Read Avg. Disk sec/Write	≤ 5ms ≤ 5ms	Average 1-5ms Maximum 5ms

Testing shows that this solution meets or exceeds all design goals for the building architectures.

Best Practices

These best practices for the design of the virtualized SQL Server 2008 environment are based on Hitachi Data Systems testing of the building block architecture. For a deployment of SQL Server 2008 using VMware vSphere 4 and one of the Adaptable Modular Storage 2000 family members, follow these guidelines:

- For enhanced storage utilization and usability, use Hitachi Dynamic Provisioning software.
- When using standard LUs, use RAID-1+0 6D+6D for SQL Server database LUs and RAID-1+0 2D+2D for SQL Server transaction logs LUs for the a 500GB building block
- When using dynamically provisioned LUs:
  - Use RAID-1+0 for all pools.
  - Adjust the size of the RAID group for scalability. When using 146GB disks, 2D+2D RAID groups provide the best scaling option for SQL Server 2008 environments.
  - Create separate pools for database and transaction log LUs.
- Use a minimum of two vSphere servers to take full advantage of vSphere features such as High Availability, Dynamic Resource Allocation and Fault Tolerance.
- Configure at least two redundant paths to the Adaptable Modular Storage 2000 from each ESX host.
- Use the building block architecture for easy scaling.
- Use round robin multipathing policy to take best advantage of the Adaptable Modular Storage 2000's symmetric active-active front end ports.
- Use virtual disks for disk management.

Conclusion

The virtualization of a SQL Server environment can be easily implemented with solutions and features provided by Hitachi Data Systems and VMware. Although a virtualized environment requires a few additional tasks during setup compared to a non-virtual environment, this small investment of time easily pays off with reduced data center server sprawl and full environment utilization. This solution offers a number of other benefits, including increased availability, reduced complexity, simplified management and scalability, and reduced power consumption. The building block architecture described in this document optimizes memory, storage and CPU utilization for your SQL Server environment while ensuring a reliable and easily managed environment.

For more information about Hitachi products and solutions, see the Hitachi Data Systems Web site, your sales representative or a channel partner.

Hitachi Data Systems Corporation

Corporate Headquarters 750 Central Expressway, Santa Clara, California 95050-2627 USA
Contact Information: + 1 408 970 1000 www.hds.com / info@hds.com

Asia Pacific and Americas 750 Central Expressway, Santa Clara, California 95050-2627 USA
Contact Information: + 1 408 970 1000 www.hds.com / info@hds.com

Europe Headquarters Sefton Park, Stoke Poges, Buckinghamshire SL2 4HD United Kingdom
Contact Information: + 44 (0) 1753 618000 www.hds.com / info.uk@hds.com

AS-022-01 March 2010