HITACHI
Inspire the Next

Deploying Microsoft SQL Server 2008 R2 with Logical Partitioning on the Hitachi Virtual Storage Platform with Hitachi Dynamic Tiering

Reference Architecture Guide

By Eduardo Freitas

May 2011

Microsoft GOLD CERTIFIED Partner

Feedback

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

Table of Contents

Deploying Microsoft SQL Server 2008 R2 with Logical Partitioning on the Hitachi Virtual Storage Platform with Hitachi Dynamic Tiering

Reference Architecture Guide

All business sizes need to deploy and maintain storage solutions for virtualized applications. These solutions include high availability and flexible scalability, with predictable performance delivery. These businesses want to improve resource utilization with reduced server sprawl. This creates green data centers with smaller footprints and lower energy costs.

When the virtualized application is a mission-critical Microsoft® SQL Server 2008 deployment, the stakes are even higher. This solution reduces the complexity of SQL server environments by using servers and storage from Hitachi that work together seamlessly.

The Hitachi Virtual Storage Platform helps you leverage your information, the new currency in today's data-driven economy. Information must be protected and be readily accessible to ensure the survival and success of your business.

The Virtual Storage Platform maximizes cost efficiency and return on investment by creating an agile storage infrastructure. This reduces costs and increases performance, availability, scalability, and reliability.

The Hitachi Compute Blade 2000 is an enterprise-class platform that offers these features:

- Balanced system architecture that eliminates bottlenecks in performance and throughput
- Embedded logical partitioning virtualization
- Unprecedented configuration flexibility
- Eco-friendly power-saving features and capabilities
- Fast recovery from server failures due to N+1 cold standby design that allows you to replace failed servers within minutes instead of hours or days

With its unique combination of power, efficiency, and flexibility, the Hitachi Compute Blade 2000 extends the benefits of virtualization to new areas of the enterprise data center — including mission-critical application servers and database servers like Microsoft SQL Server 2008 R2 — with minimal cost and maximum simplicity.

Many SQL server instances deployed in data centers operate on dedicated server hardware. This can lead to server sprawl and underutilized hardware resources. Each dedicated server requires the finite resources of management, physical space within the data center, power, and cooling.

Each database may require different levels of service agreements. This makes it difficult to consolidate multiple databases or SQL server instances to save costs while maintaining acceptable latency values.

Deploying Microsoft SQL Server 2008 R2 using Hitachi Virtual Storage Platform and Microsoft Hyper-V helps resolve the problem of ever-increasing demands on finite resources through consolidation and virtualization.

This reference architecture consolidates the number of dedicated servers running SQL server instances by virtualizing Microsoft SQL Server 2008 R2, with logical partitioning virtualization, and Hitachi Virtual Storage Platform as virtual machines running on a Hitachi Compute Blade 2000 server. Hitachi Dynamic Tiering allows you to combine these different workloads on a single pool of disks for greater flexibility.

This reference architecture guide describes a building block design that helps large and enterprise deployments achieve easy-to-manage, highly available, and scalable virtualized SQL server deployments. Hitachi Dynamic Tiering improves resource utilization, reduces costs, and substantially minimizes the database latency and maximizes the amount of capacity available for use. It provides the best practices required for a successful deployment of Microsoft SQL Server 2008 R2 in a virtualized environment using Microsoft Hyper-V, Hitachi Dynamic Tiering, and the Hitachi Virtual Storage Platform.

This reference architecture guide is written for IT administrators responsible for Microsoft SQL Server 2008 deployments, virtualization, or storage administration. It assumes familiarity with general storage concepts, Microsoft Hyper-V, Microsoft Windows Server 2008, and Microsoft SQL Server 2008.

Solution Overview

This solution provides a building block architecture that helps large and enterprise-scale deployments of Microsoft SQL Server 2008 R2 achieve their critical business objectives. It provides best practices required to deploy Microsoft SQL Server 2008 R2 in a virtualized environment using Hitachi Compute Blade 2000 logical partition virtualization configurations with the Hitachi Virtual Storage Platform.

This solution describes a reference architecture for deploying a SQL server environment that can scale from 75,000 user accounts to 150,000 user accounts. However, this number is not a limitation on the scaling of the environment from a storage or a server perspective. Hitachi Data Systems testing used an industry-standard OLTP workload that simulates a stock brokerage scenario to populate the databases and exercise the building block architecture.

To achieve high availability for this solution, redundant physical paths were enabled using multiple host bus adapters (HBAs) on the server and multiple paths to the Hitachi Virtual Storage Platform. Proper zoning within the storage fabric and the use of multipathing software allows for continued operation in the event of a hardware component failure.

Figure 1 shows the topology of this reference architecture.

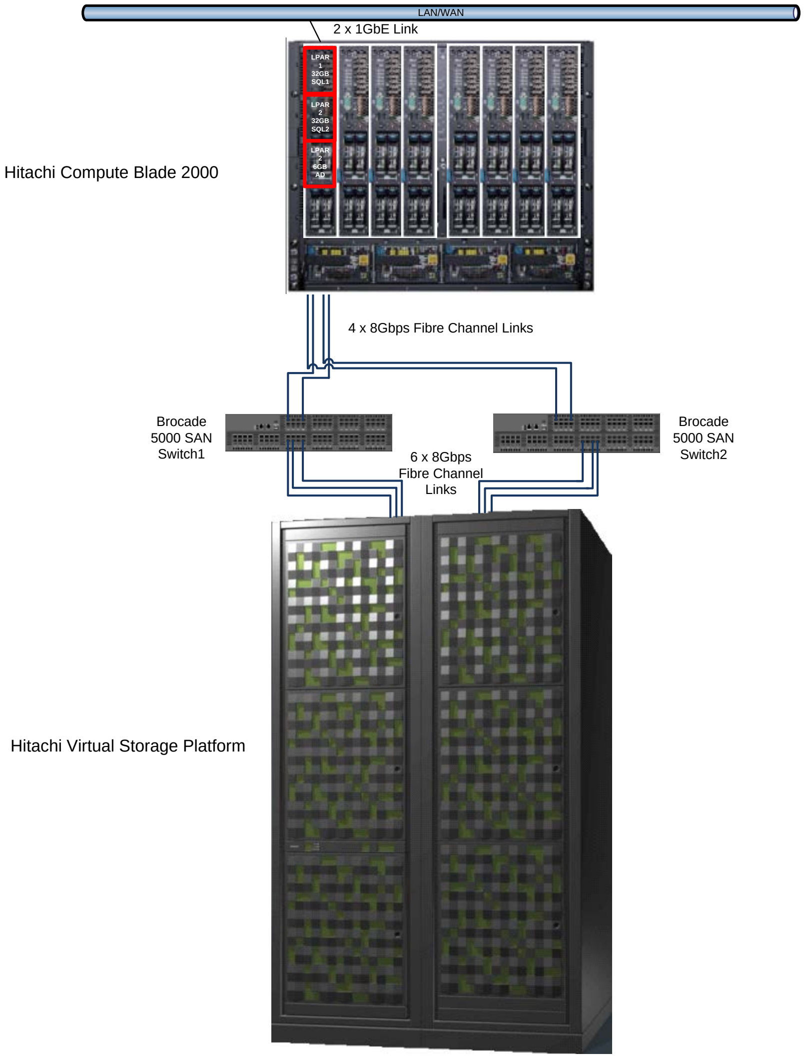

Figure 1

Key Solution Components

These are the key components needed to deploy this solution.

Hitachi Virtual Storage Platform

The Hitachi Virtual Storage Platform is the industry's only 3D scaling storage platform. It has the ability to do the following in a single storage system:

- **Scale Up—**Increase virtual server consolidation to improve utilization of resources and reduce costs.
- **Scale Out—**Combine multiple chassis into a single logical system with shared resources.
- **Scale Deep—**Extend the advanced functions of the Virtual Storage Platform to external multivendor storage.

The Hitachi Virtual Storage Platform flexibly adapts for performance, capacity, connectivity, and virtualization. No other enterprise storage platform can dynamically scale in three dimensions.

Find out more information about the Hitachi Virtual Storage Platform.

Hitachi Dynamic Tiering

Hitachi Dynamic Tiering is a revolutionary solution that eliminates the time-consuming manual processes of data classification and movement between storage tiers. This optimizes tiered storage usage while improving performance.

Most data is accessed rarely after creation. As a result, it should not be stored on your most expensive storage tier. Instead, it can be moved to a lower, less expensive storage tier. Defining where and for how long data should reside at any point in its life cycle can be complex and problematic.

Many organizations manually provision space using several storage technologies with different performance and cost characteristics. With this approach, data specialists typically examine past usage patterns to determine how to configure tiers. This makes the storage infrastructure unable to respond effectively to dynamic application and data use. If usage patterns change rapidly, manually tiered storage systems produce less than optimal results.

Hitachi Dynamic Tiering takes the automation of tiered storage to a new level. It enables the management of multiple storage tiers as a single entity. By leveraging the existing features of Hitachi Dynamic Provisioning, Hitachi Dynamic Tiering presents a new kind of virtual volume with embedded smart tiering. It monitors access and moves data at the 42MB page level.

Hitachi Dynamic Provisioning breaks the volume into pages and Hitachi Dynamic Tiering automatically moves infrequently referenced pages to lower cost tiers of storage. Moving pages instead of entire data sets or files reduces the time and storage space required to migrate data.

Hitachi Dynamic Tiering automatically moves pages of data within virtual volumes configured on a dynamic provisioning pool to the most appropriate media, according to workload. After an initial setup process, Hitachi Dynamic Tiering monitors data access in real time and makes decisions on moving data between the available storage tiers based on actual use. If a page on a lower tier is accessed frequently, Hitachi Dynamic Tiering moves it to a higher tier.

This maximizes service levels while it minimizes total cost of storage ownership. Additionally, Hitachi Dynamic Tiering improves the availability and performance of your storage systems and the applications using that storage.

Previously, each dynamic provisioning pool had to be created using one RAID level and one disk type. Hitachi Dynamic Tiering on the Virtual Storage Platform allows a single pool to contain tiers made up of multiple types of RAID groups and any type of disk. Hitachi Dynamic Tiering manages the various tiers within a dynamic provisioning pool automatically. This eliminates most user management of storage tiers within a storage system, and maintains peak performance under dynamic conditions without storage administrator intervention.

Additionally, Hitachi Dynamic Tiering inherits the thin provisioning and wide striping functionalities of Hitachi Dynamic Provisioning. This provides virtual storage capacity to eliminate application service interruptions, reduce costs, and simplify administration by:

- Optimizing or “right-size” storage performance and capacity based on business or application requirements
- Supporting deferring storage capacity upgrades to align with actual business usage.
- Simplifying the storage administration process.
- Providing performance improvements through automatic optimized wide striping of data across all available disks in a storage pool.
- Eliminating hot spots across different RAID groups by smoothing the combined workload.
- Improving capacity utilization significantly.

For more information, see the Hitachi Dynamic Tiering datasheet.

Hitachi Compute Blade 2000

Figure 2 shows the Hitachi Compute Blade 2000. It features a modular architecture that delivers unprecedented configuration flexibility.

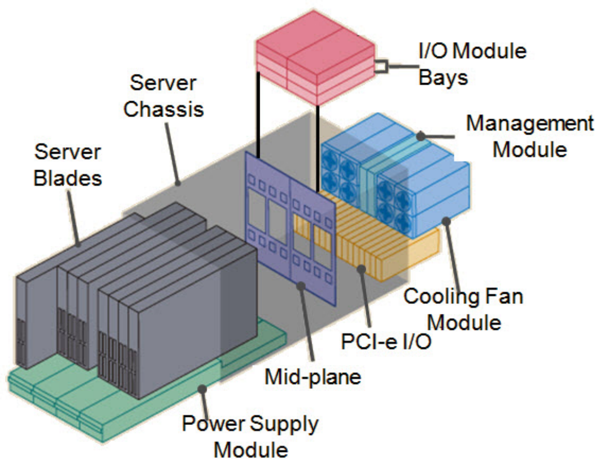

Figure 2

The Hitachi Compute Blade 2000 combines all the benefits of virtualization with all the advantages of the blade server format:

- Simplicity
- Flexibility
- High compute density
- Power efficiency

Take advantage of these benefits with the Hitachi Compute Blade 2000:

- Consolidate more resources.
- Extend the virtualization benefits, whether logical partitioning, VMware vSphere, Microsoft Hyper-V, or all three.
- Reduce costs without sacrificing performance.

The Hitachi Compute Blade 2000 enables virtualization of servers to consolidate application and database servers for backbone systems. These are areas where effective consolidation was difficult in the past. By removing performance and I/O bottlenecks, Hitachi Compute Blade 2000 opens new opportunities for increasing efficiency and utilization rates and reduces the administrative burden in your data center.

The Hitachi Compute Blade 2000 is a manageable and flexible blade system. Configure and administer the Hitachi Compute Blade 2000 in one of two ways:

- Use an HTML browser that supports secure encrypted communications
- Leverage the optional management suite to manage multiple chassis using a unified GUI-based interface.

Chassis

The Hitachi Compute Blade 2000 chassis is a 19-inch rack compatible, 10U-high chassis with a high degree of configuration flexibility. The front of the chassis has slots for eight server blades and four power supply modules. The back of the chassis has six bays for I/O switch modules, eight fan modules, two management modules, 16 half-height PCIe slots, and two AC power input modules.

The four power supply modules can be configured as mirrored power supplies. This provides backup on each side of the chassis and higher reliability.

Efficient, variable-speed, redundant fan modules, each with three fans, provides cooling. This tolerates fan failures within a module. If an entire module fails, the other fan modules continue to cool the chassis.

Server Blades

The Hitachi Compute Blade 2000 supports two server blade options that can be combined within the same chassis. Table 1 lists the specifications for each option.

Table 1. Hitachi Compute Blade 2000 Server Blade Specifications

Feature	***X55A2***	***X57A1***
Processors (up to two per blade)	Intel Xeon 5600—4 or 6 cores	Intel Xeon 7500—6 or 8 core
Processor cores	4,6,8 or 12	6, 8, 12 or 16
Memory slots	18	32
Maximum memory	144GB (with 8GB DIMMS)	256GB (with 8GB DIMMS)
Hard drives	Up to 4	N/A
Network interface cards (on-board)	Up to 2 1Gb Ethernet	Up to 2 1Gb Ethernet

Feature	*X55A2*	*X57A1*
Other interfaces	2 USB 2.0 ports and 1 serial port	2 USB 2.0 ports and 1 serial port
Mezzanine slots	2	2
PCIe 2.0 (8x) expansion slots	2	2

Up to four X57A1 blades can be connected using the SMP interface connector to create a single eight-socket SMP system with up to 64 cores and 1024GB of memory.

I/O Options

The connections from the server blades through the chassis' mid-plane to the bays or slots on the back of the chassis consist of the following:

- The two on-board NICs connect to switch bays one and two
- The optional mezzanine card in the first mezzanine slot connects to switch bays three and four
- The optional mezzanine card in the second mezzanine slot connects to switch bays five and six
- Two connections to dedicated PCIe slots

For more information about currently supported mezzanine cards, switch modules and PCIe cards, see the Hitachi Compute Blade 2000 datasheet.

The I/O options supported by the optional mezzanine cards and the switch modules are either 1Gb/sec Ethernet or 8Gb/sec Fibre Channel connectivity.

Logical Partitioning

The Hitachi Compute Blade 2000 logical partitioning feature is embedded in the firmware of the Hitachi Compute Blade 2000 server blades. This is a proven, mainframe-class technology that combines the logical partitioning expertise of Hitachi with Intel VT technologies to improve performance, reliability, and security.

Unlike emulation solutions, the embedded logical partitioning virtualization feature does not degrade application performance. Unlike third-party virtualization solutions, it does not require the purchase and installation of additional components, keeping total cost of ownership low.

A blade can operate in one of two modes, basic or Hitachi Virtualization Manager (HVM). There are three licensing options available:

- **Basic**—The blade operates as a standard server without logical partitioning support.
- **HVM with essential license**—The blade supports 2 logical partitions. No additional purchase is required for this mode.
- **HVM with enterprise license**—The blade supports up to 16 logical partitions. The enterprise license costs additional.

You can use the embedded logical partitioning feature alone or combine it with Microsoft Hyper-V, VMware vSphere, or both in a single system. This provides additional flexibility.

Hitachi Blade Server Management Modules

The Hitachi Compute Blade 2000 supports up to two management modules for redundancy. Each module is hot-swappable and supports live firmware updates without the need for shutting down the blades. Each module supports an independent management LAN interface from the data network for remote and secure management of the chassis and all blades. Each module supports a serial command line interface and a web interface. SNMP and email alerts are supported.

N+1 or N+M Cold Standby Failover

The Hitachi Compute Blade 2000 maintains high uptime levels through sophisticated failover mechanisms. The N+1 cold standby function enables multiple servers to share a standby server. This increases system availability while decreasing the need for multiple standby servers or costly software-based high-availability servers. It enables the system to detect a fault in a server blade and switch to the standby server, manually or automatically. The hardware switching is executed in the absence of the administrator, enabling the system to return to normal operations within a short time.

The N+M cold standby function has “M” backup server blades for every “N” active server blades for cascading failover. In the event of multiple hardware failures, the system automatically detects the fault, identifying the problem by indicating the faulty server blade. This allows immediate failure recovery. This approach can reduce total downtime by enabling sharing of the application workload among the working servers.

Microsoft SQL Server 2008 R2

Microsoft SQL Server 2008 R2 provides a scalable, high performance database engine for mission-critical applications that require the highest levels of availability and security. SQL Server 2008 R2 also provides enhanced enterprise-class manageability for large OLTP deployments like the one described in this reference architecture guide.

With the Hitachi Virtual Storage Platform, Microsoft SQL Server 2008 R2 provides a scalable, high-performance database engine for any midrange to enterprise level application.

Hitachi Dynamic Link Manager Advanced

Hitachi Dynamic Link Manager Advanced is a software package that consists of Hitachi Dynamic Link Manager multipathing software and Hitachi Global Link Manager software. Hitachi Dynamic Link Manager, installed on each SQL server host, includes capabilities such as path failover and failback, and automatic load balancing to provide higher data availability and accessibility.

Hitachi Dynamic Link Manager was used for storage area network (SAN) multipathing, configured with the round-robin multipathing policy. The round-robin load balancing algorithm automatically selects a path by rotating through all available paths. This balances the load across all available paths and optimizing IOPS and response time.

Hitachi Global Link Manager provides a centralized view and management interface for the SAN multipathing configuration, as show, in Figure 3.

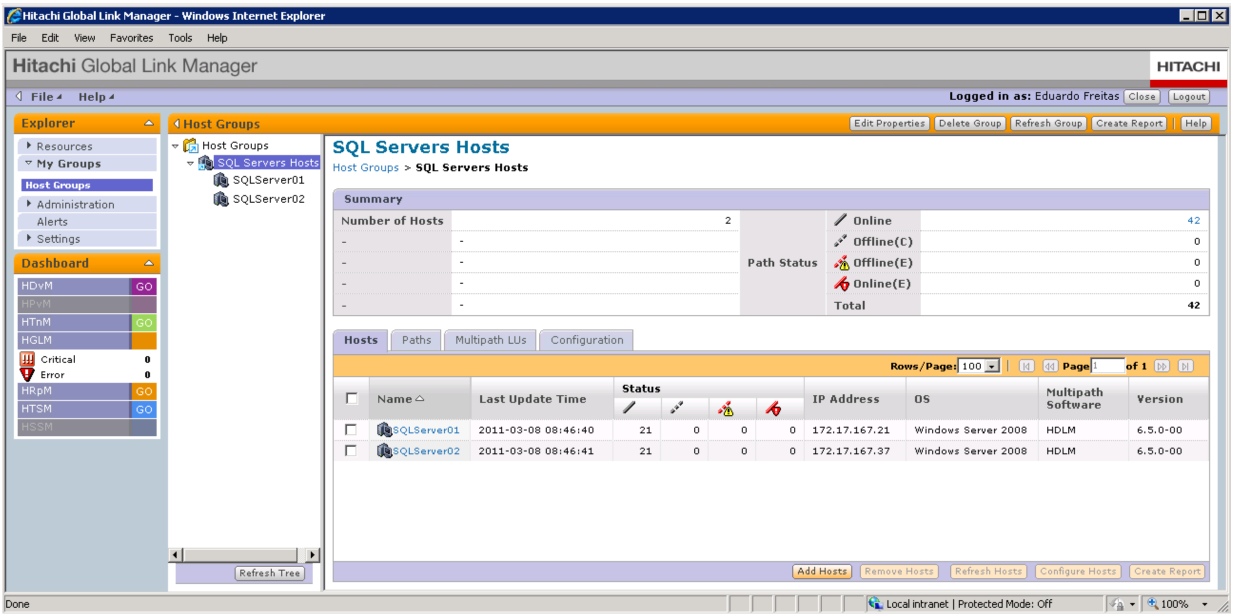

Figure 3

Solution Design

This section provides detailed information on the large-scale Microsoft SQL Server 2008 R2 environment design used for this reference architecture. Although it includes application and hardware design information, it focuses primarily on the storage design required to build the basic infrastructure for the SQL environment.

High-level Infrastructure

For ease of management, scalability, and predictable performance, this solution uses a building block approach. This reference architecture takes advantages of the logical partitioning feature of the Hitachi Compute Blade 2000 with the capability of using the N+1 cold standby function to establish a reliable and highly available virtualized SQL server solution.

Each building block includes a single logical partition provisioned on a server blade using:

- Microsoft Windows Server 2008 R2
- Microsoft SQL Server 2008 R2
- Underlying storage from the Hitachi Virtual Storage Platform

This reference architecture scales from a 75,000-user deployment on one 900GB database with a single SQL server instance on a dedicated logical partition to a 150,000-user deployment on two 900GB databases. The scaling of the reference architecture takes place by implementing a second logical partition with a SQL server instance. The use of Hitachi Dynamic Tiering enables a substantial decrease in I/O latency as well as better utilization of disk capacity by using RAID-5 instead of RAID-1+0.

The building block for the 75,000-user database consists of:

- Two 450GB LUNs for the SQL database data files
- One 90GB LUN for `tempdb`
- One 225GB LUN for the SQL transaction log

This storage building block was replicated along with adding a second logical partition on the Hitachi Compute Blade 2000 and a second SQL server instance to support 150,000 users. All LUNs were provisioned from dynamic tiering pools on the Hitachi Virtual Storage Platform. Figure 4 shows the high level storage design of the Microsoft SQL Server 2008 R2 reference architecture on the Hitachi Virtual Storage Platform.

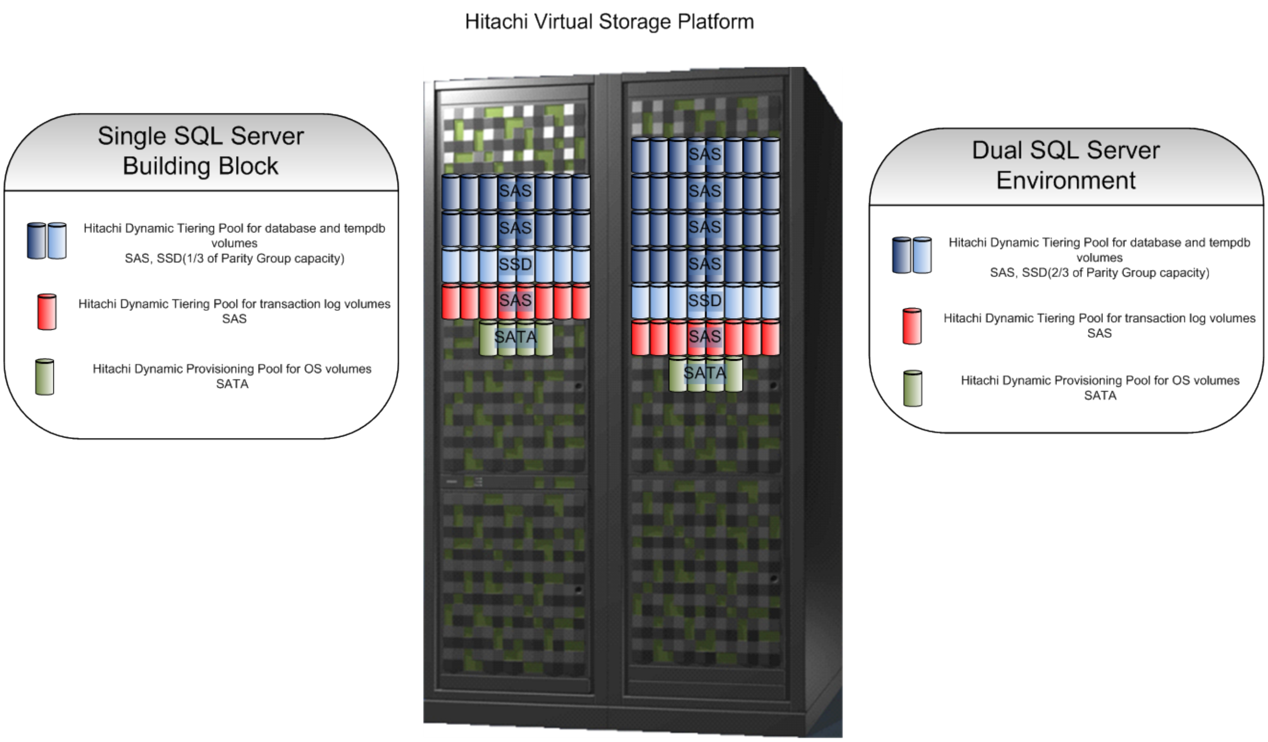

Figure 4

LUNs from the Hitachi Virtual Storage Platform are allocated to the servers using host groups configured on the Virtual Storage Platform. These LUNs are available to the Windows Server 2008 R2 operating system and are used for:

- Microsoft SQL Server 2008 R2 databases
- `tempdb`, transaction logs
- BLOB storage (outside of this paper's scope)

The solution building block supports databases starting at 900GB capacity. It scales up to 1.8TB database capacity across multiple SQL server instances. Nevertheless, this is not a limitation of the building block approach, as more resources can be assigned to a given environment from the storage and host perspectives.

This reference architecture includes a pool shared within the infrastructure to host the OS boot volumes for the virtual machines.

For this reference architecture, Hitachi Data Systems connected the Hitachi Compute Blade 2000 server and the Hitachi Virtual Storage Platform to two Fibre Channel switches for redundancy and high availability purposes. Another option is to deploy the environment using an enterprise-level director that contains multiple blades that can support high availability and redundancy.

Table 2 describes the storage, storage area network (SAN), and server components used in this reference architecture.

Table 2. Hardware Components

Hardware	Detail Description	Version	Quantity
Hitachi Virtual Storage Platform	128GB cache memory, 6 × 8Gb/sec front- end ports, 35 × 300GB, 10K RPM, SAS disks 8 × 300GB, SSD disks 4 × 2TB 7.2K RPM, SATA disks	Microcode 70-00-63	1
Hitachi Compute Blade 2000 Chassis	4 × 8Gb/sec dual-port HBAs 2 × 1Gb/sec LAN Switch Module 2 × management modules 8 × cooling fan modules 4 × power supply modules	A0154-E-5234	1
Hitachi Compute Blade 2000 - X57A1 Blade	2 × 8-Core Intel Xeon X7560 2.26GHz 128GB memory	58.22	1
Brocade 5340 switch	SAN switch with 8Gb/sec Fibre Channel ports	FOS 6.4.0b	2

Hitachi Compute Blade 2000 Blade Configuration

This reference architecture uses:

- One SMP capable blade (X57A1)
- Two 1Gb/sec LAN switch modules
- Two Hitachi dual port 8Gb/sec HBA PCIe cards

The blade has two on-board NICs that are internally connected to the LAN switch modules. There are also two PCIe slots available.

Support of the logical partitioning feature requires Hitachi HBAs. With Hitachi HBAs, enabling logical partitioning creates eight virtual WWNs for the logical partitions on each server blade.

Figure 5 shows the front and back view of the Hitachi Compute Blade 2000 used in this solution. Present in the chassis for this solution are unused blades and HBAs.

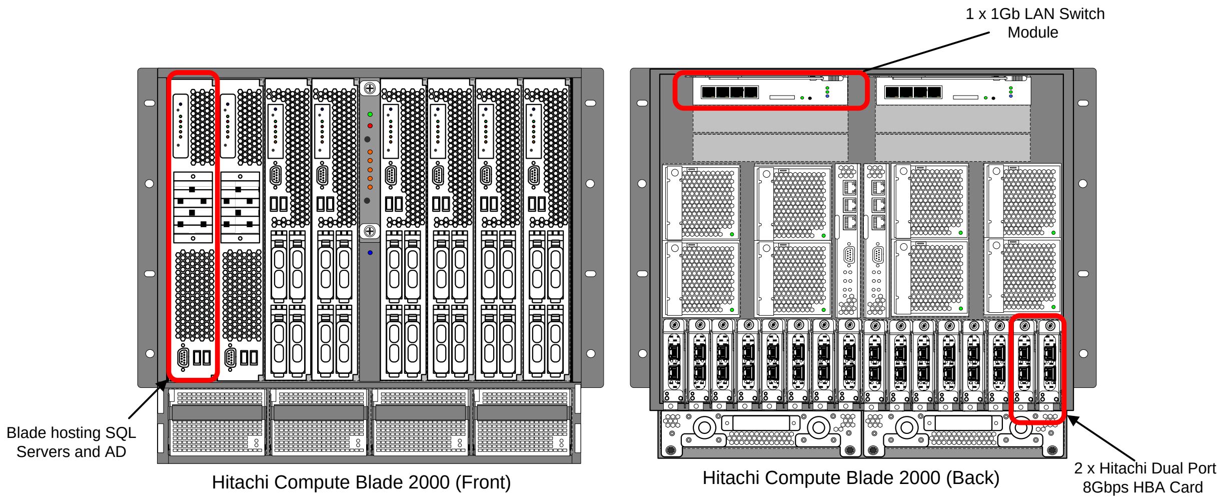

Figure 5

All modules, including fans and power supplies, can be configured redundantly and hot swapped. This maximizes system uptime.

To meet the 150,000 user requirement and to use the hardware resources efficiently, the Microsoft Windows servers listed in Table 3 were deployed on one physical server blade.

Table 3. Hitachi Compute Blade 2000 Configuration

Server Blade	*Logical Partition*	*Server Name*	*Role*	*Number of CPU Cores*	*Memory (GB)*
Blade 1	LPAR1	SQLSERVER1	SQL server instance	4	32
	LPAR2	SQLSERVER2	SQL server instance	4	32
	LPAR3	AD2	Active Directory DNS	4	8

Software

Table 4 describes the software used in this reference architecture.

Table 4. Software Components

Software	*Version*
Hitachi Storage Navigator	Dependent on microcode version
Hitachi Dynamic Tiering	Dependent on microcode version
Hitachi Dynamic Link Manager	6.5.0-00
Hitachi Global Link Manager	6.5.0-00
Microsoft Windows Server 2008 R2 (for SQL Server and Domain Controller hosts)	Enterprise edition
Microsoft SQL Server 2008 R2	Enterprise edition

Multipathing

Multipathing software, such as Hitachi Dynamic Link Manager or Microsoft Windows Server 2008 native multipath IO (MPIO), is a critical component of a highly available system. Multipathing software allows the Windows operating system to see and access multiple paths to the same LUN. This enables data to travel down any available path for increased performance or continued access to data in the case of a failed path.

Hitachi Data Systems recommends using the round robin load-balancing algorithm when using either Hitachi Dynamic Link Manager or MPIO to distribute load evenly over all available HBAs. Hitachi Data Systems used Hitachi Dynamic Link Manager Advanced to test this reference architecture.

Disk Management

Use of Hitachi's logical partitioning technology in this solution requires that the virtual servers use SAN boot. This reference architecture uses the Hitachi Virtual Storage Platform to host the OS volumes used by the SQL server hosts as well as the Active Directory host.

The OS volumes, the SQL database, tempdb, and transaction log volumes are presented to the physical blade using the virtual WWNs created automatically after enabling the logical partitioning feature. This ensures that the volumes maintain a dedicated connection to only the logical partitions that should have access to them.

Storage Area Network

The storage area network (SAN) configuration for this reference architecture uses two Fibre Channel switches for high availability. There are four redundant paths configured from the Hitachi Compute Blade 2000 for the Microsoft SQL Server 2008 R2 databases, `tempdb,` and transaction logs to the Hitachi Virtual Storage Platform. The SQL server host's OS volumes have two redundant paths configured to the Virtual Storage Platform. The server has two dual port host bus adapters (HBAs) installed for high availability purposes.

Hitachi Dynamic Link Manager is used for multipathing, employing the round-robin multipathing policy. Hitachi Dynamic Link software's round-robin load balancing algorithm automatically selects a path by rotating through all available paths. This balances the load across all available paths and optimizing IOPS and response time.

Figure 6 shows the Fibre Channel SAN architecture for the Microsoft SQL Server 2008 R2 implementation described in this reference architecture guide.

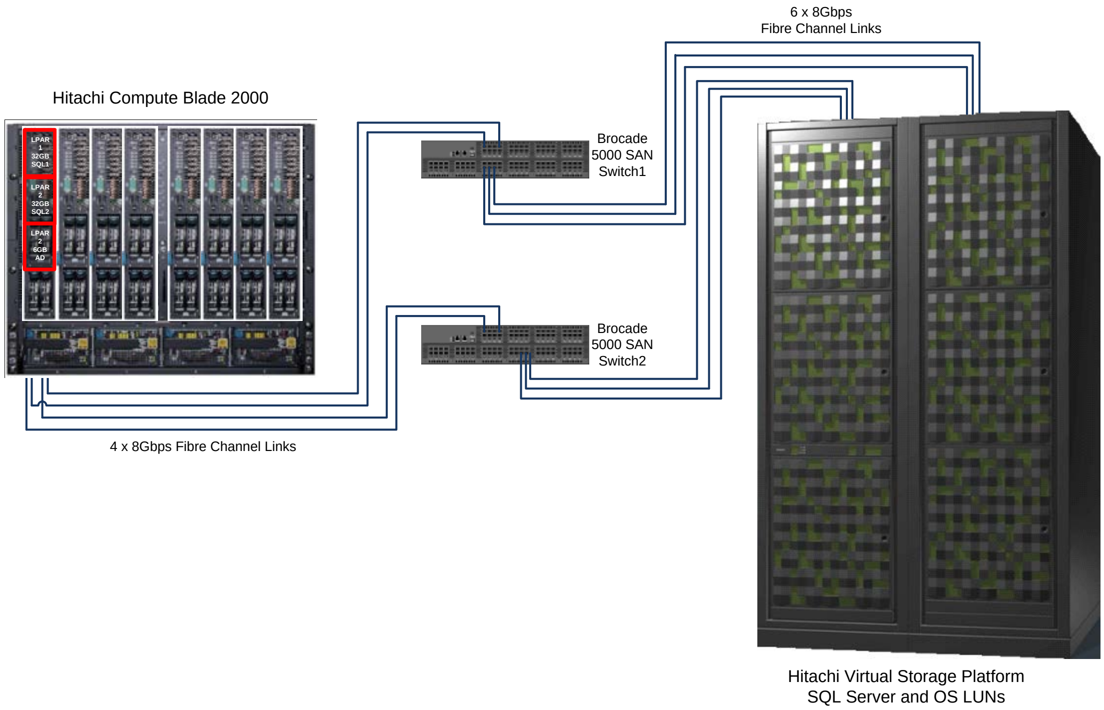

Figure 6

Path Configuration

All LUNs are masked with the ports from the Hitachi Virtual Storage Platform zoned to all four HBAs on each of the SQL server logical partitions. Four Fibre Channel ports on the Hitachi Compute Blade 2000 are used. Table 5 lists the connections between the Hitachi Compute Blade 2000 and the Hitachi Virtual Storage Platform ports.

Table 5. Path Configuration

Host HBA and Port	*Switch*	*Zone Name*	*Storage System Port*	*Storage System Host Group*
HBA 1 port 1	Brocade 5100 - 1	HCB2K_01_B0_L1_HBA1_1_ASE45_36_1C_4B	1C 4B	SQL_Server_1_HBA1_1
HBA 1 port 2	Brocade 5100 - 2	HCB2K_01_B0_L1_HBA1_2_ASE45_36_7C	7C	SQL_Server_1_HBA1_2
HBA 2 port 1	Brocade 5100 - 1	HCB2K_01_B0_L1_HBA2_1_ASE45_36_3C	3C	SQL_Server_1_HBA2_1
HBA 2 port 2	Brocade 5100 - 2	HCB2K_01_B0_L1_HBA2_2_ASE45_36_5C_7B	5C 7B	SQL_Server_1_HBA2_2
HBA 1 port 1	Brocade 5100 - 1	HCB2K_01_B0_L2_HBA1_1_ASE45_36_1C_4B	1C 4B	SQL_Server_2_HBA1_1

Host HBA and Port	Switch	Zone Name	Storage System Port	Storage System Host Group
HBA 1 port 2	Brocade 5100 - 2	HCB2K_01_B0_L2_HBA1_2_ASE45_36_7C	7C	SQL_Server_2_HBA1_2
HBA 2 port 1	Brocade 5100 - 1	HCB2K_01_B0_L2_HBA2_1_ASE45_36_3C	3C	SQL_Server_2_HBA2_1
HBA 2 port 2	Brocade 5100 - 2	HCB2K_01_B0_L2_HBA2_2_ASE45_36_5C_7B	5C 7B	SQL_Server_2_HBA2_2
HBA 1 port 1	Brocade 5100 - 1	HCB2K_01_B0_L3_HBA1_1_ASE45_36_4B	4B	AD2_HBA1_1
HBA 2 port 2	Brocade 5100 - 2	HCB2K_01_B0_L3_HBA2_2_ASE45_36_7B	7B	AD2_HBA2_2

Storage Building Block

Designing a SQL server implementation using logical partitioning virtualization is no different than performing the same activities in a non-virtualized environment. Deploying Microsoft SQL Server 2008 R2 using a building block approach allows you to easily manage and scale your environment. While the two reference architectures differ in capacity and in number of SQL server instances, they both use the same underlying storage building block. All the LUNs created for the environment followed these best practices:

- For dynamically provisioned environments, place database and log files on separate dynamic tiering pools.
- Place the `tempdb` files on a separate LUN within the dynamic tiering pool for performance and ease of monitoring. While `tempdb` files can be placed in the same dynamic tiering pools as the database files, sharing the same LUN for data files and tempdb files can negatively affect performance.
- When using Hitachi Dynamic Tiering, use both SAS and SSD disks in the dynamic tiering pool used for the database and tempdb volumes. This ensures the best performance as well as the best use of your disks capacity. RAID groups used in dynamic tiering pools are configured as RAID-5.
- When calculating required capacity, always account for approximately 20 percent additional overhead for the database, `tempdb`, and transaction logs LUs. Additional capacity might be required for your transaction logs, depending on your backup recovery model.
- When increasing the storage capacity for your environment, maintain the pool tier capacity ratios to scale the performance appropriately. This ensures that the performance requirements continue to be met.
- Scale the SQL server logical partition's CPU and memory to meet expected database workloads.

Keeping these guidelines in mind, you can build a scalable architecture that meets your performance and capacity requirements, and ensure that you have a scalable and a highly available environment.

Figure 7 shows how an architecture built on a Hitachi Virtual Storage Platform that starts with a 900GB database, a 90GB `tempdb`, and a 225GB transaction log can scale up following the building block guidelines.

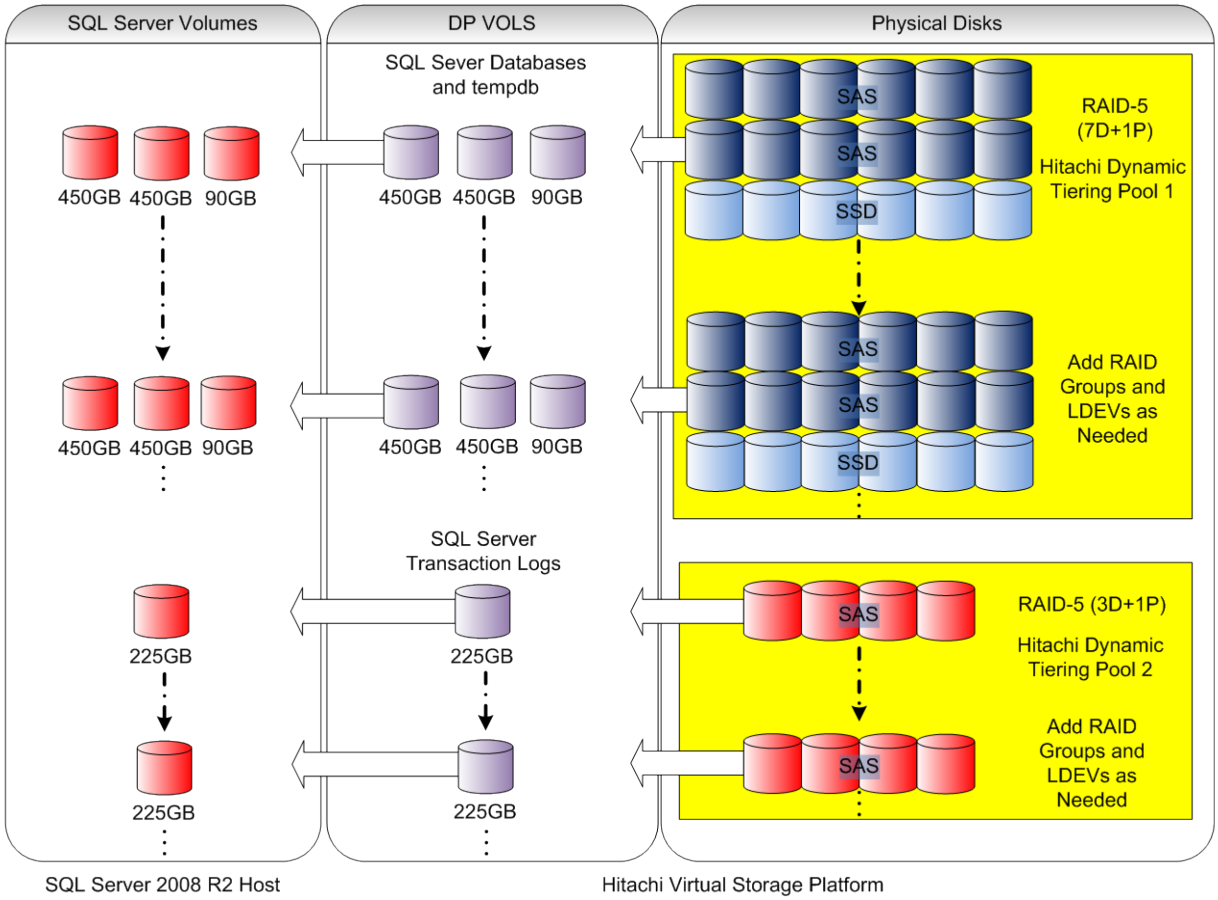

Figure 7

Only a third of the capacity available on the SSD RAID group is allocated per building block. This enables using the rest of the SSD RAID group capacity for other dynamic tiering pools that require high performance while maintaining the performance levels expected for the SQL server environment.

LUNs are presented to the SQL servers using virtual WWNs created during the logical partition configuration. Hitachi Data Systems recommends using mount points for the SQL server database, `tempdb`, and transaction log volumes. This avoids the limitation of the number of drive letters available when making additional volumes available to your host.

Following Microsoft's recommendations, each LUN has at least 20 percent additional capacity. There is also additional capacity provided to ensure that the architecture continues to meet performance requirements for the environment.

Two separate dynamic tiering pools were used in this environment. For the database and `tempdb` dynamic tiering pool two RAID-5+0 (7D+1P) groups using SAS drives and a third of the capacity from the RAID-5 (7D+1P) group using SSD drives were used. For the transaction log volumes dynamic tiering pool, a single RAID-5 (3+1) group was used. Independent LUNs were created for each file type.

Although a single dynamic tiering pool can be created for database and transaction log files, a dual dynamic tiering pool architecture provides the same level of protection you have on a standard provisioned environment at the RAID group level to a Hitachi Dynamic Tiering environment.

Table 6 describes the 75,000-user building block architecture for a single server on a logical partition.

Table 6. Building Block Resources for Single Server on a Logical Partition

Resource	Building Block
Logical Partition Server	32GB memory 4 virtual processors
Number of supported users	75,000
LUNs	2 × 450GB DP-VOLs for database files 1 × 90GB DP-VOL for `tempdb` files 1 × 225GB DP-VOL for transaction logs
Disk configuration	2 × RAID-5 (7D+1P) with 16 × 300GB 10K RPM SAS and 1 × RAID-5 (7D+1P) with 8 × 300GB SSD drives* for databases and `tempdb` 1 × RAID-5 (3D+1P) with 4 × 300GB 10K RPM SAS drives for transaction log

*A single LDEV using one-third of the total capacity was created from the SSD RAID group

To scale this environment to support 150,000 users, add a second building block. This building block uses a 900GB database for each SQL server instance in a dynamically provisioned environment. The databases all use RAID-5+0 with a 7D+1P drive configuration for both the SAS as well as the SSD tier of the pool.

When increasing the number of SQL server machines in your environment, you can increase the size of the existing pool or create another pool. Base your decision on performance requirements.

For scalability purposes, this reference architecture uses a third dynamic provisioning pool to host the virtual machine OS. As with the data pools, this approach enables you to more efficiently scale up your architecture.

The Hitachi Compute Blade 2000 server blade used in this test environment contains 128GB of memory and two eight-core Intel Xeon CPUs. If you require a high availability environment, use a minimum of two Hitachi Compute Blade 2000 server blades.

You can add SQL server hosts to the blade by creating additional logical partitions and allocating more storage resources on the Hitachi Virtual Storage Platform to scale up to larger databases. Additionally up to four X57A1 blades can be connected using the SMP interface connector to create a single eight-socket SMP system with up to 64 cores and 1024GB of memory.

Storage Configuration

This reference architecture uses the Hitachi Virtual Storage Platform with the Hitachi Dynamic Tiering to provision the LUNs used by SQL server databases, tempdb, and the transaction logs on a Microsoft SQL Server instance. Two SQL server instances were deployed to test and validate that the building block architecture scales as resources are added to the dynamic tiering pool. The RAID group that hosts the host's OS resides on a separate RAID-5 (3D+1P) dynamic tiering pool.

Hitachi Dynamic Tiering distributes the I/O workload across all drives within each tier of the dynamic tiering pool.

Table 7 lists the LUN configuration for a single building block using Hitachi Dynamic Tiering. Note that all storage is in dynamic tiering pools 1 and 2 except for the LUNs used for the guest OS VHD, which are in pool 3.

Table 7. Building Block Storage Configuration

LUN	Size (GB)	Description	Pool ID
3	150	OS for SQL Server VM operating system	3
4	450	Database for SQL Server	2
5	450	Database for SQL Server	2
6	90	`tempdb` for SQL Server	2
7	225	Transaction logs for SQL Server	1

Scaling the Storage

You can scale your environment by adding RAID groups to the existing dynamic tiering pools, whether using the single building block with 900GB deployment, or the two building blocks with 1.8TB environment. Hitachi Data Systems testing shows that this reference architecture meets Microsoft's throughput guidelines as it scales. Most importantly, the tests results indicate that the overall latency values dropped by almost half and the IOPS increased by approximately 40% in comparison to tests using RAID-1+0 RAID groups and twice the amount of SAS disks.

If deploying additional SQL workloads, you can choose to increase the capacity of the existing pools currently being used. This should be done only if the performance characteristics of the additional workload and the effect on other workloads in the pool are understood.

If an existing pool cannot support additional workloads, an alternative is to create an additional dynamic tiering pool. When establishing your workload requirements, consider additional workloads, such as backup and replication operations (internal to the storage system or external such as transaction log shipping). These operations also can affect the overall pool performance bandwidth.

A dynamic tiering pool is made up of one or more logical devices (LDEVs) and up to three different tiers. To increase your pool capacity for a tier, add LDEVs to the pool. When adding new LDEVs, the dynamic tiering pool tier is unbalanced because the data is not spread evenly across all the spindles in that tier. Additionally, the new data blocks that are saved to the pool may not be allocated to the correct tier, as the pool was optimized for the previously existing data in the pool.

Hitachi Dynamic Tiering contains all of the features available in Hitachi Dynamic Provisioning, including a feature called reclaim zero pages. This corrects unbalanced conditions by optimizing the tiers within each dynamic tiering pool. Execute the optimization carefully. It can affect performance, depending on how heavily the pool is used.

The nature of the optimization process is to relocate data by distributing it across all available spindles within one or more tiers in a dynamic tiering pool, depending on the capacity of those tiers. This feature does not automatically move the data blocks across different tiers within a dynamic tiering pool. The move of data blocks between tiers within a dynamic tiering pool can be done manually by a user or automatically by selecting a monitoring cycle between 1 and 24 hours.

Hitachi Data Systems test environment used a 1 hour monitoring cycle for the reference architecture described in this paper. Your monitoring cycle may vary depending on the type of workload and, most importantly, the type of SQL server deployment you have.

Always test your design before deploying it in your production environment. 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 SQL servers to achieve processor utilization rates of 90+ percent in Windows Performance Monitor while maintaining latency levels that meet Microsoft recommendations.

A second Hitachi Virtual Storage Platform might be needed for more complex environments that require failover capability between multiple data centers. You can automatically or manually failover SQL servers to other resources within the same data center or between data centers using one or more of the following:

- Hitachi Storage Cluster
- Failover clustering
- Hitachi Compute Blade 2000 N+1 or N+M cold standby

Windows Server 2008 and Hitachi tools enable you to monitor the resource utilization in your environment. Windows Performance Monitor enables you to monitor the CPU and memory utilization of the Hyper-V hosts. Monitor SQL server hosts with Performance Monitor on the machine's OS. Hitachi Tuning Manager provides a holistic view of all the Virtual Storage Platform performance related counters. Tuning Manager helps to do the following:

- Identify potential bottlenecks that might require you to add disks to the environment.
- Monitor the SQL server hosts and the SQL server instance and databases.

Find out more information about the Hitachi Tuning Manager software.

Scaling Microsoft SQL Server 2008 R2

This section describes planning and deployment considerations to keep in mind when scaling your Microsoft SQL Server 2008 R2 environment.

Planning

Scaling SQL server databases and instances requires planning and testing. When scaling up, both capacity and performance are concerns. Properly test any production environment to ensure it satisfies end-user requirements.

Calculate the maximum number of logical partitions that a single Hitachi Compute Blade 2000 server blade can support by noting the total number of CPUs and total memory requirements. Hitachi Data Systems recommends deploying a single instance of SQL server per logical partition for improved performance and security isolation. Deploy additional logical partitions if you need more SQL server instances to support your environment.

When deploying SQL server databases, it is important to understand the key database file types. By understanding the type of workload that each file type has, along with the database type and size, storage and database administrators (DBAs) can establish storage requirements for the SQL server environment.

SQL Server Transaction Log File

Every SQL server database has at least one log volume and file that records database modifications made by each transaction. They are critical components of the database for availability and data integrity.

In the event of a system failure, at least the active transaction log is required to bring the database back to a consistent state. The transaction logs are written before the data records are updated to the database file via the checkpoint process.

The logs can be used to roll back transactions if corruption is discovered later or to recover a committed transaction if an error writing to the database occurs.

Response time and performance are critical considerations for separating the transaction log from the database files. Microsoft suggests aiming for log I/O response times between 1 and 5 milliseconds. Hitachi's optimized caching and proper storage design ensure that the logs can be written with minimal delay.

A feature built into the Hitachi Virtual Storage Platform is the ability to optimize physical I/O based on recognition of I/O patterns. The Virtual Storage Platform controller makes optimized timing decisions about when to move the data between mirrored, protected cache, and physical disks when it encounters a series of I/O requests. By combining multiple logical I/O requests into a single physical I/O or by optimizing the order of individual reads and writes, the Virtual Storage Platform can significantly increase overall performance.

SQL Server tempdb Files

SQL server `tempdb` files are used for storage of temporary data structures and can be subject to intense and unpredictable I/O. Many best practice recommendations suggest locating `tempdb` files on separate RAID groups from the database and using the fastest disks available. This is generally a safe recommendation. The load on `tempdb` is highly dependent upon database and application design.

However, if the `tempdb` load is well understood and monitored regularly, testing by Hitachi Data Systems shows that the `tempdb` files can reside on the same RAID group as the databases without adverse effects.

If the environment does not have sufficient physical disk I/O resources to meet the combined requirements of `tempdb` and database files, performance for all databases in the SQL server instance degrades. Have a good understanding of your `tempdb` usage, regardless of where you choose to place the `tempdb` files.

By default, `tempdb` has a single data file group and a single log file group with a default number of files set to 1. Additional data files can be created for `tempdb`. Create at least as many data files of equal size as you have CPU cores. The number of concurrent threads is less than or equal to the number of CPU cores.

If the `tempdb` file does not perform well, the system does not perform well. New allocations favor the newer file because of availability free space. This can cause allocation problems. You must restart SQL server when you add a new file. After adding or modifying the files on `tempdb`, restart the SQL server instance.

SQL Server Database Files

In most cases, SQL server database file I/O is made up of random small record reads and writes. A database might include only a single database file, while those designed to support heavy transactional workloads or large schemas might use a variety of file group architectures to improve performance, operational convenience, or availability.

Performance considerations include monitoring the Average Disk sec/Read and Average Disk sec/Write counters. This indicates latency values for reads and writes, respectively. I/O size, RAID configuration, and other factors in the data path can affect average response time for logical or physical disks. Lower measures of disk latency are better than higher measures, but can vary depending on the size and nature of the I/Os being issued. These numbers also vary across different storage configurations. Cache size and utilization can greatly affect these measures. Acceptable database response I/O response times typically range between 5 and 20 milliseconds.

Storage Design Considerations

Configure storage for the SQL server databases first for performance and then capacity, as based on the performance levels that a SQL server logical partition can handle. This is no different from how you configure SQL server storage for a standalone server deployment. The architectures described in this section require completing the following calculations for an existing or planned Microsoft SQL Server 2008 environment:

- IOPS needed for SQL server user databases and transaction logs
- Overall capacity needed for user databases and transaction logs, including planned growth for SQL server instances

Design Goals

This solution's building block architecture achieves the following design goals for an OLTP workload:

- Optimize storage configuration on the Hitachi Virtual Storage Platform with Hitachi Dynamic Tiering for best I/O throughput, database latency, and ease of management.
- Maximize storage utilization while maintaining or improving the environment performance levels using Hitachi Dynamic Tiering.
- Deliver sustainable and acceptable levels of IOPS falling within the 1ms to 5ms response time range for transaction log file I/O and 1ms to 20ms response time range for data file I/O.
- Deliver at least 80 percent disk capacity utilization for the database volumes.

Several design factors were considered to arrive at the building block architecture, including:

- For a very heavy user profile, you might need additional disks available to support the required IOPS. As with any SQL server deployment, make sure to test your environment to ensure the proper number of disks are available to the database from an IOPS perspective.
- Determine the granularity of scale by the database size along with the SQL server logical partition provisioned hardware. Provision additional storage by adding RAID-5 groups for database, tempdb files, and the transaction logs. For the dynamic tiering pools, the specific drive type depends on the initial building block environment that meets both performance and capacity requirements for your environment.

- In a dynamically provisioned configuration, allocate dynamic tiering pool space by adding an LDEV from a RAID-5 group to the dynamic tiering pool. After that, create the required LUNs and assign them to a SQL server host.
- Using a 1 hour monitoring cycle, allow between 1 to 8 cycles to take place prior to establishing whether the correct ratio of tiers is devised for your environment. It may take between at least 4 monitoring cycles for your SQL server data to move to the appropriate tier in order to provide you with the best performance for your environment.

Engineering Validation

To validate this reference architecture, an industry-standard OLTP workload that simulates a stock brokerage scenario was exercised on the building block architecture. An additional SQL server host was added along with added storage capacity for the same previously used dynamic tiering pool 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 the 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 8 lists the OLTP workload parameters used as a baseline for the tests.

Table 8. 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	≥4 hours

Table 9 lists results for the OLTP workload tests for the 75,000 user configuration.

Table 9. OLTP Workload Test Results for 75,000 User Configuration

Metric	*Microsoft Success Criterion*	*Result*
Achieved Database IOPS	Varies	10,573 IOPS
Database Avg. Disk sec/Read	≤ 20ms	11ms
Database Avg. Disk sec/Write	≤ 20ms	1ms
Transaction Log Avg. Disk sec/Read	≤ 5ms	N/A
Transaction log Avg. Disk sec/Write	≤ 5ms	1ms

Table 10 lists test results for the OLTP workload tests for the 150,000 user configuration. The latency values for the database and transaction logs were calculated by taking the average of the volumes in each SQL server.

Table 10. OLTP Workload Test Results for 150,000 User Configuration

Metric	Microsoft Success Criterion	Result
Achieved Database IOPS	Varies	21,562 IOPS
Database Avg. Disk sec/Read	≤ 20ms	8.75ms
Database Avg. Disk sec/Write	≤ 20ms	2.5ms
Transaction Log Avg. Disk sec/Read	≤ 5ms	N/A
Transaction Log Avg. Disk sec/Write	≤ 5ms	1ms

Testing shows that this solution meets or exceeds all design goals for the reference architecture. Testing also shows that we can achieve a better capacity utilization of the drives used in this building block in comparison to a building block using only RAID-1+0 RAID groups while decreasing the latency time and increasing the IOPS in comparison to that building block.

Conclusion

This reference architecture guide describes a building block-based storage solution for virtualized Microsoft SQL Server 2008 environments that is easy to deploy and maintain, provides high availability and flexible scalability, and delivers predictable performance. It also improves resource utilization, reduces server sprawl, shrinks data center footprints, maximizes the capacity utilization of the drives being used, and lowers energy costs.

Hitachi Data Systems Global Services offers experienced storage consultants, proven methodologies, and a comprehensive services portfolio to assist you in implementing Hitachi products and solutions in your environment. For more information, see the Hitachi Data Systems Global Services web site.

Live and recorded product demonstrations are available for many Hitachi Data System products. To schedule a live demonstration, contact a sales representative. To view a recorded demonstration, see the Hitachi Data Systems Corporate Resources web site. Click the **Product Demos** tab for a list of available recorded demonstrations.

Hitachi Data Systems Academy provides best in class training on Hitachi products, technology, solutions, and certifications. Hitachi Data Systems Academy delivers on-demand web based training (WBT), classroom based instructor led training (ILT), and virtual instructor led training (vILT) courses. For more information, see the Hitachi Data Systems Academy web site.

For more information about Hitachi products and services, contact your sales representative or channel partner, or visit the Hitachi Data Systems web site.

Hitachi Data Systems Corporation

Corporate Headquarters
750 Central Expressway,
Santa Clara, California 95050-2627 USA
www.hds.com

Regional Contact Information
Americas: +1 408 970 1000 or info@hds.com
Europe, Middle East and Africa: +44 (0) 1753 618000 or info.emea@hds.com
Asia Pacific: +852 3189 7900 or hds.marketing.apac@hds.com