Deploying SQL Server 2008 R2 with Hyper-V on the Hitachi Virtual Storage Platform

Reference Architecture Guide

By Rick Andersen

October 2010

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 SQL Server 2008 R2 with Hyper-V on the Hitachi Virtual Storage Platform

Reference Architecture Guide

Businesses of all sizes need storage solutions for their virtualized applications that are easy to deploy and maintain, provide high availability and flexible scalability, and deliver predictable performance. These businesses also need to improve resource utilization, reduce server sprawl, and create green data centers with smaller footprints and lower energy costs. When the virtualized application is as mission-critical as SQL Server 2008, the stakes are even higher.

The Hitachi Virtual Storage Platform can help you leverage your information, which is the new currency in today's data-driven economy. Information, which exists in many forms, must be protected and readily accessible to ensure business survival and success. The Virtual Storage Platform maximizes cost efficiency and return on investment by creating an agile storage infrastructure that reduces costs and increases performance, availability, scalability and reliability.

The Hitachi Virtual Storage Platform is the industry's only 3D scaling storage platform. With the unique ability to concurrently scale up, scale out and scale deep in a single storage system, the Virtual Storage Platform flexibly adapts for performance, capacity, connectivity and virtualization. No other enterprise storage platform can dynamically scale in three dimensions. Scaling up allows you to increase virtual server consolidation, improve utilization of resources, and reduce costs. Scaling out allows you to meet increasing demands by combining multiple chassis into a single logical system with shared resources. Scaling deep extends the advanced functions of the Virtual Storage Platform to external multivendor storage.

Many SQL Server instances deployed in data centers operate on dedicated server hardware that can lead to server sprawl and underutilized hardware resources. Each dedicated server requires management, physical space within the data center, and power and cooling, all of which are finite resources. Deploying technologies such as the Hitachi Virtual Storage Platform, Hyper-V, and SQL Server 2008 R2 help to resolve the problem of ever-increasing demands on these finite resources through consolidation and virtualization.

By virtualizing SQL Server 2008 R2 on the Hitachi Virtual Storage Platform, you can consolidate the number of dedicated servers running SQL Server instances by deploying these as virtual machines running under a Hyper-V host. Hitachi Dynamic Provisioning software allows you to combine these different workloads in a single frame for greater flexibility.

This reference architecture guide describes a building block design that helps large and enterprise deployments achieve easy-to-manage, highly available, scalable virtualized SQL Server deployments that improve resource utilization and reduce costs. It provides best practices required to successfully deploy Microsoft SQL Server 2008 R2 in a virtualized environment using Microsoft Hyper-V, Hitachi Dynamic Provisioning software 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. Readers need to be familiar with general storage, Hyper-V, Windows 2008 and SQL Server 2008 concepts.

Solution Overview

The solution described in this white paper provides a building block architecture that helps large and enterprise-scale deployments of SQL Server 2008 R2 achieve their critical business objectives. It provides best practices required to successfully deploy Microsoft SQL Server 2008 R2 in a virtualized environment using Microsoft Hyper-V 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 to150,000 user accounts. 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. In addition, a SQL Server VM and a storage building block were added to scale the configuration.

To achieve high availability for this solution, redundant physical paths were enabled via 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.

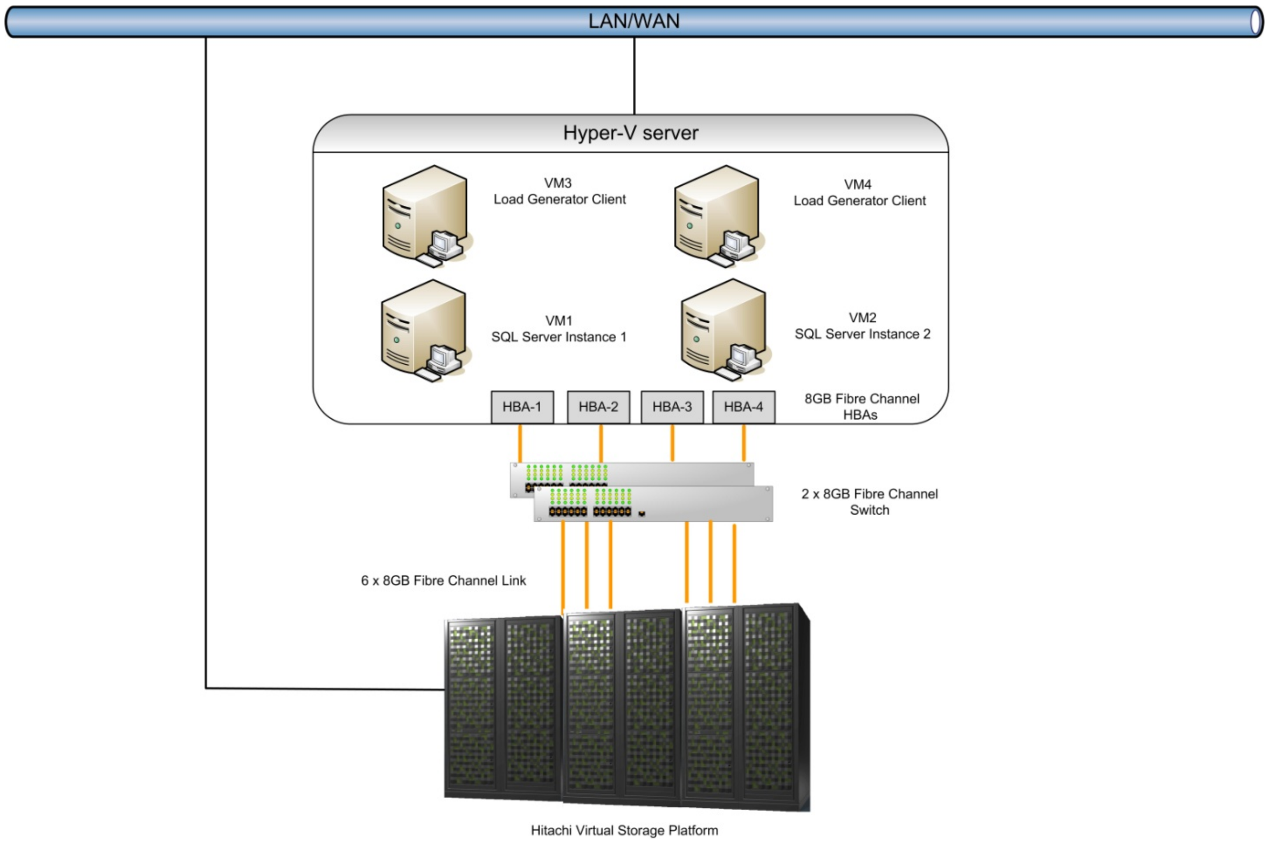

Figure 1

Key Solution Components

The following sections describe the key components needed to deploy this solution.

Virtual Storage Platform

The Hitachi Virtual Storage Platform is the industry's only 3D scaling storage platform. With the unique ability to concurrently scale up, scale out and scale deep in a single storage system, the new Virtual Storage Platform flexibly adapts for performance, capacity, connectivity and virtualization. No other enterprise storage platform can dynamically scale in three dimensions. The Virtual Storage Platform provides virtual storage that meets the growing demands of server virtualization.

The trend in server virtualization is to consolidate the I/O workload of many servers onto a single storage system. As more virtual machines are consolidated onto a physical host, storage systems must be able to dynamically add more storage resources to keep up with I/O demand. The 3D scaling capability of the Virtual Storage Platform meets that requirement.

Scaling up allows you to increase virtual server consolidation, improve utilization of resources, and reduce costs. With the Hitachi Virtual Storage Platform, you can increase performance, capacity and connectivity by adding cache, processors, connections and disks to the base system. A virtual server that accesses the storage system can use all these resources, which act as one system managed as a common pool of resources.

Scaling out allows you to meet increasing demands by combining multiple chassis into a single logical system with shared resources. By scaling out you can support increased resource needs in virtualized server environments.

Scaling deep extends the advanced functions of the Virtual Storage Platform to external multivendor storage. By dynamically virtualizing new and existing storage systems, those systems become part of the Virtual Storage Platform's pool of storage resources. Once virtualized, external data can then be migrated, tiered, replicated and managed by the Virtual Storage Platform. In this manner, older data storage systems can gain a longer useful life. You can extend distance replication for business continuity to lower-cost, lower-function storage systems by virtualizing them behind a Virtual Storage Platform.

The switch matrix architecture of the Virtual Storage Platform makes all of this possible. It connects the basic components, front-end directors, back-end directors, global cache modules and virtual storage directors. You can add redundant pairs of directors and cache modules as required without disruption to connected host servers. All these resources are tightly coupled through a global cache that creates a common pool of storage resources. These resources can include external storage that is connected through front-end director initiator ports.

Virtual Storage Platform Architecture

The Virtual Storage Platform offers an entirely new level of scalable enterprise storage, capable of handling the most demanding workloads while maintaining great flexibility. The Virtual Storage Platform offers much better performance, higher-performance scalability, higher reliability and greater flexibility than any storage system on the market today.

The Virtual Storage Platform offers these features:

- The HiStar-E PCI Express Switched Grid acts as the interconnection among front-end directors, back-end directors, data cache adapter boards and virtual storage director boards.
- Data accelerator processors on the front-end directors and back-end directors work with central processor boards called virtual storage directors that manage all I/O by sets of assigned logical devices (LDEVs).
- Dual SAS controllers on back-end director boards contain eight 6Gbps SAS links per board.
- The control memory function resides in global cache and each VSD board contains a local copy with information for its LDEVs. Most control memory accesses are lookups to the local copy.
- Global cache is backed up to solid state drives (SSDs) on the cache boards.
- Each virtual storage director board controls all I/O operations for a discrete group of LDEVs. LDEVs are assigned round-robin across the installed virtual storage director boards as they are created. If necessary, you can manually reassign LDEV ownership to a different virtual storage director.

- Each virtual storage director board executes the code for initiator mode (hosts), external mode (virtualization), back-end director mode, or the copy products send and receive modes. Code execution is done on a per-job basis.
- A Virtual Storage Platform can be scaled from a single-chassis system to a dual-chassis system. Each chassis has a control rack and a logic box.
- Up to 1,280 3.5-inch large form factor (LFF) drives or 2,048 2.5-inch small form factor (SFF) drives can be installed in a dual-chassis system. If you install both LFF and SFF disk containers and drives in a storage system, the limits change based on the configuration you choose.

The Virtual Storage Platform is built as a single-chassis or dual-chassis storage system. Each chassis has one control rack and up to two disk expansion racks. The control rack has the logic box that holds all of the control boards for a chassis and one or two disk containers. The disk expansion racks can hold three disk containers each. Disk containers come in two types: small form factor (up to 128 2.5-inch drives) and large form factor (up to 80 3.5-inch drives). When using two chassis as a single integrated storage system, the two units are cross connected at the grid switch level. The storage system behaves as a single unit, not as a pair of units operating as a cluster.

Figure 2 shows the two types of racks available for a Virtual Storage Platform.

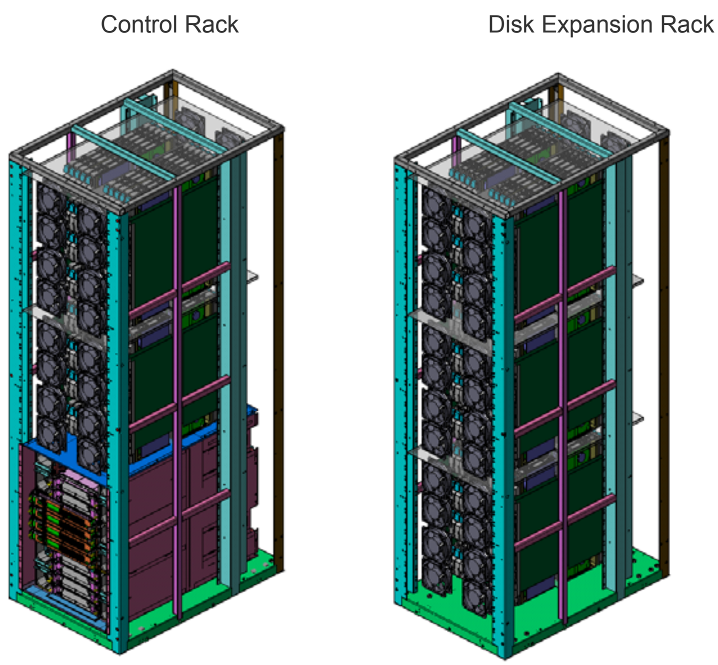

Figure 2

Figure 3 shows the logic boards in a fully populated single-chassis Virtual Storage Platform.

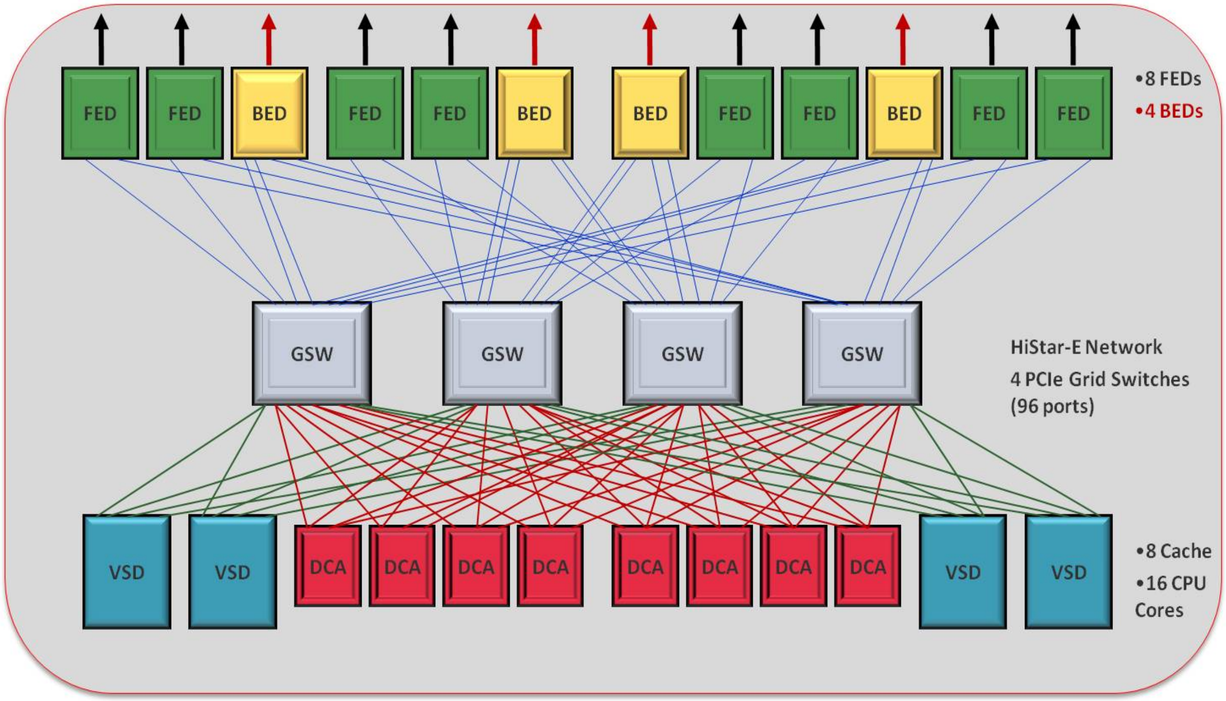

Figure 3

Figure 3 shows the following chassis components (note that a feature is a pair of boards on two separate power domains):

- **GSW** — Grid Switch PCI Express Switch. One or two features (two or four boards) per control unit with 24 2GB/sec HiStar-E ports each can be installed
- **DCA** — Data cache adapter cache memory. One to four features (two, four, six or eight boards) per control unit with up to 32GB of RAM each can be installed
- **VSD** — Virtual storage director processor module. One or two features (two or four boards) per control unit can be installed.
- **FED** — Front-end director host port module. One to four features (two, four, six or eight boards) per control unit of four or eight 8Gbps Fibre Channel ports can be installed.
- **BED** — Back-end director disk controller module. One or two features (two or four boards) per control unit with eight 6Gbps SAS links per board can be installed. If the back-end director options are not installed (available for the single-chassis configuration only), two additional front-end director options can be used in those chassis slots.

3D Scaling Architecture

The Hitachi Virtual Storage Platform allows for optimal infrastructure growth in all dimensions by scaling up, scaling out and scaling deep.

Scale Up

Scale up to meet increasing demands by dynamically adding processors, connectivity and capacity in a single unit, providing the highest performance for both open and mainframe environments.

In the basic single chassis configuration, the number of logic boards, disk containers, and drives is highly scalable. You can start with the minimum set of logic boards,10, and one disk container, then add more boards (up to a total of 28 boards in a single chassis) and disk containers (up to a total of eight disk containers in a single chassis). Disk container types may be intermixed within a chassis.

Scale Out

Scale out to meet multiple demands by dynamically combining multiple units into a single logical system with shared resources, support increased demand in virtualized server environments, and ensure safe multitenancy — that is, the ability to run multiple servers simultaneously without the risk of corruption or modification of data from one server to another — and quality of service through partitioning of cache and ports.

You can double the scalability of the Virtual Storage Platform with a dual-chassis system with up to six racks. The logic box in each chassis is the same, using the same types and numbers of logic boards. Any front-end port can access any back-end RAID group; no division within the storage system exists between the chassis.

A dual-chassis Virtual Storage Platform can manage up to 247PB of total storage capacity.

Table 1 lists the capacity differences between a single-chassis and a dual-chassis Virtual Storage Platform storage system.

Table 1. Virtual Storage Platform Chassis Capacity Comparison

Maximum Capacity	*Single Chassis*	*Dual Chassis*
Data cache	256GB	512GB
Raw cache bandwidth	64GB/s	128GB/s
Solid state drives	128	256
2.5” SFF drives	1,024	2,048
3.5” LFF drives	640	1,280
Logical volumes (LDEVs)	65,280	130,560

Scale Deep

Scale deep to extend storage value by dynamically virtualizing new, existing external storage systems, extend Hitachi Virtual Storage Platform advanced functions to multivendor storage and offload less demanding data to external tiers to optimize the availability of your tier one resources.

The Virtual Storage Platform provides the virtualization mechanisms that allow other storage systems to be attached to some of its front-end director Fibre Channel ports and accessed and managed via hosts that are attached to the host ports on the Virtual Storage Platform. As far as any host is concerned, all virtualized logical units passed through the Virtual storage Platform to the hosts appear to be internal logical units from the Virtual Storage Platform. The front-end ports on the Virtual Storage Platform that attach to the external storage system's front-end ports are operated in external or SCSI initiator mode (attached to servers), rather than the usual SCSI target mode (attached to hosts).

For more information about the Hitachi Virtual Storage Platform, see the Hitachi Data Systems web site.

Hitachi Dynamic Provisioning Software

On the Virtual Storage Platform, Hitachi Dynamic Provisioning software provides wide striping and thin provisioning functionalities. In the most basic sense, Hitachi Dynamic Provisioning software is similar to the use of a host-based logical volume manager (LVM), but with several additional features available within the Hitachi Virtual Storage Platform and without the need to install software on the host or incur host processing overhead. Hitachi Dynamic Provisioning software provides for one or more pools of wide striping across many RAID groups within a Virtual Storage Platform. One or more Dynamic Provisioning virtual volumes (DP-VOLs) of a user-specified logical size of up to 60TB (with no initial physical space allocated) are created against each pool.

Primarily, you deploy Hitachi Dynamic Provisioning software to avoid the routine issue of hot spots that occur on logical devices (LDEVs) from individual RAID groups when the host workload exceeds the IOPS or throughput capacity of that RAID group. By using many RAID groups as members of a striped Dynamic Provisioning pool underneath the virtual or logical volumes seen by the hosts, a host workload is distributed across many RAID groups, which provides a smoothing effect that dramatically reduces hot spots.

Hitachi Dynamic Provisioning software also carries the side benefit of thin provisioning, where physical space is only assigned from the pool to the DP-VOL as needed using 42MB pages, up to the logical size specified for each DP-VOL. A pool can also be dynamically expanded by adding more capacity or reduced by withdrawing pool capacity. Either operation is performed without disruption or requiring downtime. Upon expansion, a pool can be rebalanced so that the data and workload are wide striped evenly across the current and newly added RAID groups that make up the pool.

Hitachi Dynamic Provisioning software's thin provisioning and wide striping functionalities 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 the storage administration process.
- Provides performance improvements through automatic optimized wide striping of data across all available disks in a storage pool.
- Eliminates hot spots across the different RAID groups by smoothing the combined workload.
- Significantly improves capacity utilization.

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

Microsoft Windows Hyper-V

Microsoft Windows Hyper-V is a hypervisor-based virtualization technology that is integrated into Windows Server 2008 x64 and Windows Server 2008 R2 versions of the operating system. It allows for the reduction of hardware footprints and capital expenses through server consolidation. This is accomplished by consolidating multiple physical servers that are hosting SQL Server instances into a single Hyper-V server.

Additional options are available with Hyper-V, such as quick and live migration, which provide high availability for SQL Server virtual machines. One of the requirements for making the SQL Server virtual machines highly available is that they must be hosted in a Hyper-V failover cluster configuration.

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 Server2008 R2 also provides enhanced enterprise-class manageability for large OLTP deployments like the one that is described in this reference architecture guide. Together with the Hitachi Virtual Storage Platform, SQL Server 2008 R2 provides a scalable, high-performance database engine for any midrange to enterprise level application.

Solution Design

This section provides detailed information on the large-scale SQL Server 2008 R2 environment design used for this reference architecture. Although it includes both application and hardware design information, it is focused 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. A single virtual machine (VM) running Windows Server 2008 R2 and SQL Server 2008 R2 with underlying storage from the Virtual Storage Platform make up each building block. The Virtual Storage Platform takes advantages of Hyper-V features such as live migration via storage replication technologies to establish a reliable and highly available virtualized SQL Server solution. This reference architecture scales from a 75,000-user deployment on one 900GB database to a 150,000-user deployment on two 900GB databases by implementing a second virtual machine with a SQL Server instance.

The building block for the 75,000-user database consists of two 450GB LUNs for the SQL database, a 90GB LUN for `tempdb` and a 225GB LUN for the SQL transaction log. This building block was replicated along with adding a second VM and SQL Server instance to support 150,000 users on the Virtual Storage Platform. All LUNs were provisioned from Dynamic Provisioning pools on the Hitachi Virtual Storage Platform. Figure 4 shows the high level storage design of the SQL Server 2008 R2 reference architecture on the Virtual Storage Platform.

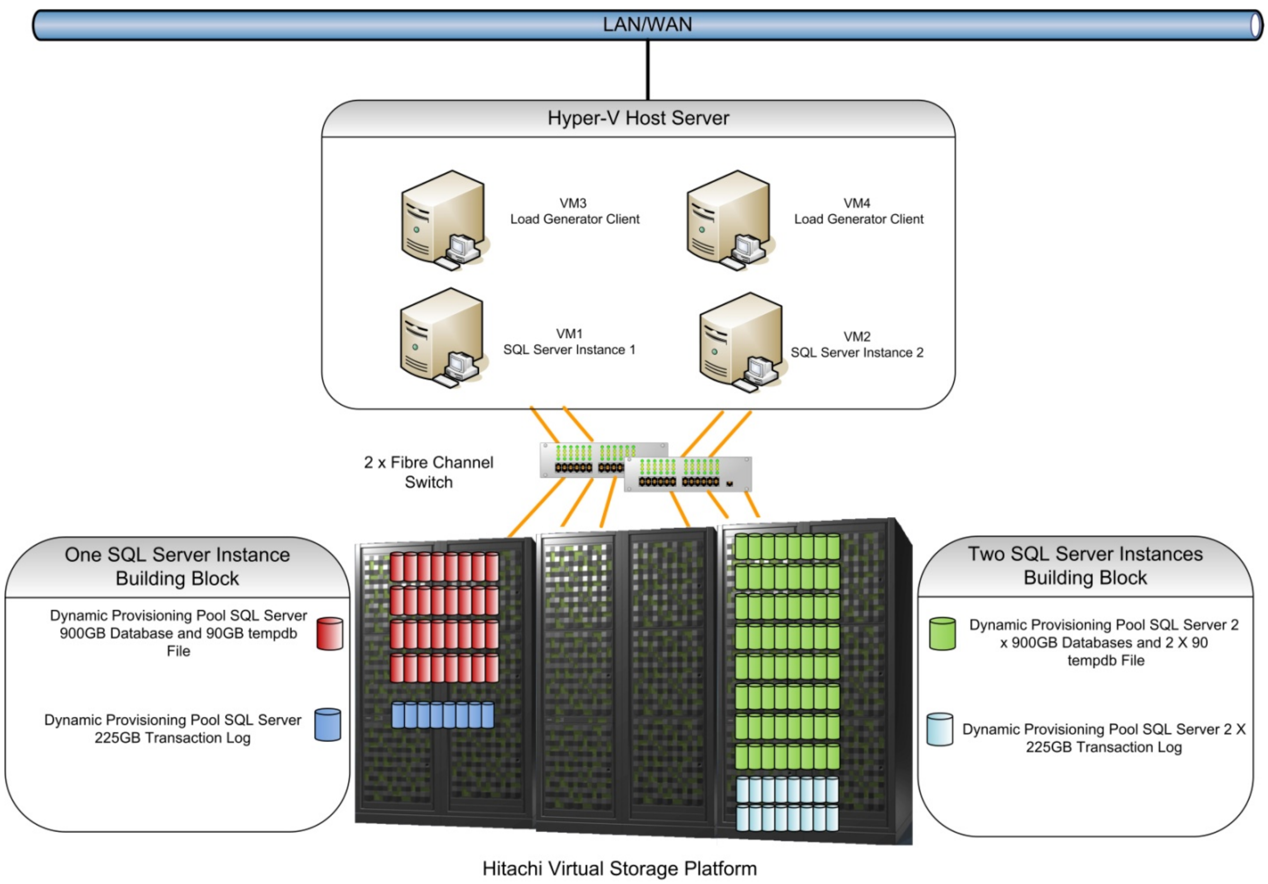

Figure 4

LUNs from the Virtual Storage Platform are allocated to the Hyper-V servers and made available to the virtual machines by using the pass-through disks approach For further information on pass-through disks please see the “Disk Management” section of this paper. These disks are made available to the Windows Server 2008 R2 guest operating system and can be partitioned and used for NTFS file systems for the SQL Server 2008 R2 databases, `tempdb`, transaction logs and BLOB storage, although BLOB storage is outside of this paper’s scope. The solution building block supports databases starting at 900GB capacity and 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, both from storage and host perspectives.

For this solution, Hitachi Data Systems connected the Hyper-V host server and the 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, SAN and server components used in this reference architecture.

Table 2. Hardware Components

Hardware	Detail Description	Version	Quantity
Hitachi Virtual Storage Platform storage system	128GB cache memory, 6 x 8Gb per front- end ports, 72 x 300GB, 10K RPM, SAS disks	Microcode 70-00-53	1
Brocade 5340 switch	SAN switch with 8Gb Fibre Channel ports	FOS 6.4.0b	2
Dell PowerEdge R905 server	4 x Quad-Core AMD Opteron processor 1.9 GHz, 128GB RAM. Equipped with 2 x Emulex LPe111-H 4 GB HBAs	Windows 2008 R2 64-bit edition	1

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, enabling 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 in both Hitachi Dynamic Link Manager software and MPIO to distribute load evenly over all available HBAs. Hitachi Data Systems used MPIO to test this reference architecture.

Host Software

Table 3 describes the software used in this reference architecture.

Table 3. Software Components

Software	*Version*
Hitachi Storage Navigator	Dependent on microcode version
Hitachi Dynamic Provisioning	Dependent on microcode version
Microsoft MPIO	006.0001.7600.16385
Windows Server 2008 R2 (for Hyper-V server)	Datacenter edition
Windows Server 2008 R2 (for SQL Server virtual machines)	Enterprise edition
SQL Server 2008 R2	Enterprise edition

Disk Management

Hyper-V servers can access LUNs in two ways, using passthrough disks or using virtual hard disks (VHDs) that are created on a LUN that is mapped to the physical host. Passthrough disks enable the LUNs to be available to a virtual machine as if they were directly connected to it. This reference architecture uses passthrough disks for the database, `tempdb` and log files. For the virtual machines operating systems, the LUNS are presented to the physical host and used as containers to house the virtual machine OS VHD file or additional data files.

Storage Area Network

The storage area network (SAN) configuration for this reference architecture uses two Fibre Channel switches for high availability. Four redundant paths are configured from the Hyper-V host for the SQL Server 2008 R2 databases, `tempdb` and transaction logs to the Virtual Storage Platform. For the SQL Server virtual machines the VHDs for the guest OS are configured with two redundant paths to the Virtual Storage Platform. The server has two dual port host bus adapters (HBAs) installed for high availability purposes.

The Microsoft MPIO software is used for multipathing, employing the round-robin multipathing policy. Microsoft MPIO software’s round-robin load balancing algorithm automatically selects a path by rotating through all available paths, thus balancing the load across all available paths and optimizing IOPS and response time.

Figure 5 shows the Fibre Channel SAN architecture for the SQL Server 2008 R2 implementation described in this reference architecture guide.

Figure 5

Path Configuration

All LUNs are masked and zoned to all four HBAs on each of the Hyper-V servers using four dedicated Fibre Channel ports on the Virtual Storage Platform. Table 4 lists the connections between the Hyper-V server and the storage system ports.

Table 4. Path Configuration

Host HBA and Port	*Switch*	*Zone Name*	*Storage System Port*	*Storage System Host Group*
HBA 1 port 1	Brocade 5100 - 1	HyperV_1_HBA1_1_VSP_1C_3B	1C 3B	HyperV_1_HBA1_1
HBA 1 port 2	Brocade 5100 - 2	HyperV_1_HBA1_2_VSP_2C	2C	HyperV_1_HBA1_2
HBA 2 port 1	Brocade 5100 - 1	HyperV_1_HBA2_1_VSP_5C_4B	5C 4B	HyperV_1_HBA2_1
HBA 2 port 2	Brocade 5100 - 2	HyperV_1_HBA2_2_VSP_6C	6C	HyperV_1_HBA2_2

On the Hyper-V server, the round-robin multipath policy was used on MPIO for all LUNs presented to the host. Note that you can also use Hitachi Dynamic Link Manager software with the round-robin algorithm. Hitachi Dynamic Link Manager Advanced software can provide a central location for management and monitoring of all of your Hyper-V host links to the Virtual Storage Platform.

Storage Building Block

Designing a SQL Server implementation in a virtual server environment is no different than performing the same activities in a non-virtualized environment. Deploying SQL Server 2008 R2 using a building block approach allows you to easily manage and scale your environment. Additional VMs are easy to deploy using System Center Virtual Machine Manager 2008 R2 templates, and storage can be provisioned quickly using dynamic provision pools. While the two reference architectures differ in capacity, they both use the same underlying storage building block. All the LUNs for the environment were created following these best practices:

- For dynamically provisioned environments, place database and log files on separate Dynamic Provisioning pools.
- `tempdb` files can be placed in the same Dynamically Provisioning pools as the database files. However, place them on a separate LUN within the Dynamic Provisioning pool for performance and ease of monitoring. Sharing the same LUN for data files and tempdb can negatively affect performance.
- For the best performance and availability, use RAID-1+0 for the RAID groups for your SQL database and tempdb files. Design for performance first, then capacity.
- When calculating required capacity, always account for approximately 20 percent additional overhead for the database, `tempdb`, and transaction logs LUs. Additional capacity might also be required for your transaction logs depending on your backup recovery model.
- If you increase capacity, use the same RAID group type. This ensures that performance requirements continue to be met.
- Scale SQL Server virtual machines CPU and memory to meet 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 6 shows how an architecture built on a Virtual Storage Platform that starts with a 900GB database and 90GB `tempdb` can scale up while following the building block guidelines.

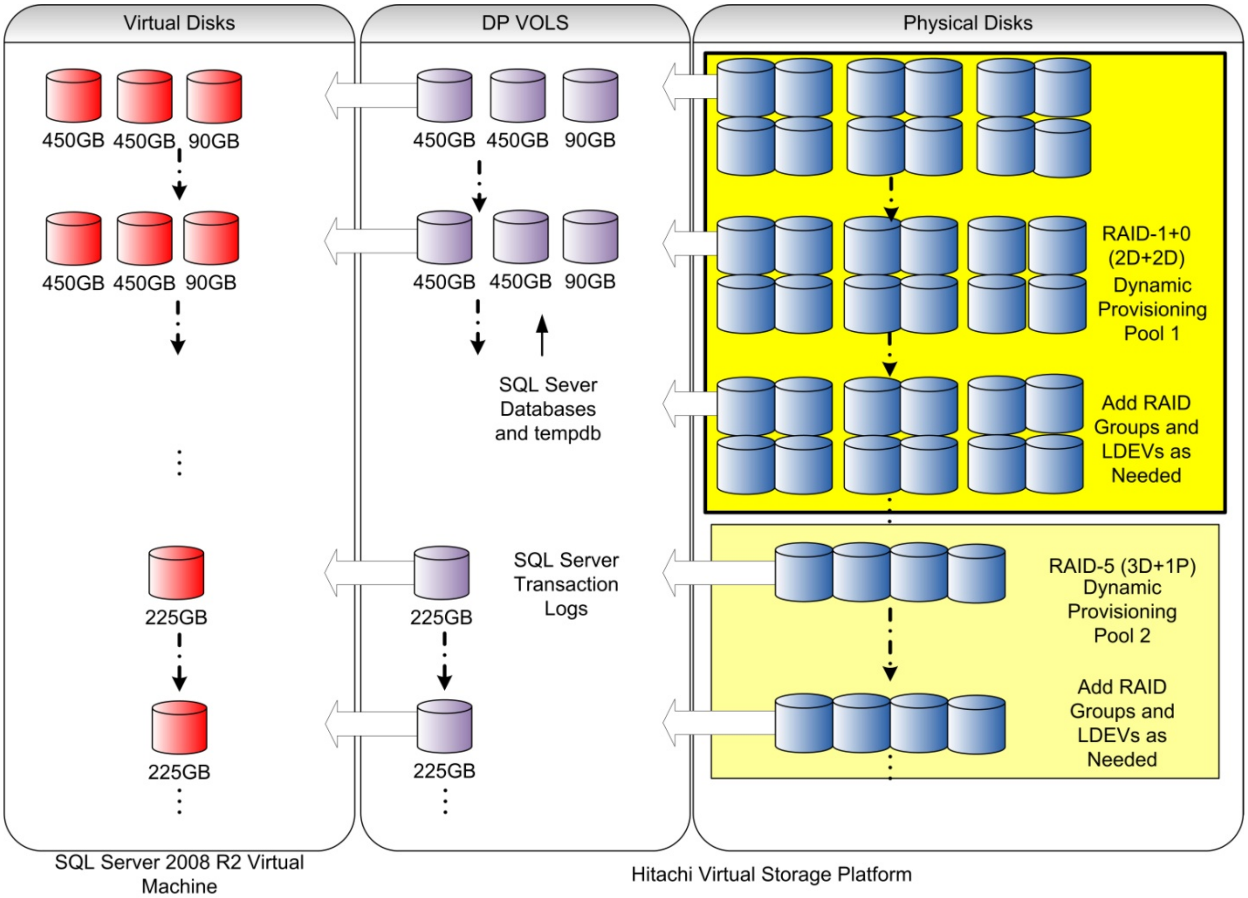

Figure 6

LUNs are presented to the Hyper-V server and made available to the VMs through the passthrough disk capability. Note that when deploying environments that use Microsoft Cluster Services, Hitachi recommends using the LUNs' GUIDs when mapping them to the virtual machines, rather than using the disk number. For more information about best practices and deployment guidelines, see the *Optimizing High Availability and Disaster Recovery with Hitachi Storage Cluster in Microsoft® Virtualized Environments with Live Migration* white paper.

Note that each LUN is sized with at least 20 percent additional capacity as Microsoft recommends. Additional capacity is also provided to ensure that the architecture continues to meet performance requirements for the environment. The Hitachi Dynamic Provisioning software architecture uses eight RAID-1+0 (2D+2D) groups and two RAID-5 (3+1) groups in two separate Dynamic Provisioning pools. One pool is used for the user database and `tempdb` files, the other for the transaction log files. Independent LUNs were created for each file type. 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.

Table 5 describes the 75,000-user building block architecture. 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-1+0 with a 2+2 drive configuration.

Table 5. Building Block Resources for Single Virtual machine

Resource	*Building Block*
Virtual machine	32GB memory 4 x virtual processors
Number of supported users	75,000
LUNs	2 x 450GB DP-VOLs for database files 1 x 90GB DP-VOL for `tempdb` files 1 x 225GB DP-VOL for transaction logs
Disk configuration	8 x RAID-1+0 (2D+2D) with 32 x 300GB 10K RPM SAS drives for databases and `tempdb` 2 x RAID-5 (3D+1P) with 8 x 300GB 10K RPM SAS drives for transaction log

When increasing the number of SQL Server virtual machines in your environment, you can create another pool or you can increase the capacity of the existing 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 Hyper-V server used in this test environment contains 128GB of memory and four quad-core AMD Opteron CPUs. If you require a high availability environment, use a minimum of two Hyper-V servers. You can add SQL Server VMs to the Hyper-V server and allocate more storage resources on the Virtual Storage Platform to scale up to larger databases.

Storage Configuration

This reference architecture uses the Virtual Storage Platform with the Hitachi Dynamic Provisioning software to provision the LUNs used by SQL Server databases, tempdb, and the transaction logs on a 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 Provisioning pool. The RAID group that hosts the VM's OS resides on a separate RAID-5 (3D+1P) Dynamic Provisioning pool.

Hitachi Dynamic Provisioning software distributes the I/O workload across all drives in the Dynamic Provisioning pools. All LUNs for the building block using Hitachi Dynamic Provisioning software are allocated from Dynamic Provisioning pools.

Table 6 lists the LUN configuration for a single building block using Hitachi Dynamic Provisioning software. Note that all storage is in Dynamic Provisioning pools 1 and 2 except for the LUNs used for the guest OS VHD, which are in pool 3.

Table 6. 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 o the existing Dynamic Provisioning pools, whether you're using the single building block, 900GB deployment, or the two building block, 1.8TB environment. Hitachi Data Systems testing shows that this reference architecture meets Microsoft's throughput guidelines as it scales from 900GB to 1.8TB.

If additional SQL workloads are being deployed, you can choose to increase the capacity of the existing pool currently being used, but this should only be done if the performance characteristics of the additional workload and the affect on other workloads in the pool are understood. If an existing pool cannot support additional workloads, an alternative option is to create an additional Dynamic Provisioning 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) as these operations can also affect the overall pool performance bandwidth.

A Dynamic Provisioning pool is made up of one or more logical devices (LDEVs). When increasing your pool capacity, you increase it by adding LDEVs to the pool. When new LDEVs are added, the Dynamic Provisioning pool is unbalanced because the data is not spread evenly across all the spindles in the pool.

Hitachi Dynamic Provisioning software provides a feature called reclaim zero pages that corrects unbalanced conditions by optimizing the Dynamic Provisioning pool. Execute the optimization carefully as it can affect performance depending on how heavily the pool is currently being used. This is due to the nature of the optimization process, which relocates data from by spreading it across all available spindles within a Dynamic Provisioning pool.

Always test your design before deploying it in 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 achieve processor utilization rates of 80 percent in Windows Performance Monitor while maintaining latency levels that meet Microsoft recommendations for avoiding an I/O bottleneck that can affect performance.

A second Virtual Storage Platform might be needed for more complex environments that require failover capability. Hitachi Storage Cluster, failover clustering and Hyper-V quick or live migration allow you to automatically or manually failover SQL Servers or other virtual machines to an additional Hyper-V cluster node and Hitachi Virtual Storage Platform system. Note that live migration is only available on the R2 version of Windows Server 2008.

For more information, see the *Optimizing High Availability and Disaster Recovery with Hitachi Storage Cluster in Microsoft® Virtualized Environments with Live Migration* white paper.

Windows Server 2008 and Hitachi tools enable you to monitor the resource utilization in your environment. Windows Performance Monitor allows you to monitor the CPU and memory utilization of the Hyper-V hosts. You can monitor guest machines with Performance Monitor on the virtual machine's OS. Hitachi Tuning Manager software provides a holistic view of all the Virtual Storage Platform performance related counters. Tuning Manager can help identify potential bottlenecks that might require you to add disks to the environment.

Scaling SQL Server 2008 R2

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. When scaling up, both capacity and performance are concerns; any production environment must be properly tested to ensure it satisfies end-user requirements. You must calculate the maximum number of VMs that a single Hyper-V server can support by noting the total number of CPUs and total memory requirements. Best practice is to not over-allocate the resources available on your physical Hyper-V server. Microsoft recommends deploying a single instance of SQL Server per virtual machine for improved performance and security isolation. Deploy additional virtual machines if you need more SQL Server instances to support your environment. Note that at this point, Hyper-V is limited to 64 VMs per clustered node and 384 VMs per host in standalone environments.

Assign one virtual processor per physical CPU core. If you decide to over-commit your physical processors, frequently monitor the performance of your Hyper-V server. As in a physical server, take both CPU and memory requirements into consideration when planning to deploy SQL Server on Hyper-V. Keep in mind that while you are able to allocate additional memory and virtual processors to a VM, you might need to shut down the VM prior to doing so.

When deploying SQL Server databases, it is important understand of the key file types that are essential for a database. Understanding the type of workload that each file type has, along with the database type and size, enables both storage and database administrators (DBAs) to 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 record 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, the active transaction log, at a minimum, 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 later discovered 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 without delay.

One of the features built into the Virtual Storage Platform is the ability to optimize physical I/O based on recognition of I/O patterns. The Virtual Storage Platform controller can make 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 because the load on `tempdb` is highly dependent upon database and application design. However, if the `tempdb` load is well understood and monitored regularly, Hitachi’s testing shows that it can reside on the same RAID group as the databases without adverse effects. Accordingly, 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. Therefore, you need 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. This is because 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`, you must restart 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 which indicate 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 based on the performance levels that a SQL Server VM can handle. This is no different than how you configure SQL Server storage for physical server deployments. The architectures described in this section require the following calculations be derived from the requirements for an existing or planned 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 Virtual Storage Platform with Hitachi Dynamic Provisioning software for best I/O throughput and ease of management.
- 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 these:

- 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 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 by adding RAID-1+0 groups for database and tempdb files, and RAID-5 groups for the transaction logs. The specific RAID group type depends on the initial building block environment that meets both performance and capacity requirements for your environment.
- In a dynamically provisioned configuration, you can allocate Dynamic Provisioning pool space by adding an LDEV from a RAID-1+0 group to the Dynamic Provisioning pool. After that, you can create the required LUNs and assign them to a Hyper-V server.

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 and then an additional SQL Server VM and storage building block 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 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 7 lists the OLTP workload parameters used as a baseline for the tests.

Table 7. 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 8 lists results for the OLTP workload tests for the 75,000-user configuration.

Table 8. OLTP Workload Test Results for 75,000-user Configuration

Metric	Microsoft Success Criterion	Result
Achieved Database IOPS	Varies	7,748 IOPS
Database Avg. Disk sec/Read	≤ 20ms	14ms
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 9 lists test results for the OLTP workload tests for the 150,000-user configuration.

Table 9. OLTP Workload Test Results for 150,000-user Configuration

Metric	*Microsoft Success Criterion*	*Result*
Achieved Database IOPS	Varies	15,426 IOPS
Database Avg. Disk sec/Read	≤ 20ms	17ms
Database Avg. Disk sec/Write	≤ 20ms	2ms
Transaction Log Avg. Disk sec/Read	≤ 5ms	N/A
Transaction Log Avg. Disk sec/Write	≤ 5ms	2ms

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

Conclusion

This reference architecture guide describes a building block-based storage solution for virtualized 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 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 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.

For more information about Hitachi products, 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