Hitachi Universal Storage Platform™ V and Universal Storage Platform VM

Cost Effective Tiered Storage Options

Design Guide

By John Harker

Executive Summary

Designing storage systems to encompass multiple service levels and price points delivers superior storage service levels for less cost. This paper introduces and explains a design methodology for developing tiered storage architectures. It then discusses the options for configuring the Hitachi Universal Storage Platform™ system tiers using available software facilities, disk drive types and RAID group levels to meet performance, reliability and cost goals.

For best results use Acrobat Reader 8.0.

Contents

Hitachi Universal Storage Platform™ V and Universal Storage Platform VM

Cost Effective Tiered Storage Options

Design Guide

By John Harker

Introduction

Designing storage systems to encompass multiple storage service levels and price points delivers superior storage service for less cost. That is what tiered storage systems are about. This paper introduces a design methodology for developing tiered storage systems and explains and discusses best practices for drive selection and configuration for the major options. It also shows relative pricing for variously configured versions of a reference system.

This paper was designed as an introduction to the topic for use by customers, Hitachi TrueNorth Channel Partners and Hitachi Data Systems employees who specify, design or configure Hitachi Universal Storage Platform™ models.

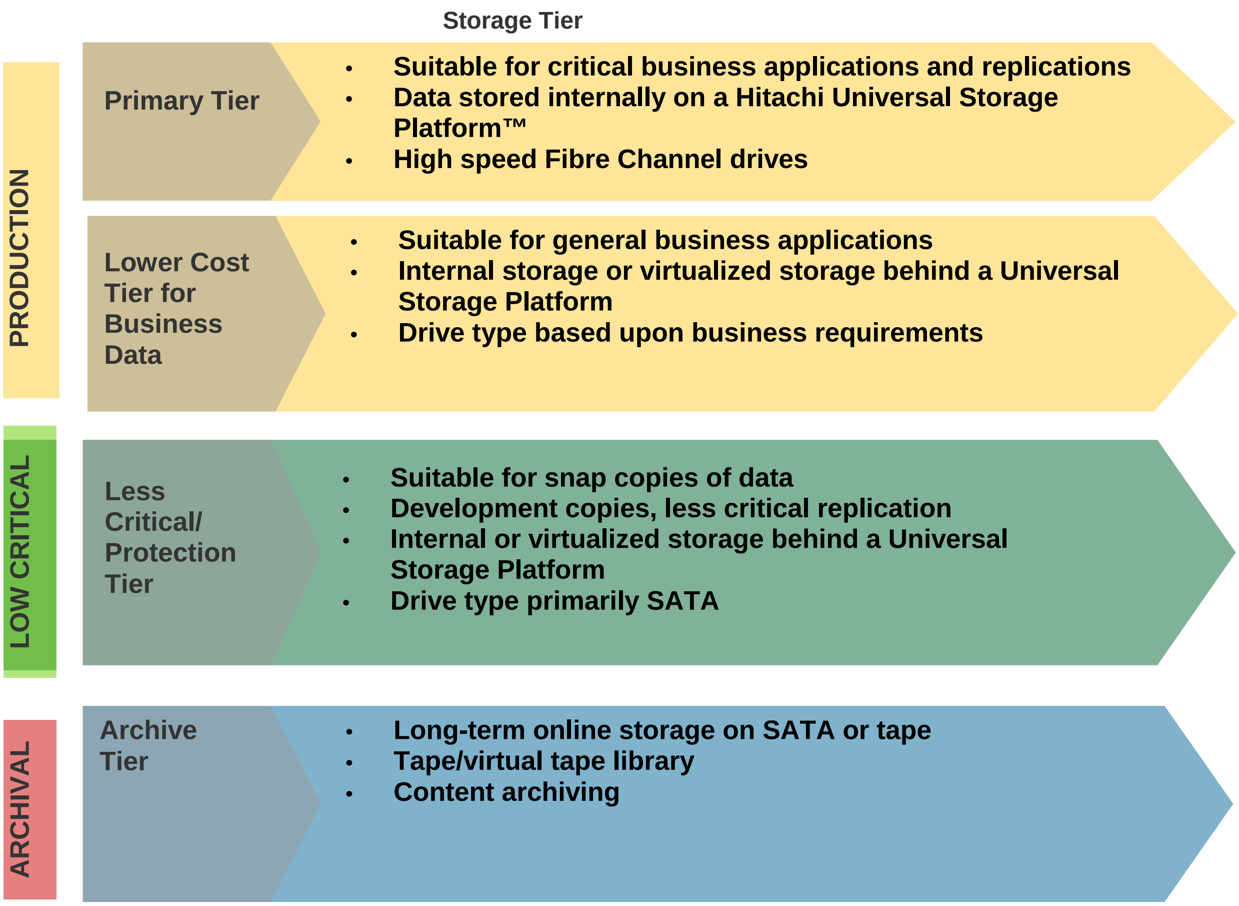
Tiered Storage Design

It Starts with the Applications

The design of a virtualized tiered storage system starts with the applications. It is the business needs and applications that drive the storage requirements which guide tier configuration. Most applications can benefit from a mix of storage service levels, using high performance where it is important and less expensive storage where it is not.

But operationally it is not efficient to configure unique tiers for each application. Individually configuring a unique scheme for each application leads to extra work, cost and provisioning delays. Instead, the recommended practice is to develop a catalog of pre-defined tiers with pre-defined characteristics and then allocate storage to applications as needed out of the catalog. Figure 1 outlines a four tier model; your individual requirement may call for more or less.

Figure 1. Sample Pre-defined Model for Storage Tiers



Aligning Storage Tiers with Applications

In designing storage tiers you should start by looking at your application requirements in a variety of areas. The goal is to align each application to the right tier of storage with correct storage characteristics. What are “storage characteristics”? There are a number of characteristics of storage, each of which is a topic unto itself. As a starting point, individual consideration needs to be given to:

- Availability — How important is it that the data always be available? A retail Web storefront might need guaranteed 100 percent uptime, while an archival system might only be needed during business hours.
- Performance — What is the I/O response time needed? How much bandwidth?
- Cost — Drive types and RAID configuration options offer a wide range of price points for different service levels.
- Protection and recoverability — What are the recovery time objective (RTO) and recovery point objective (RPO) requirements? Is a simple nightly backup acceptable? Is replication necessary for backup window timing or business continuity? How many replication copies are needed? Is distance replication needed?
- Retention and compliance — How long does the data need to be kept on this tier? Note that if you back up and archive to a lower tier you do not need elaborate retention controls on higher tiers.

It is useful to develop a specific map of the individual characteristics of each attribute for every tier. See Figure 2. These will be the storage service characteristics an application gets when assigned storage from that tier. You need to determine what each of these will be. Consider your current set of applications; you need to develop a single set of characteristics for each tier such that from the total set of tiers you can satisfy the storage service level requirements of all of the applications.

One way is to develop this in an iterative fashion for each of the starting set of application requirements in each area, devising a starting set of numbers/characteristics for each tier, then examining the next application and optimizing for both; do the same for the next application, until arriving back to the first application, which starts the cycle over again until converging on a design that meets the requirements. This is shown in the black box overlay of Attributes by Tier on Figure 2. During this process you may discover that you need more or fewer tiers.

Figure 2. Sample Map of Individual Characteristics of Each Attribute for Every Tier

Desired Attributes by Tier

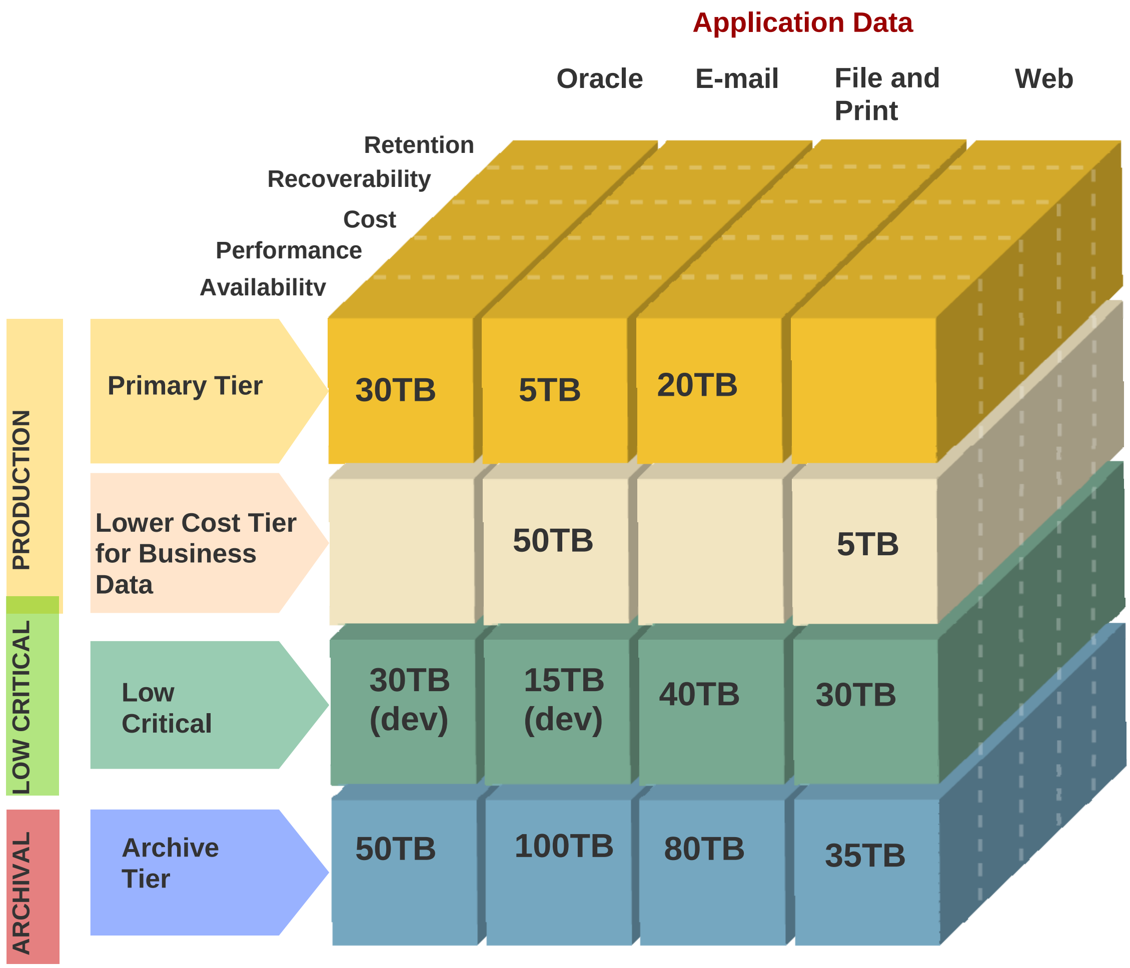
Data Tier		Availability	Performance	Cost (relative)	Recoverability	Retention
PRODUCTION	Primary Tier	99.999%	5	$$$$$	A	6 months
PRODUCTION	Lower Cost Tier for Business Data	99.99%	4	$$$$	A	6 months
LOW CRITICAL	Less Critical	99.9%	3	$$	B	1 year
ARCHIVAL	Archive Tier	99.9%	1	$	C	7–100 years

It is worth noting that in the example above, for purposes of illustrating the layered design effort, tier characteristics such as “Performance,” “Recoverability” and “Retention” are simplified; in actual design each of these areas in turn faces a number of design options that aren’t fully captured with the linear values specified.

The last step is to map in storage for each application from the appropriate tiers as needed. Figure 3 provides an example of how this might look for a variety of applications.

For the company's mission critical Oracle business applications, primary tier storage is used. However, the lower priority development copies are run on a more cost efficient less critical tier. The Oracle databases are backed up regularly to the Archive tier.

Figure 3. Sample Map of Applications and Appropriate Storage Tiers



- The Microsoft® Exchange® e-mail application needs to keep its log files on primary storage for performance, but the bulk of the storage for the mailboxes themselves can be mapped to the less expensive but still performing lower cost business tier. A small amount of storage space is also mapped in from a less critical tier for development purposes. With stringent retention policies and an expanding amount of e-mails with large attachments, a large amount of archive tier storage is needed.
- The NAS Head File and Print functions need some primary tier storage for several critical image processing applications, but the bulk is file sharing used for shared directories within the company and print spooling, and can use inexpensive low critical tier.
- Additionally, the company's Web server uses the lower cost business data tier for the core set of often accessed pages, but the bulk of what is online which is infrequently accessed can be kept on less critical storage.

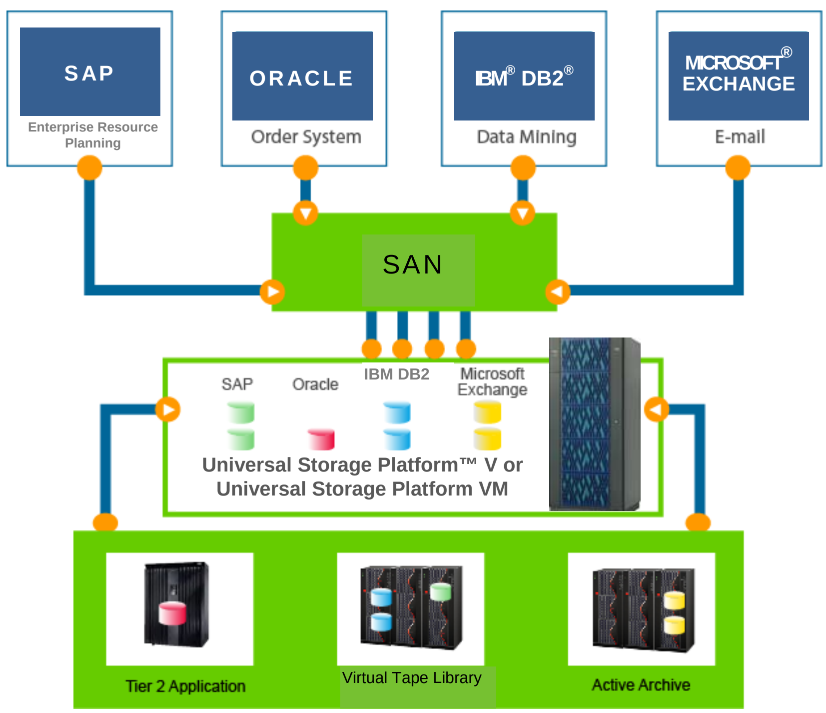
Storage Options

Now that we have designed our tiers from a requirements standpoint, how do you configure a system to match? There are a variety of ways to configure tiered storage architectures. You can dedicate specific storage systems for each tier, or you can use different types of storage within a system for an "in-the-box" tiered storage system. The Hitachi best practice is to use the virtualization capabilities of the Universal Storage Platform to eliminate the inflexible nature of dedicated tiered storage silos and seamlessly combine both, allowing for the best overall solution possible.

For example, for the highest tier you could start with a Hitachi Universal Storage Platform V or Universal Storage Platform VM configured with Fibre Channel drives and a high-performance RAID configuration. Here the highest levels of performance and availability for mission critical applications are required. As a second tier you could add Fibre Channel drives to the Universal Storage Platform that are configured at a RAID level that is more cost effective and still highly reliable but with a little less performance. As a third tier you could add even more cost effective internal or external SATA drives.

The Hitachi storage virtualization architecture is differentiated by the way in which Hitachi storage virtualization maps its existing set of proven storage controller-based services such as replication and migration across all participating heterogeneous storage systems (see Figure 4).

Figure 4. Hitachi Virtualized Tiered Storage Solution Example



Major Hitachi Data Systems customers such as Fidelity National and HDFC banks realize multiple benefits from a Hitachi virtualized tiered solution, including operational efficiencies in change management, simplified storage management, a streamlined migration process and an extension of the life of storage assets.

Most customer environments will contain a mix of workload and applications within a single system so it is important to design a system where I/O load or activities (such as formatting, shredding, replicating or parity rebuilding) on a tier do not disrupt or degrade other tiers. In particular, it is important that a lower tier load does not affect higher tier performance. Hitachi enables this to be done safely on one platform through the extremely high internal bandwidth of the Universal Storage Platform's advanced crossbar switch architecture and advanced performance and balancing tools, such as those that provide the ability to create independent storage or cache partitions and to control I/O bandwidth at the virtual port level.

In looking at performance requirements for designing or selecting a tier it is also important to look at the application's use of that tier's storage in terms of read and write characteristics. How much is sequential read? How much random read? How much sequential write? How much random write? The mix of read/write characteristics is a major determining factor in the storage performance it will see. As an illustration of why, consider an I/O's data path inside the Universal Storage Platform:

- Random read hits are serviced at electronic speed by cache and do not reach the back end.
- Random read misses go through cache unaltered and go straight to the appropriate back end disk drive.
  - This is the only type of "I/O" operation where the host always "sees" the performance of the back end disk drive.
- Random writes
  - Host sees random writes complete at electronic speed; it only sees delay if too many pending writes build up.
  - Each host random write is transformed going through cache into a multiple I/O pattern that depends on RAID type.
- Sequential I/O reads and writes
  - Host sequential I/O is at electronic speed. Cache acts like a "holding tank" and the storage system back end puts [removes] "back end buckets" of data into [out of] the tank to keep the tank at an appropriate level.

So high-performance drive type and RAID configurations will have a much bigger impact on a random read and write intensive application, versus one that does mostly sequential reads and writes. And proper cache sizing, I/O queue depth settings and cache partitioning can be equally important as drive type and configuration in tiered storage designs.

There are also a number of tiered storage design tools useful for configuring and managing storage tiers on the Universal Storage Platform. These are software products, interfacing to and supported by the platform's firmware, that provide enhanced levels of functionality which are used to optimize and tune all drive and RAID combinations.

The following sections describe several areas of options for configuring tiered storage systems, starting with available software-based capabilities and going on to RAID types and physical drive options.

As with all storage systems, when sizing it is necessary to keep in mind not only the amount of storage needed, but also the application's I/O characteristics and performance requirements. The final drive count may wind up being driven more by performance requirements than required capacity. Proper drive choices as well as the right configuration will provide the highest performance and will create a more cost effective tiered storage environment; the Universal Storage Platform V and Universal Storage Platform VM provide a number of significant value added capabilities in this regard.

Architecture Related Software

Hitachi Basic Operating System V and External Storage

One of the unique abilities of the Universal Storage Platform is to be able to connect and virtualize externally attached storage systems, creating a single pool of tiered storage. Enabled with Hitachi Basic Operating System V software, this ability brings the capabilities of the Universal Storage Platform to bear on a multitiered storage area network of heterogeneous storage systems. It allows single point access from the Universal Storage Platform to other old and new storage systems in the SAN.

This ability allows the operation of multiple storage systems connected to a Hitachi Universal Storage Platform as if they were all one system and with a common management toolset and software capabilities. Using the Basic Operating System V the shared storage pool composed of both internal and external storage volumes can be used with all Hitachi storage system–based software for data migration and replication, as well as any host-based application.

It also provides the system designer with the option to reuse existing storage and flexibility and scalability for storage expansion. One of the design factors involving the use of SATA drives is how much growth is expected. In general, external storage becomes much less expensive than internal if there is going to be a lot of it.

Hitachi Dynamic Provisioning Software

Hitachi Dynamic Provisioning software is an important option to consider when developing your tiered architecture. Hitachi Dynamic Provisioning is a thin provisioning product and is part of the Hitachi Storage Command Suite. It permits the complete virtualization of a volume provisioned to an application. Actual storage capacity is assigned on an on-demand basis when data is written by the application. Actual storage capacity is collected in one or more Hitachi Dynamic Provisioning physical storage pools. The pools are created from multiple LDEVs from multiple drive groups of any RAID level. Each Dynamic Provisioning pool supports virtual host accessible volumes referred to as HDP Volumes (short form: DPVOLs). From an application's standpoint, an HDP Volume looks no different than any other storage volume. As applications write data to HDP Volumes, Hitachi Dynamic Provisioning software assigns actual capacity from the HDP Pool to the HDP Volumes. This "just in time" method means physical storage allocations remain available until they are actually needed.

Dynamic Provisioning simplifies storage management by decoupling the provisioning of capacity to an application from the management of physical resources in the storage system. For example, physical storage is nondisruptively added as needed to the storage system, placed in centralized pools and made available to new and existing HDP volumes.

Dynamic Provisioning software also simplifies performance optimization by transparently spreading many hosts' individual I/O patterns across many physical disks. You can create pools with the RAID level desired and use as many disks as needed to support the desired target IOPS rates for the DPVOLs connected to that pool.

For example, you might create four database pools, with 16 LDEVs taken from 8 RAID-1+0 groups (two LDEVs each), and use these for the database files that contain user data. These database pools are configured to support high levels of random small block OLTP workloads.

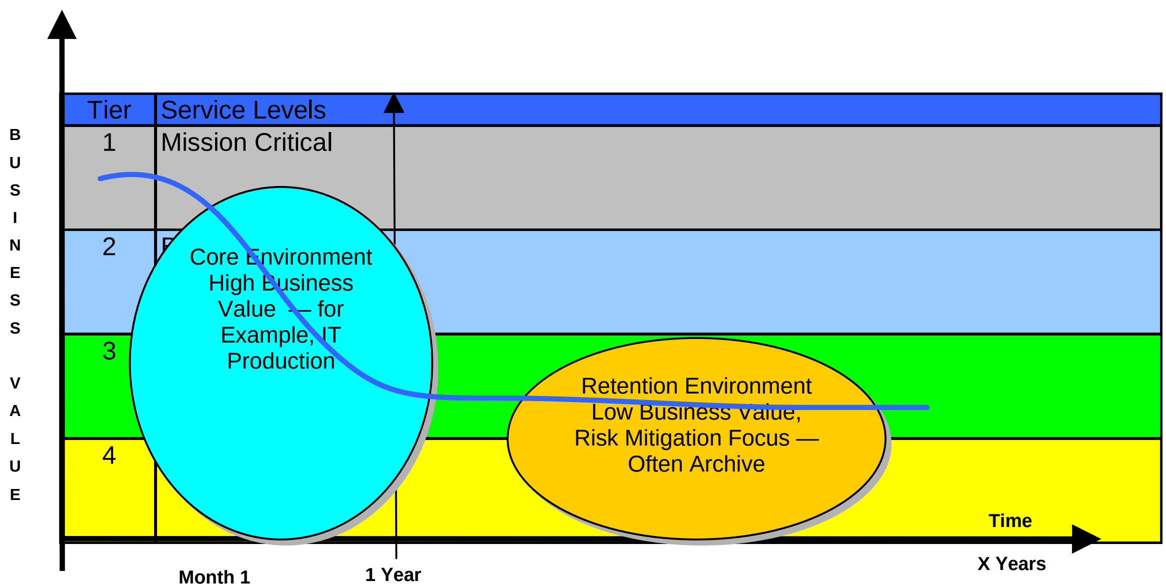
You might then create a log pool of eight LDEVs from four RAID-5 groups (two LDEVs each) and use this pool for DPVOLs assigned to database logs or temp space. This log pool is configured to accept high levels of large block sequential write workloads.

Hitachi Tiered Storage Manager Software

Hitachi Tiered Storage Manager software is an optional Hitachi Storage Command Suite module that provides easy tier definition and dynamic data movment capabilties for all storage in the Universal Storage Platform pool — both internal and external. It enables administrators to easily and interactively identify available storage to match application quality of service (QoS) requirements and then define storage tiers based on this information. As such, it is an excellent tier management tool. It also then provides the ability to migrate or schedule migrations of online data within or between tiers without disrupting application access.

It is important to be able to nondisruptively move a running application's data between tiers in a tiered storage environment (see Figure 5). Information value changes over time, driving content movement requirements across service levels to realize an optimal cost balance.

Figure 5. Data Lifecycle of a Transaction



Tiered Storage Manager is an also important option for management flexibility. The ability to nondisruptively migrate the data to a different tier, drive or RAID type is important for rapid response to new requirements, rapid deployments or data migrations. Storage can be allocated to an application, observed and then nondisruptively changed if needed.

Tiered Storage Manager is a useful tool for data migrations off of old equipment, which can be virtualized behind a Universal Storage Platform and the stored data nondisruptively migrated to another system or to storage internal to the Universal Storage Platform.

Hitachi Virtual Partition Manager Software

Hitachi Virtual Partition Manager software, included in Hitachi Basic Operating System software provides the ability to divide storage, ports and cache on the Universal Storage Platform V into four (with Basic Operating System) or 32 (with Basic Operating System V) partitions. For the Universal Storage Platform VM, Basic Operating System V enables up to eight partitions. This enables you to limit access to data and resources from individual servers or applications to the elements of a single partition and prevents consumption of resources from outside of that partition. The partitions are dynamically configurable and can be changed without bringing the storage system down. This capability is useful for various purposes, such as restricting cache-intensive applications to a cache partition so they don't affect the performance and QoS of other applications.

For example, a cache partition can be set up for a set of SATA drives so that the slower write speeds do not cause excessive cache consumption and affect the performance of other applications using the system. Alternatively, an application that has a high cache hit ratio can be assigned a fixed amount of cache to ensure resources are available for other parallel usage.

Hitachi Server Priority Manager Software Feature

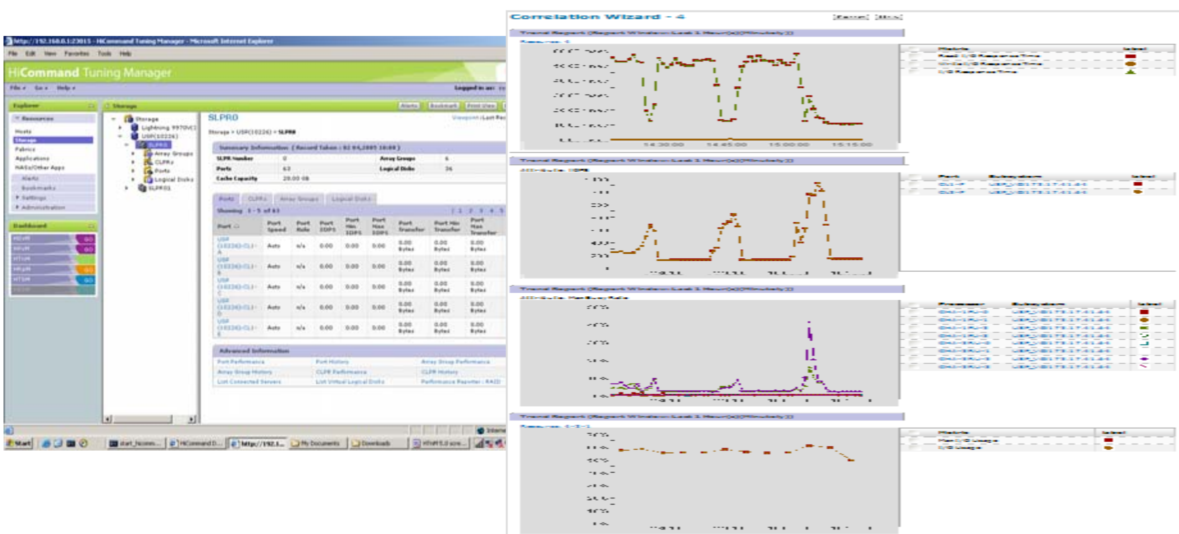
Another capability of the Universal Storage Platform using the Basic Operating System Server Priority Manager feature, is to limit I/O requests (I/O Rate or IOPS) from lower priority applications or servers. By limiting throughput (transfer rate) you can ensure that activity on lower tier storage does not affect higher tiers. You can prioritize by port or by server (WorldWide Names) and thresholds can be set so that if the I/O of the "high priority" application drops below a certain level, the noncritical applications can use the available resources.

For example, if an important transaction processing application resides on the same system as an archiving system, the volumes assigned to the archiving system can be throughput limited to prevent disk conflicts during periods of write-intensive activity.

Hitachi Tuning Manager Software

Hitachi Tuning Manager software is an optional Hitachi Storage Command Suite module that can deliver real time and historical performance and capacity information useful in defining and monitoring individual storage tier QoS to administrators and other products such as Tiered Storage Manager (see Figure 6). Tuning Manager can also feed real time performance data to Tiered Storage Manager for improved tier definition and management.

Figure 6. Tuning Manager Display of Real Time and Historical Performance Information



RAID Types

The industry and Hitachi Data Systems best practice is to always use some form of RAID configuration beyond RAID-0 as a first line of defense in the data protection war. In a tiered storage environment you can take advantage of different (multiple) RAID configurations to provide varying levels of service and capacity optimization according to individual tier requirements.

There is a price for using a RAID configuration both in the storage capacity used and performance. This price varies according to the type of RAID used. For example, with RAID-1, whenever a host writes data to the storage system, at least two disks need to be updated. The amount of extra disk drive I/O activity needed to handle write activity is the key factor in determining the lowest cost solution combination of disk drive RPM, disk drive capacity and RAID type.

Particularly for random I/O, back end disk I/Os are one of the finite performance factors — a drive head can only go "so fast." For example, as shown in Table 1, take a RAID-5 LUN and a RAID-1+0 LUN, issue four I/Os, and observe how many back end disk I/Os this generates:

Table 1. Back End Disk I/Os Example with RAID-5 and RAID-1+0

RAID-5	***Total number of Disk I/Os***
Random Read	4
Random Write	16
Sequential Read	4
Sequential Write	5
RAID-1+0	
Random Read	4
Random Write	8
Sequential Read	4
Sequential Write	8

Hitachi Universal Storage Platforms are usually configured with a mix of three types of RAID levels, as examined below.

RAID-1 and RAID-1+0 (Mirroring)

RAID-1 mirrors the entire disk. RAID-1 configurations have higher performance in most applications because each disk has a copy of the data. Thus, it is possible for the system to be responding to two I/O requests at the same time. In general, a mirrored configuration will perform as much as 30 percent better than a parity protected (RAID-5) configuration.

It is a Hitachi Data Systems best practice to use RAID-10, or more accurately RAID-1+0, instead of RAID-1. RAID-1+0 is a variation on RAID-1 that adds increased performance through added striping. Considered difficult for a storage system to implement, Hitachi provides an industry leading RAID-1+0 implementation.

Mirrored configurations also provide higher performance if there is a disk failure since another complete copy of the data is immediately available. They also offer the shortest rebuild time for the failed drive.

- Since RAID-1 and RAID-1+0 require doubling the number of disk drives to store the data, people tend to think of RAID-1 as the most expensive type of RAID.
- However, RAID-1 and RAID-1+0 offer the lowest "RAID penalty" of only having two disk I/Os per random write, compared to four I/Os for RAID-5 and six for RAID-6.
- And there is no penalty and a potential improvement for read operations.

For this reason, when the workload is sufficiently active and has a lot of random (as opposed to sequential) writes, RAID-1 or RAID-1+0 will often be the cheapest RAID type because it has the least disk drive I/O operations per random write. Other RAID configurations will have to be used at low densities to meet the random write requirements.

RAID-5 (parity protected)

RAID-5 configurations are parity protected. In the event of a physical drive failure, the missing data is rebuilt by reading the remaining drives in the RAID group and performing parity calculations.

RAID-5 achieves redundancy with less disk space overhead, but at the expense of having a higher "RAID penalty" for random writes and having a larger performance degradation upon a drive failure (while the failed drive is rebuilt from the parity data).

- For sequential reads and writes, RAID-5 is very good.
  - It is very space efficient (smallest space for parity), and sequential reads and writes are efficient, since they operate on whole stripes.
- For low access density (light activity), RAID-5 is very good.
  - The 4x RAID-5 write penalty is (nearly) invisible to the host, because it is non-synchronous.
- For workloads with higher access density and more random writes, RAID-5 can be throughput-limited due to all the extra parity group I/O operations to handle the RAID-5 "write penalty."

RAID-6 (double parity)

RAID-6 adds on to the single error correcting capabilities of RAID-5 with two types of parity correction that can continue running and rebuild the data even if two drives fail. It has generally the same read/write characteristics as RAID-5. It provides the highest level of data availability, albeit with more of an impact on write performance than RAID-5 due to increased overhead from the double error correcting capability.

The extra protection of RAID-6 is particularly well suited for SATA drives that have both lower reliability numbers and, due in part to their size, longer rebuild times. Because of this RAID-6 is considered a best practice for SATA configurations.

A Note on Parity Groups

A RAID configuration of drives involves both storage for data and storage for extra information. This extra information is referred to as “parity.” For each of the RAID types there are different ways of laying this out; for example, for the Universal Storage Platform the following are available:

RAID-1+0: 2D+2P (2+2)

RAID-5: 3D+1P (3+1), 7D+1P (7+1)

RAID-6: 6D+2P

The Universal Storage Platform also offers the ability to create “concatenated parity groups,” which are parity groups where the striping extends across multiple coupled physical parity groups. This extended striping brings more drives (spindles) to bear and improves performance.

RAID-1: 2 x (2+2), often called 4+4

RAID-5: 2 x (7+1) and 4 x (7+1)

So you can see that for each type of fixed drive type, using different RAID and parity group types, a number of different configurations are possible, each offering different performance and reliability levels. Now let’s look at the drive types.

Drive Types

The Hitachi Universal Storage Platform V and Hitachi Universal Storage Platform VM are available with various hard disk drive types and sizes varying by performance, capacity and reliability. Fibre Channel drives are available in various sizes and rotational speeds. High capacity SATA drives are available for lesser tier applications, such as low speed replication and archiving.

In understanding the performance impact of various drive options, it is important to understand that higher speed drives perform better than lower speed drives in terms of seek time and rotational latency. Drives of the same rotational speed, latency and seek time will have roughly the same performance, regardless of capacity.

The Universal Storage Platform V supports a mixture of up to 1,152 Fibre Channel and/or SATA internal disks and the Universal Storage Platform VM supports a mixture of up to 240 Fibre Channel and/or SATA internal disks. Table 2 illustrates the industry standard specified capacity (base-10), the actual (computer base-2) raw size and the typical usable (formatted) size of each type of disk. Also shown is the “rule of thumb” average maximum random IOPS rate for each type of disk when using most of the surface. The number and size of the LDEVs created per RAID group determine how much of its disks’ surfaces are in active use.

Table 2. Capacity and Size of Supported Fibre Channel and SATA Internal Disks

Hard Disk Drive Type	Industry Standard Capacity (GB)	Actual Capacity (GB)	Typical Formatted Capacity (GB)	Typical Physical Maximum IOPS
1TB 7,200 RPM SATA	1000	923	840	80
750GB 7,200 RPM SATA	750	738.62	701.7	80
400GB Fibre Channel 10K RPM	400	384	365	130
300GB Fibre Channel 15K RPM	300	288.2	273.8	180
146GB Fibre Channel 15K RPM	146	143.76	136.6	180
72GB Fibre Channel 15K RPM	72	71.5	67.9	180

Note that as the percentage of active data on the drive increases, the effective I/O rate per volume of data decreases. This effective I/O rate is referred to as ***access density***, or IOPS/GB (IOs per second per gigabyte). It decreases because the more of the disk surface that is active, the farther the disk heads must seek. This creates higher average seek times. For example (using a 15K RPM disk), when only using the outer 25 percent of the surface the average read seek time may typically be about 1.8ms (providing about 267 IOPS), whereas the 100 percent surface average seek time will be about 3.8ms (about 182 IOPS). For a SATA disk at 7,200 RPM, these values are about 4.3ms (119 IOPS) and 8.5ms (79 IOPS).

So it is important to consider not only the I/O characteristics you are optimizing for, but also the access density:

- The lowest cost configuration in terms of RPM, capacity and RAID type depends strongly on the tier's access density and read/write characteristics.
- If there is moderate or higher access density with significant random write activity, RAID-1+0 and/or 15K RPM drives will often turn out to be the lowest cost total solution, as they can fill up more of the drives' capacity with data without sacrificing performance.
- Where access densities are low and reads/writes are mostly sequential, use SATA or larger 10K RPM Fibre Channel drives.

Another important design factor in selecting the size of the drive is recoverability. How long will it take to rebuild the physical drive from a spare if the physical drive fails? The answer is, for a given RAID type, the bigger it is the longer it will take. Since a second hard failure occurring while a rebuild is happening could cause a hard failure, this is an exposure. This is one of the factors influencing the RAID configuration selected — options such as RAID-1+0 offer rapid recovery while RAID-6 offers an extra level of security. So consider disk recovery time in system design.

SATA drives offer a less expensive design option for lower tiers that are less performance critical. With Hitachi Data Systems proprietary enhancements and proper RAID configuration their reliability can be brought in line with Fibre Channel drives. In general, the differences between the SATA and Fibre Channel disks include:

- The amount of technology and precision involved with the materials used and the degree of error detection and correction
- The expected lifespan with the duty cycles experienced in enterprise class storage (but note for lower tiers where duty cycles are low this isn't an issue)
- The overall performance available

Table 3 gives a view of the relative performance of SATA and Fibre Channel drives.

Table 3. Relative Performance of Fibre Channel and SATA Drives

		750GB SATA [7,200 RPM]	***146GB Fibre Channel (Reference) [15,000 RPM]***
Read/Write Performance	***Random Read (8KB)***	800 IOPS	1600 IOPS
	Random Write (8KB)	110 IOPS	440 IOPS
	Sequential Read (512KB)	190MB/sec	380MB/sec
	Sequential Write (512KB)	60MB/sec	240MB/sec
LDEV Format Time (No I/O load)		12 hours	40 minutes
Copy Time (No I/O load)	***Correction Copy***	10 hours	2 hours 15 minutes
	Drive Copy	10 hours	55 minutes

Currently, Hitachi Data Systems assures reliability with SATA drives through several technology enhancements such as read-after-write, automatic head parking and alignment, and increased sparing. In the near future, Hitachi Data Systems will be announcing further technical improvements that will make the current read-after-write function unnecessary.

The following sections describe the performance characteristics of each drive type and give some use case examples and configuration recommendations.

Fibre Channel

Fibre Channel drives offer the highest levels of performance and reliability at a higher cost than SATA. Any standard RAID configuration can be used with all of the drives below.

146GB Fibre Channel, 15K RPM (4Gbit) — general purpose high-performance drive; suitable for high-performance applications where moderate read/write requirements are driving the sizing requirements

300GB Fibre Channel, 15K (4Gbit) — general purpose high-performance medium capacity Fibre Channel drive

450GB Fibre Channel, 15K (4Gbit) — general purpose high-performance large capacity Fibre Channel drive

400GB Fibre Channel, 10K RPM (4Gbit) — midrange performance high capacity Fibre Channel drive; lower access density means less suited for random read/write intensive applications

SATA

1TB SATA, 7,200 RPM (3Gbit, with native command queuing)

This is an inexpensive high capacity drive. Hitachi adapts the slower 3Gb/sec speeds of the SATA drive interface to the faster 4Gb/sec Fibre Channel connection, so maximum performance is achieved.

The general best practice is to configure SATA drives as RAID-6. The Hitachi position is that with our proprietary SATA enhancements and RAID-6 configuration that SATA is suitable for enterprise applications for the following purposes:

- E-mail archives, backups, tape replacement
- Second or lower tier NAS file and print
- Second and third stage replication copies
- Second or lower tier primary/secondary replication pairs
- Historical data with limited access (medical records, etc.)
- Archived documents for regulatory purposes
- Sound or video files

When making performance estimates for SATA groups that use RAID-6, expect as a rule of thumb that reads/write performance will be 50 percent to 75 percent lower than with similar RAID-6 146GB 15K RPM drive groups. Because of this, SATA drives are best suited for sequential read/write intensive applications where the performance you get from them will be closely related to the amount of cache in the data path.

Because the performance is lower it is important to closely estimate the I/O rates required and properly configure the system — not just the RAID types and sizes, but also the layout of the drives in the system.

SATA drives can be used in replication pairs with some limits. They can be used for lower tier replication pairs for applications where less performance is needed. Examples would be in SATA-SATA configurations, or for read intensive applications that generate few writes in a Fibre Channel-SATA pair combination. They are not well suited for high-performance random write intensive application.

It is the Hitachi Data Systems best practice to configure a CLPR (Cache Logical Partition) for the SATA disk drives (or, for externally attached SATA to turn off Universal Storage Platform caching) to restrict cache contention and a potential performance impact on higher tiers.

The decision on how to use SATA should be carefully considered since there are two options — use SATA drives internal to the Universal Storage Platform V or Universal Storage Platform VM or use externally attached SATA storage. Which is best often depends how much SATA storage you are configuring. Generally, this works out to be that the more SATA storage there is, the more external systems are favored. For smaller amounts of SATA it is more likely to be more economic to configure it on the Universal Storage Platform V or Universal Storage Platform VM itself. For more information, refer to some of the pricing comparisons below.

Sample Tiered Storage Configuration Cost Comparisons

Pricing is a key factor in determining how to physically provision your storage system. With the Universal Storage Platform V and Universal Storage Platform VM, users have the option of multiple types of internal and external physical storage. How do the prices for various layouts of the same net storage add up? There is no simple answer to the question because the price is related to the current configuration and how much or little needs to be added to the system beyond the new drives.

Because the pricing is complex our strategy is to select and show a representative cross set of configurations. Pricing is given in relative terms. To arrive at an actual price for a system, work with your Hitachi Data Systems sales representative. Here we compare relatively simple configurations (but representative of what customers purchase) at different storage capacities and internal/external ratios (see Figure 7).

The scenarios are:

Customer Scenario 1: "New." Customer will acquire 50TB of new capacity.

Customer Scenario 2: "25TB Just Disks Upgrade." Customer already has a 25TB Universal Storage Platform and will upgrade his current capacity with 25TB using existing drive slots and controllers.

Customer Scenario 3: "25TB Complex System Upgrade." Customer already has a 25TB Universal Storage Platform and will upgrade his current capacity with 25TB but needs to add drive bays and controllers.

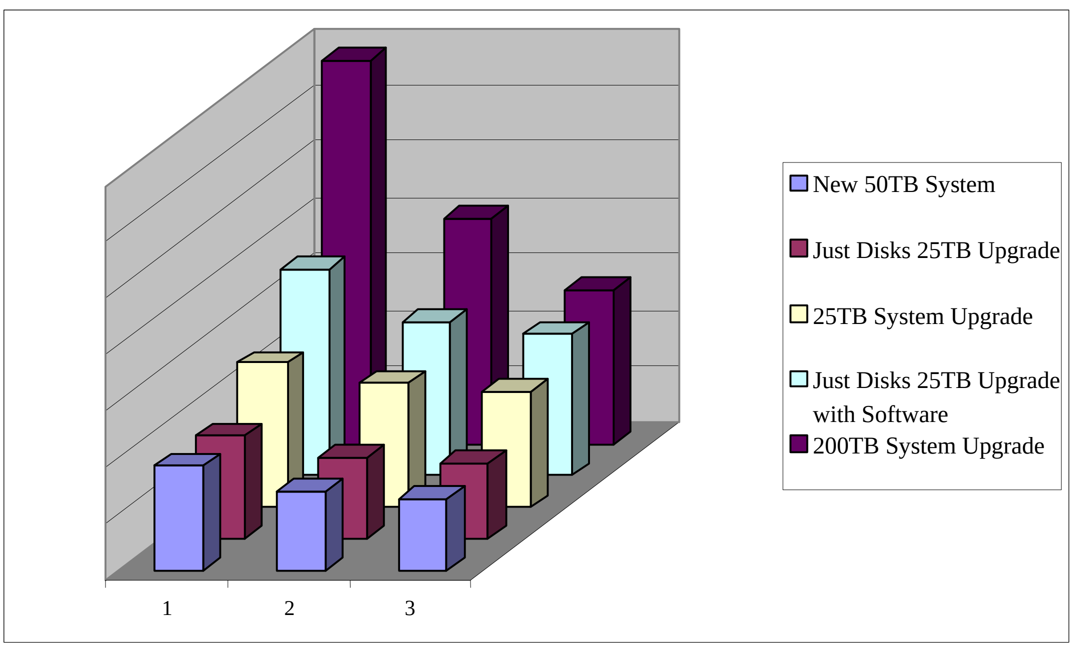
Customer Scenario 4: "25TB Upgrade with Software." Same as scenario 2 with addition of Hitachi Dynamic Provisioning, Hitachi Tuning Manager, and Hitachi Tiered Storage Manager software, as well as Hitachi In-System Replication software bundle added along with matching services.

Customer Scenario 5 "200TB Upgrade." Customer already has a 25TB Universal Storage Platform and will upgrade his current capacity with and additional 175TB.

For each of the scenarios three relative prices are shown.

1. *All internal Universal Storage Platform Fibre Channel storage* — first 25TB are 146GB 15K Fibre Channel, remainder are 300GB 15K Fibre Channel.
2. *Internal tiered storage* — first 25TB are 146GB 15K Fibre Channel, remainder are internal SATA.
3. *External tiered storage* — first 25TB are 146GB 15K Fibre Channel, remainder are external SATA.

Figure 7. Representative Sampling of Simple Configurations at Different Storage Capacities and Internal/External Ratios



Summary

The Hitachi Universal Storage Platform V and Hitachi Universal Storage Platform VM offer great flexibility and perform superbly in any environment. The ability to support a wide range of service levels within a single storage infrastructure or to combine Universal Storage Platform–based storage with legacy or new external storage systems is unique. But making the proper design decisions for tiered storage designs requires understanding what can be done and what the tradeoffs are. This paper has introduced and explained a design methodology for developing tiered storage architectures for the Universal Storage Platform. It then discussed some of the options and relative costs for configuring the system at the disk drive and RAID group level.

Hitachi has a number of add-on software products that add important design options to tiered storage systems. These include storage virtualization products such as Hitachi Dynamic Provisioning and Basic Operating System V software, tier management and data mobility products such as Hitachi Tiered Storage Manager software and performance tools such as Hitachi Tuning Manager software.

With the broad selection of software, physical media and configuration options, including SATA, available for the Universal Storage Platform, you can benefit from taking advantage of its flexibility to meet your tiered storage requirements. Financially, the larger the system is, the greater the benefits that will be seen from internal and external SATA drives. For storage systems that are or will become large, the cost benefit can be dramatic.

In order to get the maximum value from your storage system it is vital to invest the time required to design the best possible solution to meet your unique requirements, whether these include capacity, reliability, performance or cost. This paper has just touched on some of these areas. Hitachi Data Systems and Hitachi TrueNorth Channel Partner sales and support representatives have access to advanced expert configuration and sizing tools and can help in this effort. And, to assist in delivering the highest possible quality solution, Hitachi Data Systems Global Solution Services are available and should be engaged.

Hitachi Data Systems Corporation

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

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

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