HITACHI
DATA SYSTEMS

Enterprise Storage

Optimizing Cache for Individual Workloads with the Hitachi Cache Partition Manager Feature

On Hitachi TagmaStore™ Adaptable Modular Storage and Workgroup Modular Storage

An Application Brief

By Kevin Sampson

April 2006

Executive Summary

Cache partitioning is a significant feature of the Hitachi TagmaStore™ Adaptable Modular Storage and TagmaStore Workgroup Modular Storage. The Cache Partition Manager feature, included with the Hitachi Resource Manager™ utility package, gives administrators the ability to tune performance of a storage system according to the specific needs of each application. For the first time in a midrange modular storage product, cache resources can be efficiently and securely divided to accommodate a diverse range of application workloads on a single platform.

The ability to match application needs to storage system functionality is a key tenet of Application Optimized Storage™ solutions from Hitachi Data Systems. Application Optimized Storage solutions increase performance, availability, and functionality, and reduce the cost of the storage environment. These goals are achieved by consolidating multiple workloads on a single storage infrastructure, and then accurately mapping the functionality of the storage to the quality of service needs of each business application.

The Cache Partition Manager feature supports a storage strategy based on Application Optimized Storage solutions by allowing storage system cache to be fine-tuned for each application. This enables applications that might otherwise have conflicting workloads to safely share the resources of the Hitachi modular storage systems.

Contents

Optimizing Cache for Individual Workloads with the Hitachi Cache Partition Manager Feature

On Hitachi TagmaStore™ Adaptable Modular Storage and Workgroup Modular Storage

An Application Brief

By Kevin Sampson

Optimizing Consolidated Workloads

One of the core components of a strategy based on Application Optimized Storage solutions from Hitachi Data Systems is the ability to consolidate data and workloads on cost-effective, high-performance, and highly available storage systems. Consolidation provides dramatic cost savings from centralized administration of resources and from greater utility of the available disk capacity. However, consolidation can not succeed without the ability to optimize data delivery according to the specific needs of individual business applications. The Hitachi Cache Partition Manager feature, which resides in the Hitachi Resource Manager™ utility package, provides the mechanism for performing this optimization on the Hitachi TagmaStore™ Adaptable Modular Storage and Workgroup Modular Storage systems.

Introducing the Cache Partition Manager Feature

The Cache Partition Manager feature, a component of the Resource Manager software, splits the user-definable portion of the Adaptable Modular Storage and Workgroup Modular Storage system cache into sub-partitions. The storage administrator determines the capacity allocated to each partition and the partition's segment size.

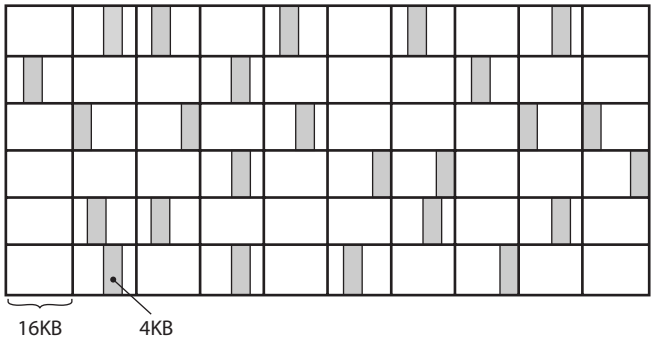
By tuning partition segment size according to the I/O block size requirements of an application the administrator can increase the amount of data that is resident in cache memory. For applications with a high percentage of read requests, the ability to access data in memory improves performance by reducing the need to replenish cache from disk.

The Cache Partition Manager feature also helps to segregate workloads that have conflicting cache requirements. For example, consolidating business-critical and lower-priority applications on a shared storage device can be problematic. The lower-priority application may attempt to hoard all cache resources, significantly impacting the performance of the business-critical application. The Cache Partition Manager feature prevents this from happening by allocating each application an appropriately sized and segmented cache partition. This eliminates the risk of one application interfering with the performance of other users of the same storage system.

Improving Cache Hit Ratios

Storage systems generally allocate cache resources in uniform 16KB segments. If the I/O of an application is processed in 4KB or 8KB blocks, then the mismatch between cache segment size and I/O block size leads to wasted memory resources, as shown in Figure 1. The inefficient use of the cache can contribute to poor cache hit ratios, causing the storage system to replenish cache from disk more frequently. The disk access adds latency to application read I/O requests, slowing response time.

Figure 1. Conventional Cache Segmentation Using a Single Block Size

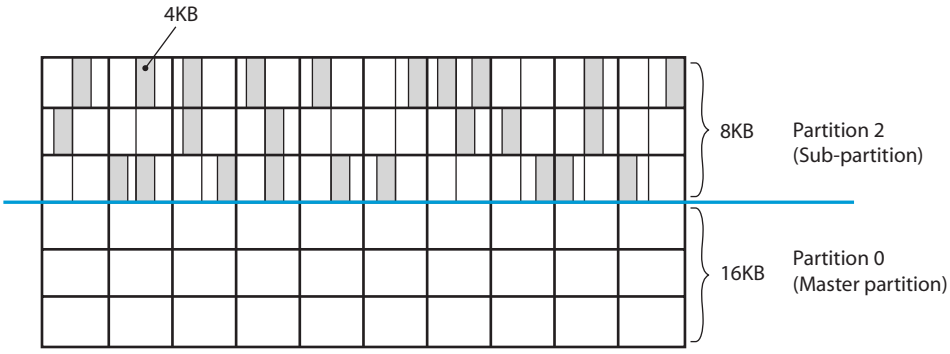


The inefficient use of the cache—smaller I/O from an application placed in a larger system-allocated cache segment—can contribute to poor cache hit ratios, causing the storage system to replenish cache from disk more frequently.

The Cache Partition Manager feature allows the cache resources of the Adaptable Modular Storage and Workgroup Modular Storage systems to be divided into multiple partitions. Each partition can be allocated a different segment size. LUNs are then assigned to a cache partition that has a segment size matching the application's I/O block size.

The close alignment of partition segment size and I/O block capacity allows more data to reside in cache, as shown in Figure 2. With more efficient use of cache resources, the potential for read request cache hits also increases, improving I/O per second performance for the application.

Figure 2. Partitioned Cache Allocates LUNs by Block Size to Optimize Data Retrieval

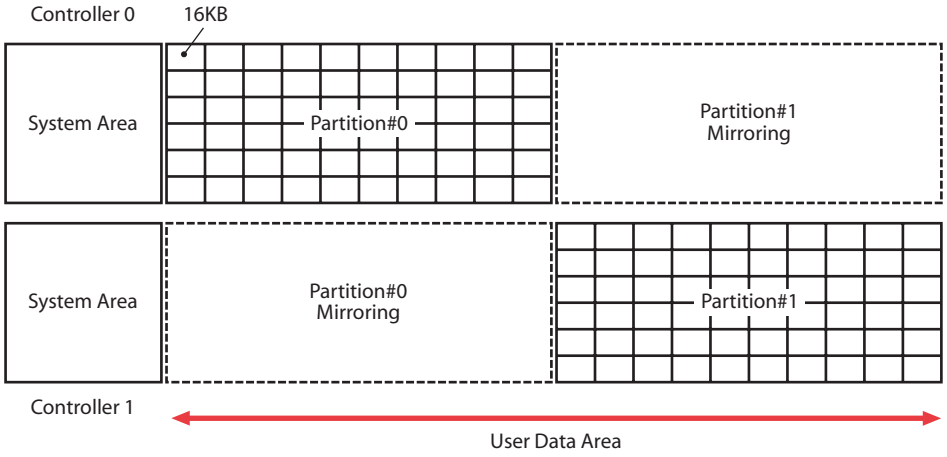


The Cache Partition Manager feature allows the cache resources of the Adaptable Modular Storage and Workgroup Modular Storage systems to be divided into multiple partitions, which can be allocated a different segment size.

How It Works

Without the Cache Partition Manager feature, storage system cache memory is allocated in consistent 16KB block sizes, as shown in Figure 3. Each storage system controller receives half of the available user data area of the cache, after taking account for mirroring and control areas.

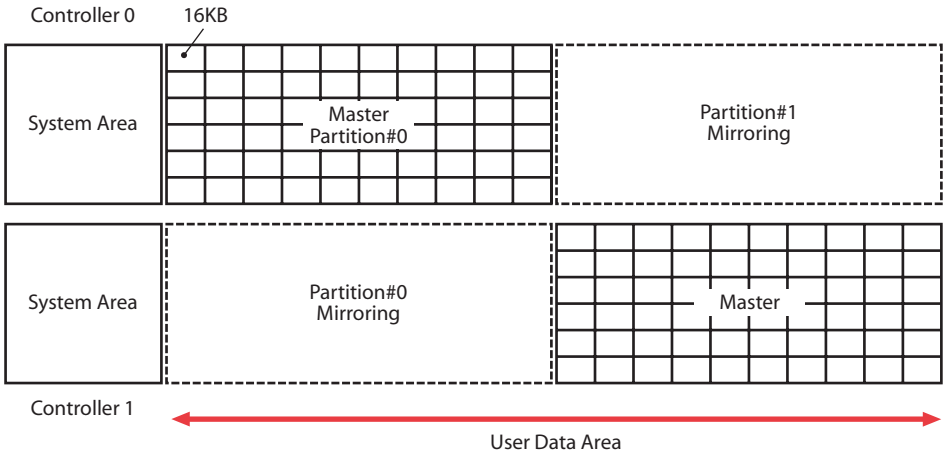
Figure 3. A Cache Memory Configuration without the Cache Partition Manager Feature



Without the Cache Partition Manager feature, the storage system cache memory allocates only 16KB block sizes.

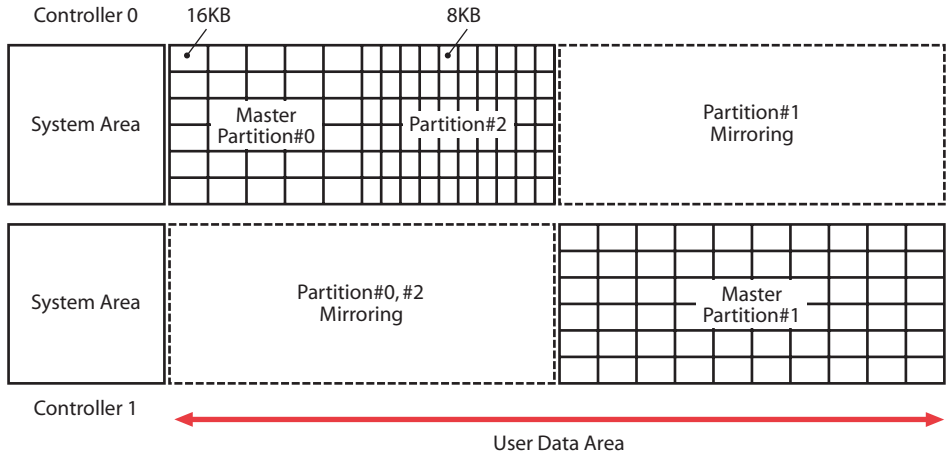
When the Cache Partition Manager feature is employed, each side of the cache is allocated to a controller and assigned a master partition. Only one master partition can exist on a controller, as shown in Figure 4, and it must have a segment size of 16KB. The master partition cannot be deleted but the memory capacity allocated to the partition can be changed. The Cache Partition Manager feature allows sub-partitions to be created in the user data area, as shown in Figure 5. And each sub-partition can be given a unique segment size— 4KB, 8KB, 16KB, 64KB, 256KB, or 512KB.

Figure 4. A Cache Partition Manager Configuration Showing Controllers 0 and 1



When the Cache Partition Manager feature is employed, each side of the cache is allocated to a controller and assigned a master partition.

Figure 5. A Cache Partition Manager Configuration Showing Partitioned Cache



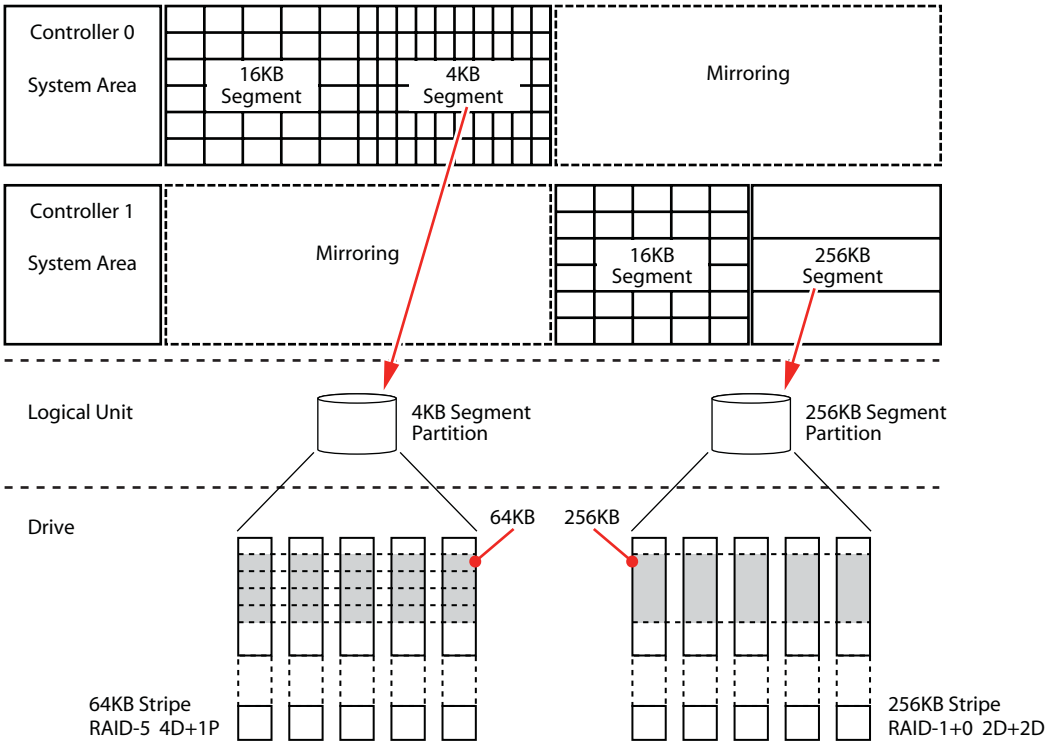
With the Cache Partition Manager feature, each sub-partition can be given a unique segment size— 4KB, 8KB, 16KB, 64KB, 256KB, or 512KB.

Changing the cache partition segment size to match the LU stripe size allows data to be sent or received to or from a disk drive more efficiently.

LU Stripe Size

One cache partition for data sending/receiving is assigned to each logical unit (see Figure 6). The stripe size of each logical unit is typically larger than the cache segment size of the assigned partition. When Cache Partition Manager feature isn't enabled, the default is a 16KB partition with a stripe size of 64KB. For some applications, such as those with frequent "big block" writes or NAS-based file systems, a stripe size of 256KB or 512KB along with a corresponding segment size may improve efficiency in getting data to and from cache. The stripe size is specified when a logical unit is created from LUN management, it is not created within the Cache Partition Manager feature. Note that while cache segment sizes may be changed, stripe sizes may not be changed without re-creating the LU.

Figure 6: Adjusting Cache Partition Segment Size



Changing the cache partition segment size to match the LU strip size allows data to be sent or received to or from a disk drive more efficiently.

Supported Cache Partition Manager Configurations

Supported partition capacity is a function of the total cache available and the number of partitions supported by the storage system. Cache is always mirrored, and partitions must fit within the memory allocated as user data area. Future product releases will allow cache mirroring to be suspended for high-performance applications.

Table 1 shows the maximum cache configurations for each of the Adaptable Modular Storage and Workgroup Modular Storage systems, while Table 2 depicts combinations of segment size and stripe size.

Table 1. Supported Cache Partitions for Each Storage System Configuration

Storage System Model and Cache Memory Configuration	Supported Partitions[1]	User Data Area[2]	Default Partition Size[3]	Minimum Partition Size		Maximum Sub-partition Cache Size[4]
				Master	Sub	
Hitachi TagmaStore™ Adaptable Modular Storage system[5]						
AMS1000 8GB per controller	2 to 32	6,760	3,380	180	100	6,580
AMS1000 6GB per controller	2 to 32	4,940	2,470	180	100	4,760
AMS1000 4GB per controller	2 to 24	3,080	1,540	180	100	2,900
AMS1000 2GB per controller	2 to 8	1,240	620	180	100	1,060
AMS500 system with 4GB per controller	2 to 16	3,240	1,620	180	100	3,060
AMS500 system with 3GB per controller	2 to 16	2,340	1,170	180	100	2,160
AMS500 system with 2GB per controller	2 to 12	1,400	700	180	100	1,220
AMS200 system with 2GB per controller	2 to 8	1,560	780	100	100	1,460
AMS200 system with 1GB per controller	2 to 6	660	330	100	100	560
Hitachi TagmaStore Workgroup Modular Storage system[5]						
WMS100 system with 1GB per controller	2 to 6	680	340	100	100	580

1 Numbers include the two default master partitions

2 Maximum user data cache per controller

3 Hitachi Cache Partition Manager feature installed but no sub-partitions defined

4 User Data Area minus Minimum Master Partition Size

5 All amounts are MB, unless otherwise noted

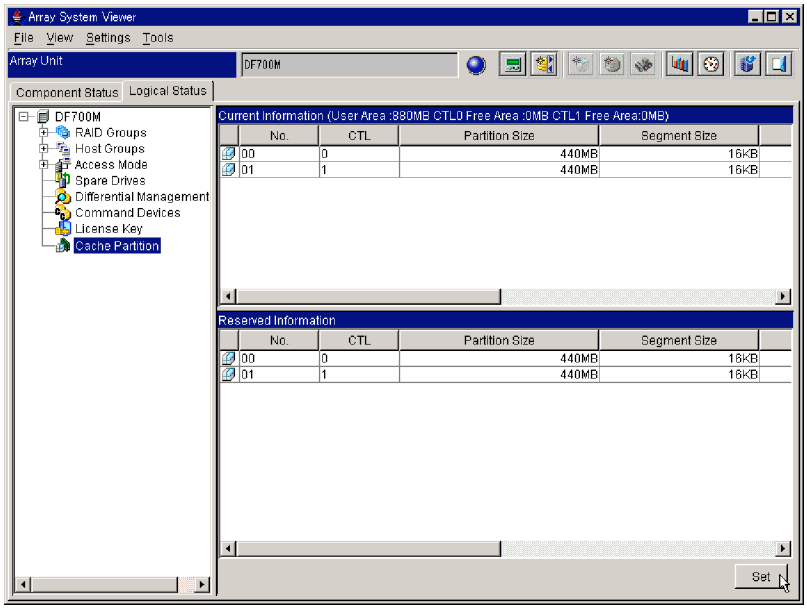
Table 2. Combinations of Segment Size and Stripe Size

	Stripe Size		
Segment Size	64KB	256KB	512KB
4KB	Yes	—	—
8KB	Yes	Yes	—
16KB	Yes	Yes	Yes
64KB	Yes	Yes	Yes
256KB	—	Yes	Yes
512KB	—	—	Yes

Allocating a Cache Partition

The Cache Partition Manager feature is accessed through the Storage Navigator module of the Resource Manager utility package. Partitions are added from the Array System Viewer panel, as shown in Figure 7.

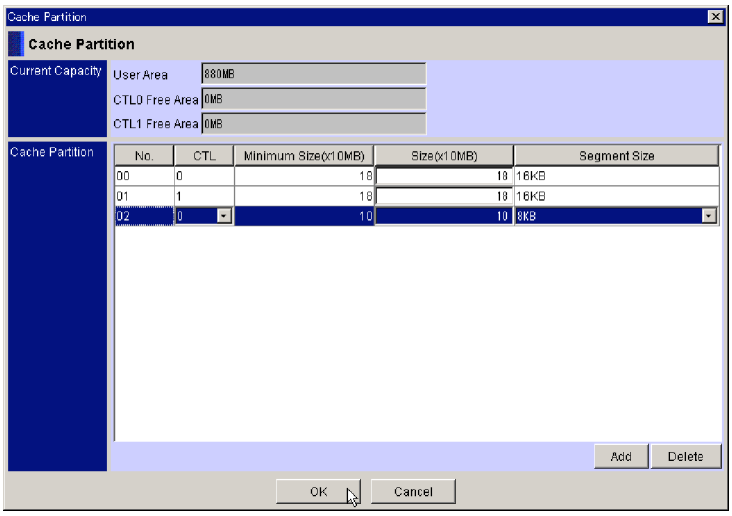
Figure 7. Array System Viewer Panel Showing Initial Cache Partitions



To add partitions, access the Cache Partition Manager feature through the Storage Navigator module of the Resource Manager utility package.

The user interface allows administrators to create partitions, allocate capacity, and assign sizes to segments, as shown in Figure 8. Once a cache partition is created it must be assigned to an LU in the storage system and the system must be restarted for changes to take effect. In future product releases, a system reboot will not be required when changing partition size.

Figure 8. Cache Partition Dialog Box



Once a cache partition is created using the Cache Partition Dialog Box, it must be assigned to an LU in the storage system and the system must be restarted for changes to take effect.

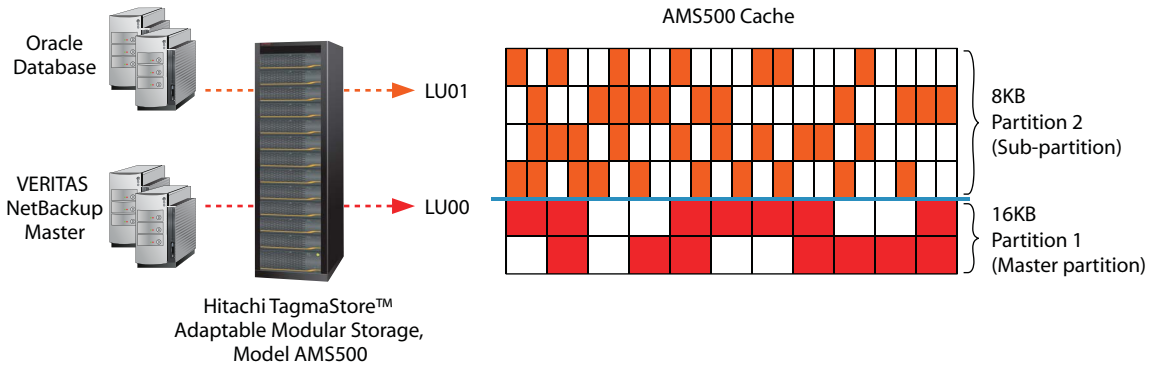
Examples of Cache Partitioning

Consolidating Diverse Application Workloads

Successful consolidation of workloads on a single storage system can occur only if safeguards are in place to prevent the actions of one application from interfering with the performance of others. The Cache Partition Manager feature provides this safeguard on the Adaptable Modular Storage and Workgroup Modular Storage systems.

Figure 9 illustrates an example in which a high-priority application—an Oracle database—is sharing storage resources with a lower priority workload—in this example, a VERITAS NetBackup server. With conventional cache management, this scenario introduces the potential for performance problems. If the NetBackup server is running a backup when the Oracle database experiences a cache miss, requiring the database to go to disk for data, NetBackup will take the opportunity to consume significant amounts of available cache. This starves the Oracle database application of cache resources, slowing database request response time and impacting end-user performance.

Figure 9. Consolidating Diverse Application Workloads with the Cache Partition Manager Feature



Unlike conventional cache management, which may encounter performance problems if a high-priority application (such as an Oracle database) is sharing storage resources with a lower priority workload (such as a NetBackup server), cache management with the Cache Partition Manager feature separates high- and low-priority workloads in cache and avoids performance issues.

With the Cache Partition Manager feature the Oracle and NetBackup workloads are separated in cache. The lower priority NetBackup application can be allocated a smaller amount of memory, with a segment size that allows the application to make efficient use of the resource. Similarly, cache allocated to the Oracle server can be sized to suit the needs of database I/O. This segregation eliminates the potential for I/O activity from one application impacting others. And, with more efficient cache allocations, application performance also improves.

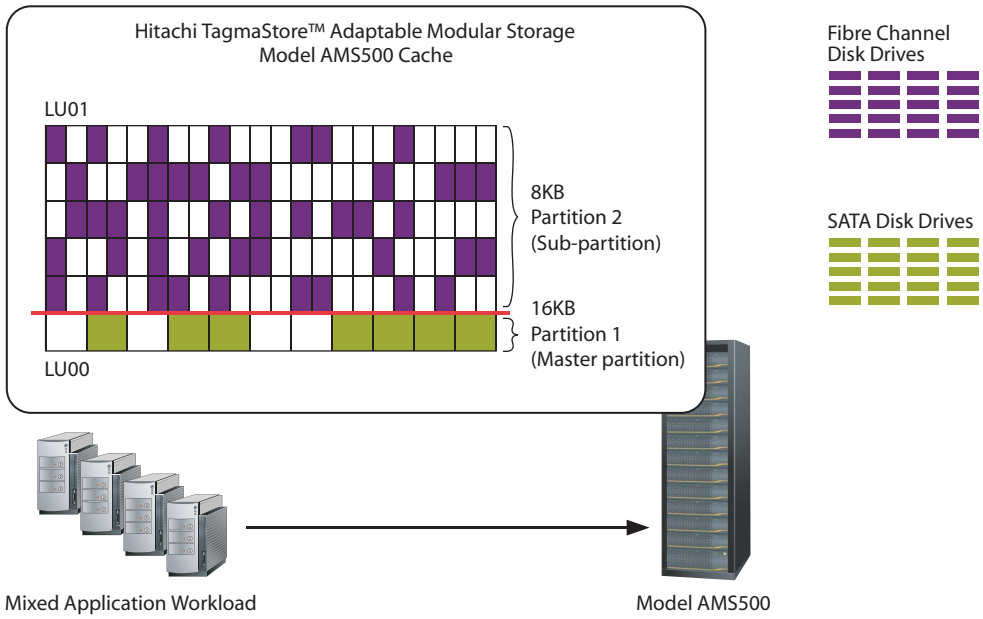
Optimizing Combined SATA and Fibre Channel Disk Drives

The versatile Adaptable Modular Storage systems support a combination of Fibre Channel and Serial ATA (SATA) disk drives in the same system, as shown in Figure 10. This allows storage managers to allocate data on disks that reflect an application's preference for speed, availability, and cost-effectiveness. However, with different spin rates, capacity, and mean-time-between-failure characteristics, Fibre Channel and SATA drives place very different demands on cache resources.

Although the same application might use both Fibre Channel and SATA storage, the cache requirements of the data stored on the two drive types are liable to be very different. For example, an e-mail application may use fast Fibre Channel disks to host the primary messaging database, but write weekly e-mail archives to lower cost SATA drives. The one-way write I/O process of e-mail archival does not make effective use of cache resources; data is written out to disk in large blocks and never re-read from cache. This use pattern is typical of applications that benefit from SATA disk.

The Cache Partition Manager feature allows the cache of the Adaptable Modular Storage system to be allocated to disk drives that are likely to make the best use of the resources. Because data stored on SATA drives is infrequently re-read while in memory, the SATA drives can be allocated a smaller portion of the available cache. This frees cache resources for use by the more demanding Fibre Channel disk drives.

Figure 10. Segregating Fibre Channel and SATA Workloads Using the Cache Partition Manager Feature



The Cache Partition Manager feature allows the cache of the Adaptable Modular Storage system to be allocated to disk drives that are likely to make the best use of the resources.

Summary

The Cache Partition Manager feature provided in the Hitachi Resource Manager utility package increases the flexibility of the Adaptable Modular Storage and Workgroup Modular Storage systems, allowing them to effectively host a diverse range of consolidated application workloads. The partitioning capability, unique to Hitachi midrange modular storage systems, allows the new storage systems to be fully integrated into an Application Optimized Storage solutions environment.

Hitachi Data Systems Corporation

Corporate Headquarters
750 Central Expressway
Santa Clara, California 95050-2627
U.S.A.
Phone: 1 408 970 1000
www.hds.com
info@hds.com

Asia Pacific and Americas
750 Central Expressway
Santa Clara, California 95050-2627
U.S.A.
Phone: 1 408 970 1000
info@hds.com

Europe Headquarters
Sefton Park
Stoke Poges
Buckinghamshire SL2 4HD
United Kingdom
Phone: + 44 (0)1753 618000
info.eu@hds.com

WHP-201-01 April 2006