HITACHI
Inspire the Next

About capacity saving

Capacity saving function: data deduplication and compression

When the capacity saving function is in use, data deduplication and compression is performed to reduce the size of data to be stored. Capacity saving can be enabled on DP-VOLs in Dynamic Provisioning pools and Dynamic Tiering pools. You can use the capacity saving function on all drive types, including data stored on encrypted flash drives and external storage. However, you cannot use the capacity saving function on SCM drives.

A data reduction shared volume (DRS-VOL) is supported to use as a volume of a Thin Image Advanced pair with the DKCMAIN microcode version 90-08-81-00/00 or later. The DRS-VOL is a DP-VOL with the capacity saving function. Therefore, the descriptions of the capacity saving function in this document also apply to the DRS-VOL unless otherwise noted.
How capacity saving works

The capacity saving function includes deduplication and compression.

Deduplication

The data deduplication function deletes duplicate copies of data written to different addresses in the same pool and maintains only a single copy of the data at one address. The deduplication function is enabled on a Dynamic Provisioning pool and then on the desired DP-VOLs in the pool. When deduplication is enabled, data that has multiple copies between data on DP-VOLs assigned to that pool is removed.

When you enable deduplication on a pool, a deduplication system data volume (DSD volume) for that pool is created. The DSD volume is used exclusively by the storage system to manage the deduplication function. A search table in the DSD volume is used to locate redundant data in the pool.

When you create DP-VOLs with the deduplication function enabled, deduplication system data volumes (fingerprint and data store) are automatically created. The deduplication system data volumes (fingerprint) store the search table for searching the deduplicated data. For one pool, 24 deduplication system data volumes (fingerprint) are created. The deduplication system data volumes (data store) store the source data for duplicated data. For one pool, 24 deduplication system data volumes (data store) are created. When deduplication is disabled for all DP-VOLs in a pool, all deduplication system data volumes in the pool are automatically deleted.
NoteIf the compression accelerator feature applies to a DRD-VOL whose capacity saving setting is Deduplication and Compression, the compression accelerator feature applies to the duplicated data on all DRD-VOLs that are created from the pool containing the DRD-VOL.

Compression

The data compression function enables you to convert the stored data into data with smaller data size by encoding without reducing the amount of data information. The data compression with software and the data compression with the hardware dedicated to data compression are available (see Compression accelerator feature for details). The LZ4 compression algorithm is used for data compression using software with the DKCMAIN microcode.

The data compression function can be enabled for each DP-VOL used for Dynamic Provisioning. Using compression with HDDs in Dynamic Provisioning or Dynamic Tiering pools is not recommended. Contact customer support for details.

The compression accelerator feature requires a higher performance compression algorithm than one used for data compression with software.

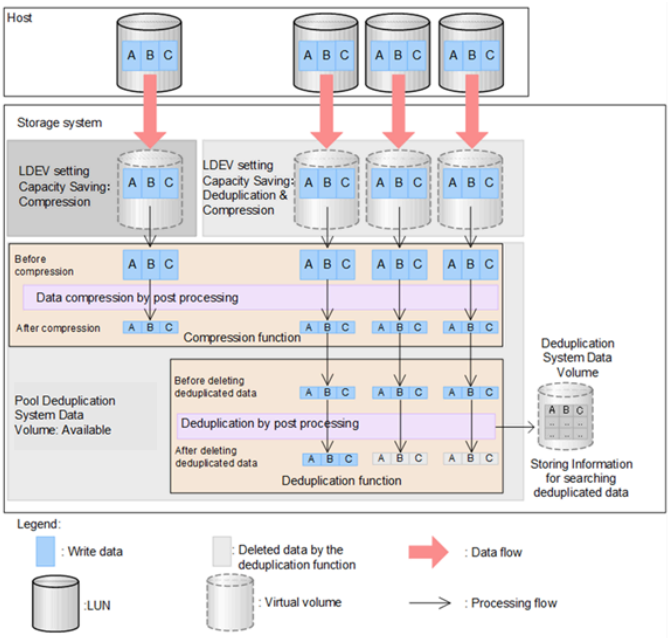
Compression accelerator feature

The compression accelerator feature provides the following features:

- Data compression is performed by using the hardware that is dedicated to data compression. Therefore, the loads on the processors can be mitigated, and the I/O performances can be improved, compared to data compression with software. The data-size saving rate due to data compression is also improved. In particular, the larger the I/O data size, the better the I/O performance.
- Even if the data is compressed with software, the data can be converted to the data compressed by using the compression accelerator feature.

How data is stored with the capacity saving function

This figure shows how data is stored by using the capacity saving function.



When the post-process mode is applied, data received by the storage controller is stored in a temporary area in the pool. When the data is classified as inactive (five minutes since the last update for Dynamic Provisioning), the capacity saving processing is performed, and the post-process data is stored in the data storage area. When post-process data is updated again, the data stored in the data storage area is no longer required. This kind of data is called garbage data. The used capacity of the pool increases until garbage collection, which collects old data that is no longer required. The pool capacity that is eventually required is the sum of the physical data capacity after capacity saving plus the amount of metadata.

Note

- The temporary area and the data storage area are not assigned fixed capacities. They share the pool and use the pool as needed.
- The temporary area is used only when the post-process mode is applied. When the inline mode is applied, capacity saving processing is performed simultaneously with receiving of data from the host, and host data is not stored in the temporary area.
- When the capacity saving function is enabled, the garbage data is created during the following processing and consumes the pool capacity:
  - Update the data on a DP-VOL with compression set.
  - Update the data that is included on a DP-VOL with deduplication and compression set but not included on a different DP-VOL.
  - Delete a Thin Image Advanced pair.

The capacity overhead associated with the capacity saving function includes these items:

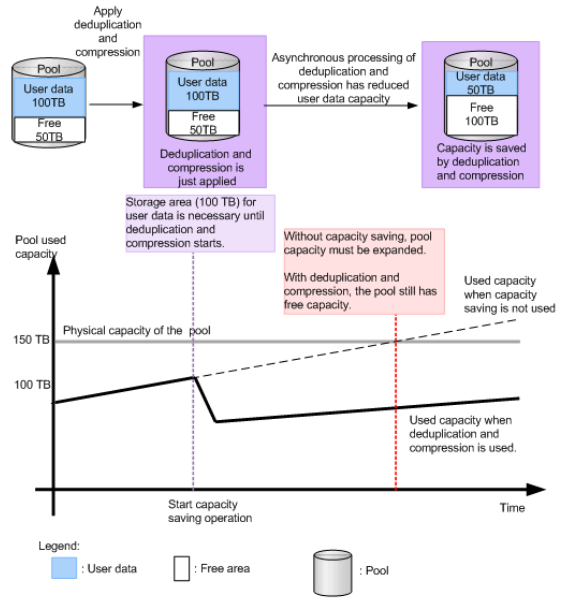
- *Capacity consumed by metadata*: The capacity consumed by metadata for the capacity saving function (deduplication and compression) is approximately 3% of the consumed DP-VOL capacity that has been processed by capacity saving. For example, if the consumed capacity of a DP-VOL is 150 TB and the capacity saving feature has processed 100 TB of the 150 TB consumed capacity and reduced it to 30 TB, the capacity consumed by metadata for the capacity saving function is approximately 3 TB (3% of 100 TB). The total consumed capacity of this DP-VOL at this instant is 83 TB (30 TB + 50 TB + 3 TB).

- *Capacity consumed by garbage (invalid) data*: The capacity consumed by garbage data is approximately 7% of the total consumed capacity of all DP-VOLs with capacity saving enabled. The capacity is dynamically consumed based on garbage data created by the capacity saving process and cleaned by the background garbage collection process. Garbage collection is a background process with a lower priority than host I/O, so the capacity consumed by garbage data depends on both the garbage created and the host I/O rate.

The total capacity consumed by these overheads is about 10% (3% for metadata + 7% for garbage data) of the consumed capacity of DP-VOLs with capacity saving enabled. During periods of high write activity from the host, this capacity might increase over 10% temporarily, and then it returns to around 10% when host write activity decreases.

For a DRS-VOL, the total capacity consumed by metadata and garbage data is about 13% of the pool capacity.
CautionWhen the available space in a pool becomes 1% or 120 GB or less, capacity deletion processing might stop, or performance might degrade.
Capacity saving processing for existing data

The deduplication and compression processing is performed asynchronously for pages that store data, and the free area of the pool can be increased, thereby reducing the cost of purchasing drives over time.



Capacity saving processing for new write data

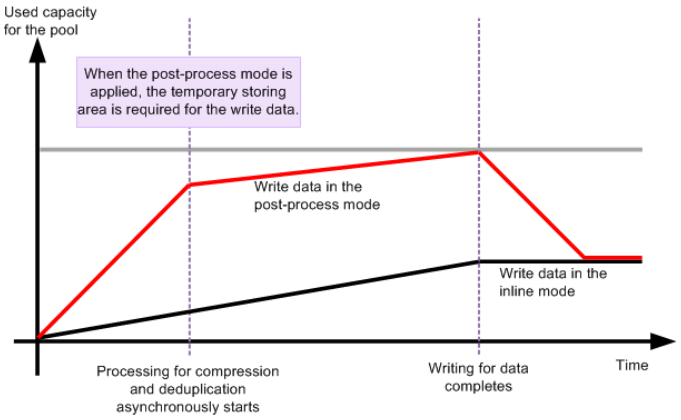
The capacity saving mode of a DP-VOL (post-process mode or inline mode) determines how capacity saving is applied to new write data from the host:

- *Inline mode (default)*: When you apply capacity saving with the inline mode to a DP-VOL, the compression and deduplication processing are performed synchronously for new write data. The inline mode minimizes the pool capacity required to store new write data but can impact I/O performance more than the post-process mode. The inline mode should be applied when writing data with sequential I/Os, for example, when writing data to target volumes of data migration or secondary volumes of copy pairs. When the data migration or copy pair creation has completed, the mode should be changed from the inline mode to the post-process mode to minimize the impact on I/O performance.

If you want to change the default mode (inline) to post-process mode, you must use CCI (raidcom add ldev [-capacity_saving_mode <saving mode>] or raidcom modify ldev [-capacity_saving_mode <saving mode>]).

- *Post-process mode*: When you apply capacity saving with the post-process mode to a DP-VOL, the compression and deduplication processing are performed asynchronously for new write data. Since capacity saving processing is not performed at the time the new data is written, the post-process mode can reduce the impact of capacity saving processing on I/O performance. However, pool capacity is required to store the new write data until the capacity saving processing is performed.

This example shows how the pool used capacity changes over time when performing data migration. The red line shows the capacity when the post-process mode is applied, and the black line shows the capacity when the inline mode is applied. This example assumes that the writing speed (GB/h) for the new data is faster than the initial capacity saving processing (GB/h).



When inline mode is applied, capacity saving processing is performed synchronously for the writing of data. When post-process mode is applied, capacity saving processing is performed asynchronously for the writing of data, and a temporary storage area is required for the write data. The capacity required for the temporary storage area depends on the writing speed of the new data or on the frequency of data updates during migration.

This table shows the processing method (synchronous or asynchronous) for initial data, new write data, and updated data. For new write data, capacity saving processing is performed at different times for post-process mode and inline mode.

Mode	Initial data[1]	New write data		Updated write data	
		Compression processing	Deduplication processing	Compression processing	Deduplication processing
Post-process mode	Asynchronous	Asynchronous[4]	Asynchronous	Synchronous[3]	Asynchronous
Inline mode	Asynchronous	Synchronous	Synchronous[2]	Synchronous[3]	Asynchronous

Notes:

1. The initial data is the existing data on the DP-VOL when the capacity saving function is enabled. The (initial)

Mode	Initial data[1]	New write data		Updated write data	
		Compression processing	Deduplication processing	Compression processing	Deduplication processing

capacity saving processing is performed for the initial data.

2. Applied to sequential I/O data, such as writing large amounts of data sequentially. Deduplication of random I/O data, such as updating files irregularly, is performed in post-process mode.
3. Indicates the compression method of the written data when compressed data is updated. If uncompressed data before initial capacity deletion is updated, compression of the written data is performed in post-process mode.
4. For a DRS-VOL, the compression processing is always performed synchronously even if the capacity saving mode is post-process mode.

Use cases for capacity saving

The results of enabling the capacity saving functions of deduplication and compression depend on the properties and access patterns of the stored data. In addition, when capacity saving is enabled, some storage behaviors are different from conventional behaviors because of the increase in load of storage controller processing caused by data scanning and garbage collection by data update. Before implementing capacity saving, you need to confirm whether it should be applied to your specific storage environment.

This table lists several storage use cases and describes the application of capacity saving to each use case.

Use case	Settings	Description
Office	Deduplication and compression	Because there are many identical file copies, deduplication is effective.
VDI	Deduplication and compression	Deduplication is very effective because of OS area cloning.
Database	Compression	Deduplication is not effective because the database has unique information for each block.
Image/video	Not suitable (Disable)	Compressed by application.
Backup/archive	Deduplication and compression	Deduplication is effective between backups.

Caution

- I/O performance to data with compression and deduplication is degraded. Verify the performance by utilizing best practices before using the capacity saving function.
- Because approximately 10% is used for metadata and garbage data, capacity saving should be applied only when the result is expected to be 10% or higher.

- In deduplication and compression, processing is performed per 8 KB. Therefore, if the block size of the file system is an integral multiple of 8 KB, capacity saving is likely to be effective.
- The capacity saving function is not a good fit for high-write workloads. If the write workload rate is higher than garbage collection throughput, Cache Write Pending increases, causing performance degradation. Contact customer support to determine the garbage collection throughput for your configuration.

If you want to perform fast operations of logical data backup, archiving, or utilization of replicated data, it is recommended to take snapshot data by using Thin Image Advanced with DRS-VOLs. Thin Image Advanced enables the fast operations of taking snapshot data by using metadata with storage location information. This table shows the examples of recommended combinations of functions in software products applied to take snapshot data.

Combination of functions in software products	Capacity saving function	Description
DP-VOLs with capacity saving enabled that do not include a DRS-VOL, and ShadowImage or clone in Thin Image	Deduplication and compression	The data copy operation takes a long time to process because deduplication and compression is performed after the user data copy operation.
DP-VOLs with capacity saving enabled that do not include a DRS-VOL and snapshot in Thin Image	Deduplication and compression Deduplication and compression are only effective for a P-VOL of Thin Image. Deduplication and compression do not apply to the snapshot data of Thin Image.	The performance degradation might be caused after the Thin Image pair split operation.
DRS-VOLs and Thin Image Advanced	Deduplication and compression Deduplication and compression are effective for a P-VOL and snapshot data of Thin Image Advanced.	• The performance degradation is not caused after the Thin Image Advanced pair split operation. • The I/O performance degradation in ShadowImage is not caused. • A pool configured with flash media (SSD, FMD, or SCM) is required.

Use cases for the compression accelerator feature

The accelerator fan module (ACLF) that includes the compression accelerator feature is installed on VSP 5200 or VSP 5600 by default. If you select Compression, or Deduplication and Compression for the capacity saving setting to create a DRD-VOL by using HDvM - SN, a DRD-VOL with the compression accelerator feature automatically set is created.

When a DP-VOL created on VSP 5100 or VSP 5500 is migrated to VSP 5200 or VSP 5600, set the compression accelerator feature. See Setting compression accelerator for a DRD-VOL. Settings from CCI are applicable. For more information, see *Command Control Interface Command Reference.*

Usage planning requirements for the capacity saving function

This table outlines the items to review and plan for before using the capacity saving function.

Classification	Item	Remarks
Implementation	Implementation method	• New implementation • Changing DP-VOL to DRD-VOL • Migrating from old model (using a program product) • Migrating an existing DP-VOL or a DRD-VOL to a DRS-VOL by using Volume Migration. • Migrating from old model (through a server) • Migrating an existing DP-VOL or a DRD-VOL to a DRS-VOL through a server.
Capacity	Total used capacity of DP-VOL	Total used capacity (before capacity saving) of DP-VOL to which the capacity saving function is applied.
	Capacity saving ratio [%]	If the data to which capacity saving is applied already exists, you can run the Data Reduction Estimation Tool. If the data to which capacity saving is applied does not exist, you can estimate the capacity saving ratio by use cases. Contact customer support for the use cases and estimated capacity saving ratio. Capacity saving ratio shown as N:1 can be converted to the capacity saving rate in % by using the following formula: Capacity saving rate

Classification	Item	Remarks
		[%]=(1-1÷N)×100
	Total used capacity of DP-VOL	Total used capacity of DP-VOL to which the capacity saving function is not applied.
Configuration	Storage system model	When planning the pool, if you want to implement capacity saving and maintain host performance, you must consider which model is suitable.
	RAID level	RAID 1, RAID 5, or RAID 6 can be used.
	Drive type	Use the same drive type (including rotational speed) in a pool. Flash media (SSD, FMD, or SCM) are only available for DRS-VOLs.
	Capacity of one parity group	None.
Performance	Requirement for throughput (IOPS)	When planning the pool, if you want to implement capacity saving and maintain host performance, these items must be considered. If you account for these items, Performance Monitor output can be used. Average I/O size can be calculated as follows: Average throughput [MB/s] ÷ Average throughput [IOPS] × 1024
	Read/Write ratio	
	Average I/O size [KB]	
	Performance boundary for one parity group [IOPS]	Calculate the performance boundary for one parity group by using performance information: • Drive type: Consider the drive you plan to use. • Read/Write ratio: Consider the throughput requirements. • I/O size: Consider the throughput requirements.
Other requirement	Use of encryption	You can use both encryption and

Classification	Item	Remarks
		capacity saving. For details, see *Encryption License Key User Guide*.

Storage planning considerations for the capacity saving function

Review this table for information about settings, configuration, and performance considerations when using the capacity saving function.

Category	Item	Remarks
Setting	Capacity saving setting	Determine the capacity saving function to use by using the capacity saving rate (%) that is estimated by the Data Reduction Estimation Tool or by guessing at the capacity saving rate (%).
Configuration	Capacity of a DP-VOL whose capacity saving setting is enabled	• If the DKCMAIN microcode version is earlier than 90-04-04-00, estimate the number and capacity of the DP-VOLs to be provided to the host. For DP-VOLs with the capacity saving setting enabled, it is recommended that the DP-VOL capacity be less than 2.4 TB. If you create a DP-VOL with 2.4 TB or more, the processing efficiencies of the capacity saving and the garbage collection will decrease due to the capacity limitation of the cache management device, and the data-size saving rate will be degraded. If the number of DP-VOLs is small, the following performances might not be obtained. ◦ Host I/O ◦ Capacity saving processing for the initial data in the post-process mode ◦ Garbage collection ◦ Data migration in the inline mode ◦ Disabling the capacity saving setting ◦ Formatting an LDEV ◦ Deleting an LDEV ◦ Initial copy operation • If the DKCMAIN microcode version is 90-04-04-00 or later, estimate the number and capacity of the DP-VOLs to be provided to the host. If the number of DP-VOLs is small, the following performances might not be obtained.

Category	Item	Remarks
		◦ Host I/O ◦ Capacity saving processing for the initial data in the post-process mode ◦ Garbage collection ◦ Data migration in the inline mode ◦ Disabling the capacity saving setting ◦ Converting LDEV data compression The transition from data compression with software to data compression with compression accelerator, or from data compression with compression accelerator to data compression with software is enabled on VSP 5200 or VSP 5600 by default. ◦ Initial copy operation
	Number of parity groups	Determine the number of parity groups when designing a pool. If you consider the number of parity groups, these cases can be considered: • The capacity, alone • The capacity and performance For details, contact customer support.
	Cache memory capacity	Determine the cache memory capacity to be installed based on the total used DP-VOL capacity. For details, contact customer support.
	Shared memory capacity	Determine the shared memory capacity to be installed based on the total used DP-VOL capacity. For details, contact customer support.
Performance	Estimated performance value	Estimate the average write throughput in a customer use case and confirm that garbage data does not keep increasing with the workload. For the average write throughput, estimate the write throughput in the operation cycle (1 day to 1 week, for example). Use information output by Performance Monitor for estimation. If garbage data increases constantly, the capacity saving function cannot be applied.

Pool capacity consumed by metadata

When you use capacity saving, these capacities are consumed for the pool capacity:

- Used capacity of the pool consumed by user data
- Used capacity of the pool consumed by garbage data
- Used capacity of the pool consumed by metadata

Metadata for the compression function: When the compression function is enabled, 2% of the total used capacity of the compression-enabled DP-VOLs is consumed as the metadata for the compression function. The capacity of the metadata for the compression function is added to the used capacity of the pool. To view the used capacity of the pool, see Pool Capacity (Used/Total) in the Pools window. To view the system data capacity for a pool, see System Data in the Pools window. The system data capacity indicates the total capacity of meta data and garbage data.

Metadata for the deduplication function: When the deduplication function is enabled, 3% of the total used capacity of the deduplication-enabled DP-VOLs is consumed as the metadata for the deduplication function. To view the capacity of the metadata for the deduplication function, see the capacity of the deduplication system data volumes (finger print). The capacity of the metadata of the deduplication function is added to the used capacity of the pool. To view the used capacity of the pool, see Pool Capacity (Used/Total) in the Pools window. To view the system data capacity for a pool, see System Data in the Pools window. The system data capacity indicates the total capacity of meta data and garbage data.

Deduplication system data volume specifications and requirements

This table lists the requirements for the deduplication system data volume (fingerprint).

Item	Description
Volume type	DP-VOL (V-VOL). When DP-VOLs using deduplication and compression are created, the deduplication system data volumes are automatically created.
Emulation type	OPEN-V
Number per pool	24 deduplication system data volumes (fingerprint) are associated with a pool. DRD-VOLs and deduplication system data volumes that belong to the same pool must belong to the same CLPR.
Volume capacity	1.7 TB

Item	Description
Cache management devices	1 deduplication system data volume (fingerprint) uses 1 cache management device.
Path definition	Cannot be defined.
LDEV format	Cannot be performed.

This table lists the requirements for the deduplication system data volume (data store).

Item	Description
Volume type	DP-VOL (V-VOL). When DP-VOLs using deduplication and compression are created, the deduplication system data volumes (data store) are automatically created.
Emulation type	OPEN-V
Number per pool	24 deduplication system data volumes (data store) are associated with a pool.
Volume capacity	From 5.98TB to 42.7 TB The DSD-VOL capacity has no impact on the DP-VOL capacity available to the host, which is described in the previous table. The subsequent table describes the maximum capacity of deduplication system data volumes (data store) for a pool and a storage system. The capacity size when the volumes are initially created is the same with the total capacity of pool volumes in a pool. When the pool operation starts, the total capacity of the 4 deduplication system data volumes (data store) automatically expands to the same size of the total pool volumes capacity. However, you can also manually expand capacities of these volumes.
Cache management devices	Number of cache management devices that are used for 1 deduplication system data volume is from 4 to 30.
Path definition	Cannot be defined.
LDEV format	Can be performed for LDEVs that are initialized of the duplicated data in a pool.

This table lists the maximum capacity of deduplication system data volumes (data store) for a pool and a storage system.

Added shared memories	Maximum capacity of deduplication system data volumes (data store) for a pool (PB)	Maximum capacity of deduplication system data volumes (data store) for a pool (PB)
Base	Smaller capacity of the following: • 1.0 • Total capacity of pool volumes in a pool	1.1
Extension 1	Smaller capacity of the following: • 1.0 • Total capacity of pool volumes in a pool	2.0125
Extension 2	Smaller capacity of the following: • 1.0 • Total capacity of pool volumes in a pool	3.125
Extension 3	Smaller capacity of the following: • 1.0 • Total capacity of pool volumes in a pool	4.15

Compatibility with the capacity saving function

The capacity saving function is not compatible with certain products and functions. This table specifies the compatibility restrictions for the capacity saving function.

Program product/function	Restrictions when using the capacity saving function
Dynamic Provisioning	The V-VOL full allocation function cannot be used. The subscription limit can only be set to Unlimited for the pool. To prevent writing failures caused by full pool capacity, you must consider monitoring the free space of a pool.
Dynamic Tiering	DRS-VOLs cannot be used in Dynamic Tiering. DP-VOLs with capacity saving enabled that do not include DRS-VOLs can be used in Dynamic Tiering.
Universal Volume Manager	The data direct mapping attribute cannot be set to DP-VOLs for which capacity saving is enabled. You must separate pools for which data direct

Program product/function	Restrictions when using the capacity saving function
	mapping is applied and pools for which capacity saving is enabled.
ShadowImage quick restore	The ShadowImage quick restore function cannot be used. Therefore, it takes time when you restore the backup data and then resume the application.
Volume Migration	The Volume Migration product cannot be used on DP-VOLs for which capacity saving is enabled. If you need to migrate capacity saving-enabled DP-VOLs, use a different method to migrate them (for example, host-based migration).
Accelerated compression	The capacity saving function can be used with accelerated compression, but accelerated compression is effective only for certain tasks. In this case, you must select the appropriate function depending on the task, as described in the next table.

This table describes the behavior when capacity saving and accelerated compression are used.

Capacity saving		Accelerated compression	Behavior
Compression	Deduplication and compression		
Disabled	Disabled	Enabled	Only accelerated compression is performed. The storage controller does not perform the compression/deduplication processing, so I/O performance is not impacted.
Enabled	Disabled	Disabled	The storage controller compresses data and stores the compressed data in the pool. Controller-based compression and accelerated compression can be used simultaneously, but it is best practice not to do so, because performance would be degraded compared to use of only accelerated compression.

Capacity saving		Accelerated compression	Behavior
Compression	Deduplication and compression		
Disabled	Enabled	Enabled	When identical data is stored in a pool, the storage controller keeps only one of copy of the data (deduplication). For compression, the storage controller automatically determines that accelerated compression can be used and uses it. The pool must consist of only FMDs, and accelerated compression needs to be enabled in all parity groups in the pool.
Disabled	Enabled	Disabled	The storage controller performs the deduplication and compression processing. The storage controller has the largest overhead of the capacity saving processing.

Capacity saving effects for DP-VOLs and pools

You can use HDvM - SN or CCI to check the capacity saving effects for DP-VOL, but you can use only HDvM - SN to check the capacity saving effects for each pool.

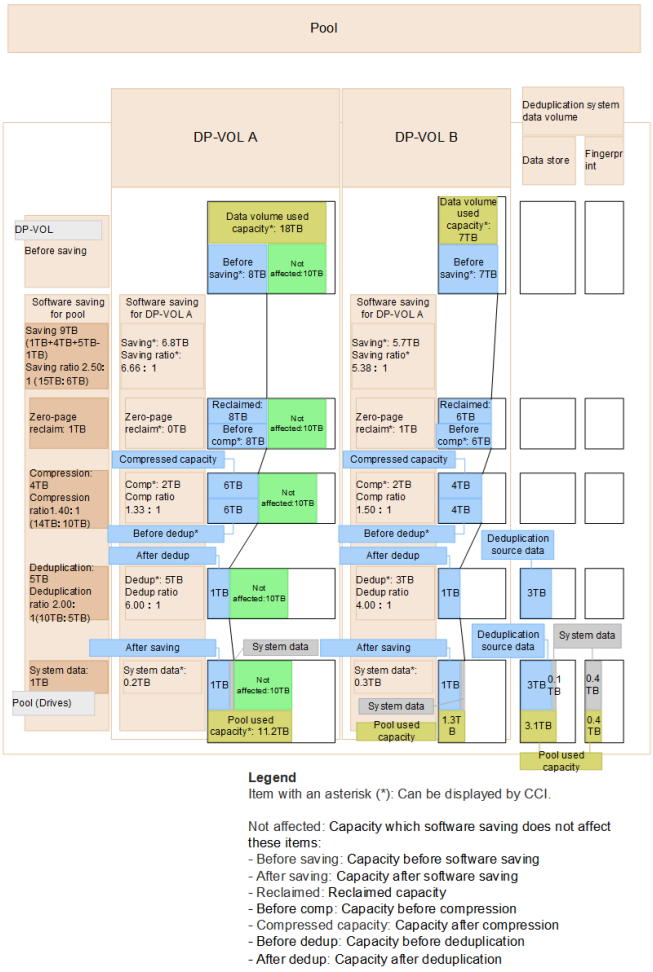
Capacity saving is supported in the pools for Dynamic Provisioning. The compression function and the deduplication function are set for each virtual volume in the pool for Dynamic Provisioning. Using compression with HDDs in Dynamic Provisioning or Dynamic Tieringpools is not recommended. Contact customer support for details.

Certain types of DP-VOLs have high capacity saving values. These types of DP-VOLs should not be used as the basis of comparison with other DP-VOLs. These conditions can cause abnormally high capacity saving values on DP-VOLs:

- DP-VOLs with large amounts of duplicated data have high capacity saving values due to the deduplication function.
- DP-VOLs that are replication targets of host-based or storage system-based replication usually have higher capacity saving values than DP-VOLs that are not replication targets.

This figure shows the capacity saving effect for pools and DP-VOLs. Items with an asterisk (*) in the figure are displayed

using CCI.



This table lists the CCI commands that show the capacity saving effect for each pool.

Command name	Output of command	**Item in the figure**
raidcom get pool -key total_saving	SE_SAVING(BLK)	Saving under Saving effect
	SES(%)	Saving ratio under Saving effect
	DATAVOL_USED(BLK)	Data volumes used capacity
raidcom get pool -key software_saving	PLS(%)	Saving ratio under Software saving
	PL_SAVING(BLK)	Saving under Software saving
	CMP(BLK)	Compression under Software saving
	DDP(BLK)	Dedup under Software saving
	RECLAIM(BLK)	Reclaim under Software saving

Command name	Output of command	**Item in the figure**
	SYSTEM(BLK)	System data under Software saving
	PL_PRE_USED(BLK)	Capacity before software saving
	PRE_CMP_USED(BLK)	Capacity before compression
	PRE_DDP_USED(BLK)	Capacity before deduplication
raidcom get pool -key fmc	FMC_LOG_USED(BLK)	FMD logical used capacity
	FMC_ACT_USED(BLK)	FMD used capacity
	FMC_PLV_USED(BLK)	FMD pool volumes used capacity

The HDvM - SN main window displays the following information about saving effects for the entire storage system.

Item	Description
Internal / External	This item switches the displayed items. • Total: Displays information on both the internal volumes and the external volumes. • Internal Only: Displays information on only the internal volumes. • External Only: Displays information on only the external volumes.
Open/Mainframe	This item switches the displayed capacity units. • Total: Displays all the open-systems, mainframe-systems and multi-platform volumes • Open Only: Displays only open-systems volumes • Mainframe Only: Displays mainframe-systems and multi-platform volumes.
Capacity Unit	This item switches the displayed units of the capacity. • Appropriate: Displays the capacity in appropriate unit depending on the capacity of each item. • TB/GB/MB: Displays the capacity of the specified unit.
Physical Summary	This item displays the capacity of physical logical devices and the number of devices. The information for open systems and mainframe systems is different. For details about this item, see the Physical Summary table.
Virtual Summary	This item displays capacity of virtual logical devices and the number of devices. The information for open systems and mainframe systems is different. For details about this item, see the table below about virtual summary.
Total Efficiency	This field is blank if the calculation for items below is not complete. A hyphen (-) is displayed if

Item	Description
	the information is not valid. For details, see the *Provisioning Guide*. • Total Efficiency: Displays the ratio of the total saving effect achieved by accelerated compression, capacity saving (compression and deduplication), snapshot, and Hitachi Dynamic Provisioning. The ratio of the system data is not included. • Data Reduction: Displays the data reduction ratio before and after performing the accelerated compression function and the capacity saving function (compression and deduplication). The ratio of the system data is not included. • Software Saving: Displays the capacity reduction ratio for data which is before and after performing the capacity saving function. The ratio of the system data is not included. ◦ Compression: Displays the capacity compression ratio for data which is before and after performing the capacity saving function. ◦ Deduplication: Displays the capacity deduplication ratio for data which is before and after performing the capacity saving function. ◦ Pattern Matching: Displays the capacity reduction ratio for data before and after performing pattern matching of the capacity saving function. • FMD Saving: Displays the capacity reduction ratio for data which is before and after performing the accelerated compression function. The ratio of the system data is not included. ◦ Compression: Displays the capacity compression ratio for data which is before and after performing the accelerated compression function. ◦ Pattern Matching: Displays the capacity reduction ratio for data before and after performing pattern matching of the accelerated compression function. • Snapshot: Displays the efficiency ratio achieved by snapshot. The ratio of the system data is not included. • Provisioning: Displays the efficiency ratio achieved by Hitachi Dynamic Provisioning. The ratio of the system data is not included. • Date and time for calculation: The start date and time and the end date and time for the calculation are displayed within the square brackets. **Note:** The date and time in the square brackets are the system date and time (date, time, and timezone) of the storage system. For Last Updated in HDvM - SN, the date and time based on the system date and time (date, time, and timezone) of the SVP are displayed. Therefore, if the settings of the system date and time for the storage system and the ones for the SVP are different, the date and time in the square brackets in this window and the date and time displayed for Last Updated are also different.
Total Saving (Software Deduplication, Software Compression, FMD Compression)	Total Saving: Displays the ratio and capacity reduced by the capacity saving function against all data in a storage system. When you use the capacity saving function, the saving ratio is calculated against metadata, garbage data, and parity data generated by the storage system in addition to user data. If the amount of used data volume before the capacity saving function is executed is smaller than

Item	Description
	the used pool capacity, a value which is invalid and smaller than the actually saved capacity might be displayed as the saved capacity. Software Deduplication: Displays the ratio reduced by the deduplication function against all data in a storage system. Software Compression: Displays the ratio reduced by the software compression function against all data in a storage system. FMD Compression: Displays the ratio reduced by the FMD compression function against all data in a storage system.
Total DP Subscription Rate	This item displays the ratio of virtual logical device capacity to physical logical device capacity in the pool volume of Dynamic Provisioning. For a Dynamic Provisioning pool in which Thin Image pairs are created, the virtual logical device capacity includes the snapshot data capacity.
Total Number of LDEVs	The number of LDEVs. The information displayed in this field depends on the type of system. See the following table.

This table lists the CCI commands that show the capacity saving effect for DP-VOLs.

Command name	Output of command	Item in the figure
raidcom get ldev	Used_Block(BLK)	Data volume used capacity
raidcom get ldev -key software_saving	TLS_R	Saving ratio of Software saving
	TOTAL_SAVING(BLK)	Saving of Software saving
	CMP(BLK)	Compression of Software saving
	DDP(BLK)	Deduplication of Software saving
	RECLAIM(BLK)	Zero page reclaim of Software saving
	SYSTEM(BLK)	System data of Software saving
	PRE_USED(BLK)	Capacity which software saving does not affect
	POOL_USED(BLK)	Pool used capacity

Total Efficiency Ratio and associated efficiency values

The total saving effect achieved by capacity saving, accelerated compression, Dynamic Provisioning, and snapshot is displayed as the Total Efficiency Ratio on the Pools window. System data (metadata and garbage data for the capacity saving function, Thin Image metadata) is not included in the Total Efficiency Ratio. These examples show the relationship between the total efficiency and the efficiency achieved by each function.
Note

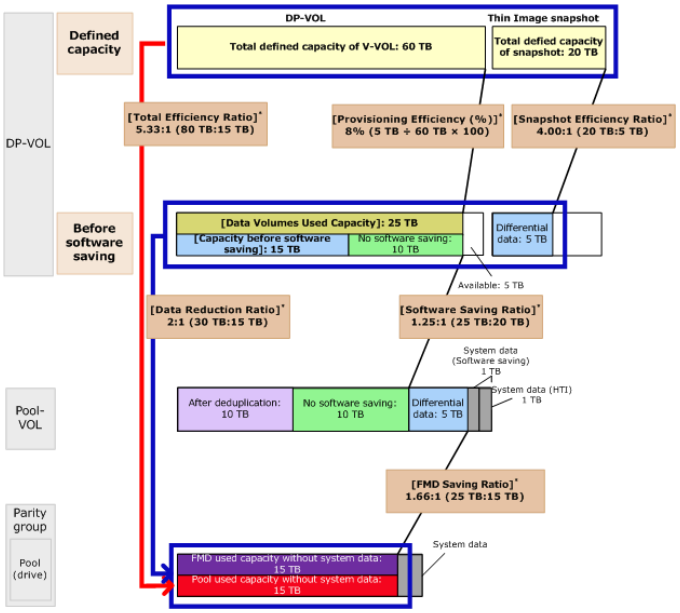
- The Total Efficiency Ratio and associated efficiency values do not include the total capacity and used capacity of these volumes:
  - Journal volumes
  - Volumes with the data direct mapping attribute
  - Command devices
  - Deduplication system data volumes (fingerprint and data store)
  - Thin Image S-VOLs
  - Quorum disks
  - Remote command devices
- When a pool is blocked, the values before the pool is blocked are displayed, and the value is updated after the pool is restored.
- When any of these operations is performed, the Total Efficiency Ratio and associated values displayed by HDvM - SN and CCI might be lower than the actual values:
  - Creating LDEVs or increasing the LDEV capacity
  - Initializing duplicated data
  - Reclaiming zero data page
  - Deleting or formatting LDEVs
  - Creating Thin Image snapshots
  - Deleting Thin Image S-VOLs
- The Total Efficiency Ratio and associated efficiency values change according to configuration changes and the type of I/O load. The efficiency values are more accurate when I/O loads are not high. You might obtain lower values immediately after performing operations.
- If data is written across the DP-VOL, the Total Efficiency Ratio might become less than 1.
- It takes up to 90 minutes to calculate the Total Efficiency Ratio and associated efficiency values.
- In a pool that contains accelerated compression-enabled parity groups, the data saving effect might not display correctly if the pool usage rate is less than 1%.
- If the total capacity of the pool volumes belonging to the parity group with accelerated compression enabled is smaller than the total capacity of the parity group, the data is always compressed on FMDs. 1.00 :1, which indicates no saving effect, might be displayed in Saving Effect in the Pools window.

  The following show when this occurs:
  - When the pool auto expansion function is disabled.
  - When the shrink pool operation is performed for some pool volumes in the parity group.

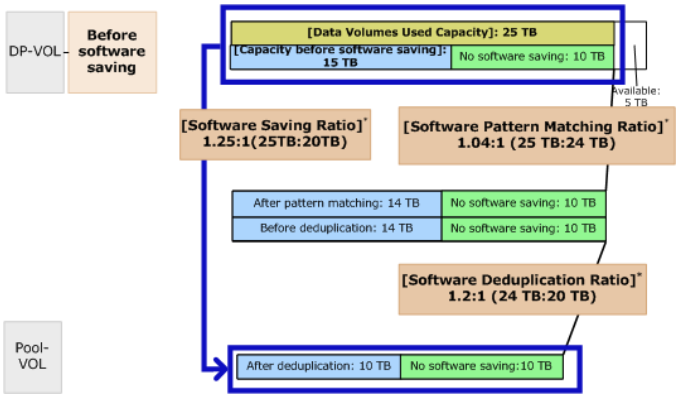
Example for deduplication and accelerated compression

These three figures show the Total Efficiency Ratio, Software Saving Ratio, and FMD Saving Ratio values achieved when deduplication and accelerated compression are used. In these figures, items enclosed by the square brackets are displayed by HDvM - SN, and items with an asterisk are displayed using CCI.

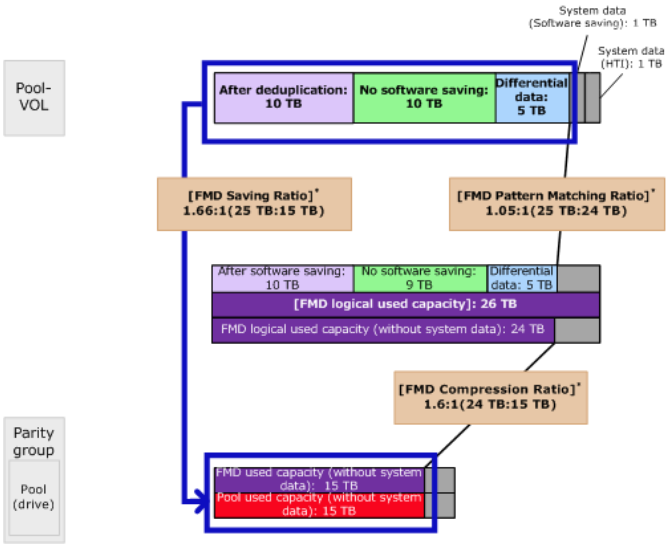
Total Efficiency Ratio for deduplication and accelerated compression



Software Saving Ratio for deduplication and accelerated compression



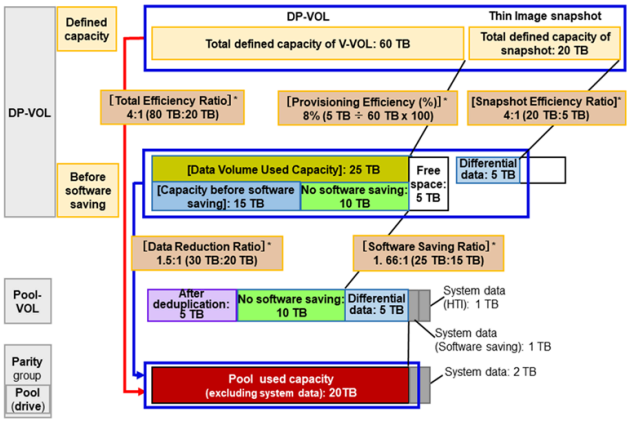
FMD Saving Ratio for deduplication and accelerated compression



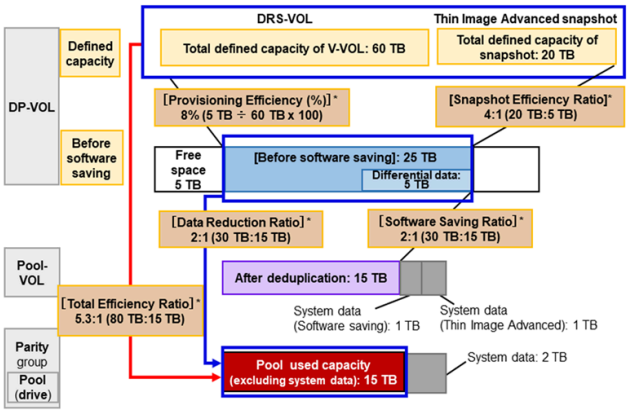
Example for deduplication and compression of the capacity saving function

These two figures show the Total Efficiency Ratio and Software Saving Ratio achieved when deduplication and compression of the capacity saving function are used. In these figures, items enclosed by the square brackets are displayed by HDvM - SN, and items with an asterisk are displayed using CCI.

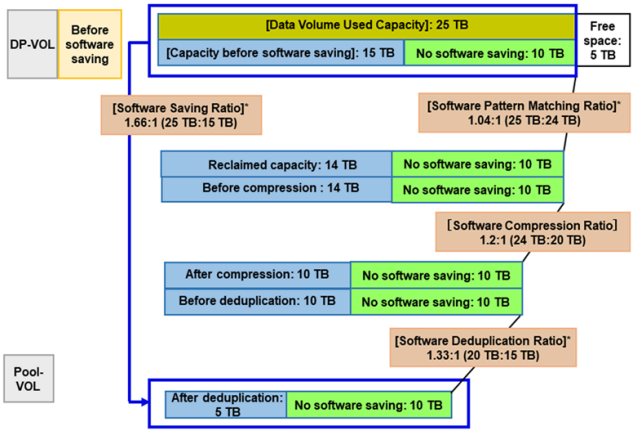
Total Efficiency Ratio for controller-based deduplication and compression on DRD-VOLs



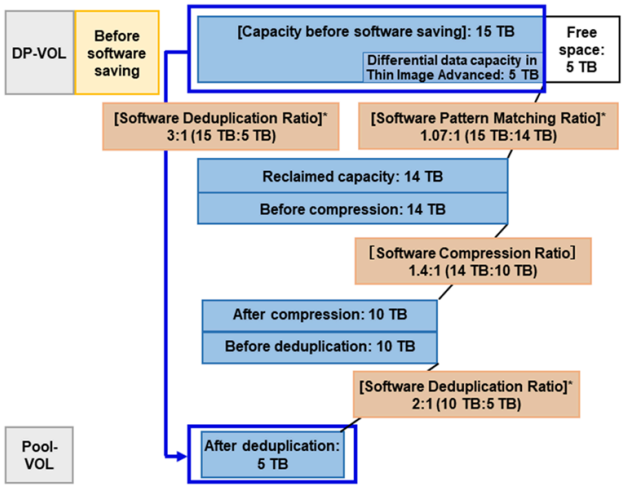
Total Efficiency Ratio for controller-based deduplication and compression on DRS-VOLs



Software Saving Ratio for controller-based deduplication and compression on DRD-VOLs



Software Saving Ratio for controller-based deduplication and compression on DRS-VOLs



Efficiency values

The raidcom get system -key efficiency CCI command displays the efficiency values for the entire storage system, and the raidcom get pool -key efficiency CCI command displays the efficiency values for the specified pool.

Item output by CCI	Corresponding item displayed by HDvM - SN
TOTAL_EFF_R	Total Efficiency Ratio
TLS_R	Data Reduction Ratio
PLS_R	Software Saving Ratio
PLS_CMP_R	Software Compression Ratio
PLS_DDP_R	Software Deduplication Ratio
PLS_RECLAIM_R	Software Pattern Matching Ratio
FMD_SAVING_R	FMD Saving Ratio
FMD_CMP_R	FMD Compression Ratio
FMD_RECLAIM_R	FMD Pattern Matching Ratio
SNAPSHOT_EFF_R	Snapshot Efficiency Ratio
PROVISIONING_EFF(%)	Provisioning Efficiency (%)

NoteThe following values are the maximum of each saving effect ratio available to display in HDvM - SN or CCI:

- HDvM - SN: 99999999999999.99
- CCI: 92233720368547758.07 (a hundredth value of hexadecimal 0x7FFFFFFFFFFFFFFF)

If no data is stored in the pool as shown in the following example, the saving effect ratio displays the maximum available value. If data is stored, the saving effect ratio calculated from the amount of actual data reduction is displayed. The following examples show the conditions where the maximum values are displayed:

Example 1. Total Efficiency Ratio after the pool and DP-VOL creations but before data writing

- HDvM - SN: 99999999999999.99
- CCI: 92233720368547758.07

Example 2. Snapshot Efficiency Ratio applied when the pool includes no Thin Image pairs but only Thin Image Advanced pairs in the PAIR status or ones with no differential data

- HDvM - SN: 99999999999999.99
- CCI: 92233720368547758.07

The GET <base URL>/v1/objects/total-efficiencies/instance REST API request displays the efficiency values for the entire storage system, and the GET <base URL>/v1/objects/pools/<object ID>?detailInfoType=efficiency request displays the efficiency values for the specified pool.

Attribute	Corresponding item displayed by HDvM - SN
totalRatio	Total Efficiency Ratio
compressionRatio	Data Reduction Ratio
dedupeAndCompression.totalRatio	Software Saving Ratio
dedupeAndCompression.compressionRatio	Software Compression Ratio
dedupeAndCompression.dedupeRatio	Software Deduplication Ratio
dedupeAndCompression.reclaimRatio	Software Pattern Matching Ratio
acceleratedCompression.totalRatio	FMD Saving Ratio
acceleratedCompression.compressionRatio	FMD Compression Ratio
acceleratedCompression.reclaimRatio	FMD Pattern Matching Ratio
snapshotRatio	Snapshot Efficiency Ratio
provisioningRate	Provisioning Efficiency (%)

Next steps

Use deduplication and compression to reduce the capacity used by your data