WHITE PAPER

SAP HANA Backup Storage Data Reduction with Capacity Saving Features in Hitachi Virtual Storage Platform Gx00

Lab Validation Report

By Milind Pathak

March 2017

Feedback

Hitachi Data Systems welcomes your feedback. Please share your thoughts by sending an email message to SolutionLab@hds.com. To assist the routing of this message, use the paper number in the subject and the title of this white paper in the text.

Revision History

Revision	Changes	Date
AS-591-00	Initial release	March 31, 2017

Table of Contents

SAP HANA Backup Storage Data Reduction with Capacity Saving Features in Hitachi Virtual Storage Platform Gx00

Lab Validation Report

Read how you can take advantage of compression and deduplication performed during backup and recovery for SAP HANA data using Hitachi Virtual Storage Platform G600 (VSP G600). See the capacity saving achieved on SAP HANA database backups when enabling compression and deduplication. Learn the impact on the time required to restore using a compressed backup. These results for a SAP HANA infrastructure are valid for all Virtual Storage Platform Gx00 models.

This lab validation report documents the savings and performance impact of the capacity saving features in Hitachi Storage Virtualization Operating System (SVOS) on SAP HANA, an in-memory and columnar store database. Columnar store allows for the efficient compression of data. This makes it less costly for the SAP HANA database to keep data in main memory. Details about SAP HANA compression can be found in Data Compression in the Column Store.

Storage Virtualization Operating System, used in Hitachi Virtual Storage Platform Gx00, offers data reduction features which can help increase the usable capacity of the storage array. The capacity saving functions include data deduplication and data compression to reduce your bitcost for the stored data. The controllers in the storage system perform the data deduplication and compression.

- **Deduplication** deletes duplicate copies of data written to different addresses in the same pool to maintain only a single copy of the data at one address.
- **Data Compression** uses the LZ4 compression algorithm to compress the data. Compression can be enabled for each dynamic provisioning volume (DP-VOL).

Since data is efficiently compressed already, capacity saving on the primary SAP HANA database storage is lower than for traditional databases. However, deduplication and compression can provide significant benefits for backup storage of SAP HANA database without any direct performance impact.

This lab validation report documents results of tests performed to identify

- Capacity savings
- Impact on restore time (when restoring from compressed backups)
- Impact on storage performance

Note — Testing of this configuration was in a lab environment. Many things affect production environments beyond prediction or duplication in a lab environment. Follow the recommended practice of conducting proof-of-concept testing for acceptable results in a non-production, isolated test environment that otherwise matches your production environment before your production implementation of this solution.

Product Features

These products are a part of the tested configuration.

Hitachi Storage Virtualization Operating System

Hitachi Storage Virtualization Operating System (SVOS) spans and integrates multiple platforms. It integrates storage system software to provide system element management and advanced storage system functions. Used across multiple platforms, Storage Virtualization Operating System includes storage virtualization, thin provisioning, storage service level controls, dynamic provisioning, and performance instrumentation.

Storage Virtualization Operating System includes standards-based management software on a Hitachi Command Suite (HCS) base. This provides storage configuration and control capabilities for you.

Storage Virtualization Operating System uses Hitachi Dynamic Provisioning (HDP) to provide wide striping and thin provisioning. Dynamic Provisioning provides one or more wide-striping pools across many RAID groups. Each pool has one or more dynamic provisioning virtual volumes (DP-VOLs) without initially allocating any physical space. Deploying Dynamic Provisioning avoids the routine issue of hot spots that occur on logical devices (LDEVs).

To enable essential management and optimization functions, this solution uses Hitachi Storage Navigator (SN), a part of Storage Virtualization Operating System. Storage Navigator runs on most browsers. A command line interface is available.

Hitachi Virtual Storage Platform Gx00 Models

Hitachi Virtual Storage Platform Gx00 models are based on industry-leading enterprise storage technology. With flash-optimized performance, these systems provide advanced capabilities previously available only in high-end storage arrays. With the Virtual Storage Platform Gx00 models, you can build a high performance, software-defined infrastructure to transform data into valuable information.

Hitachi Storage Virtualization Operating System provides storage virtualization, high availability, superior performance, capacity saving and advanced data protection for all Virtual Storage Platform Gx00 models. This proven, mature software provides common features to consolidate assets, reclaim space, extend life, and reduce migration effort.

Hitachi Compute Blade 2500

Hitachi Compute Blade 2500 delivers enterprise computing power and performance with unprecedented scalability and configuration flexibility. Lower your costs and protect your investment.

Flexible I/O architecture and logical partitioning allow configurations to match application needs exactly with Hitachi Compute Blade 2500. Multiple applications easily and securely co-exist in the same chassis.

Add server management and system monitoring at no cost with Hitachi Compute Systems Manager. Seamlessly integrate with Hitachi Command Suite in Hitachi storage environments.

Software Products

These are the key software products used in this solution.

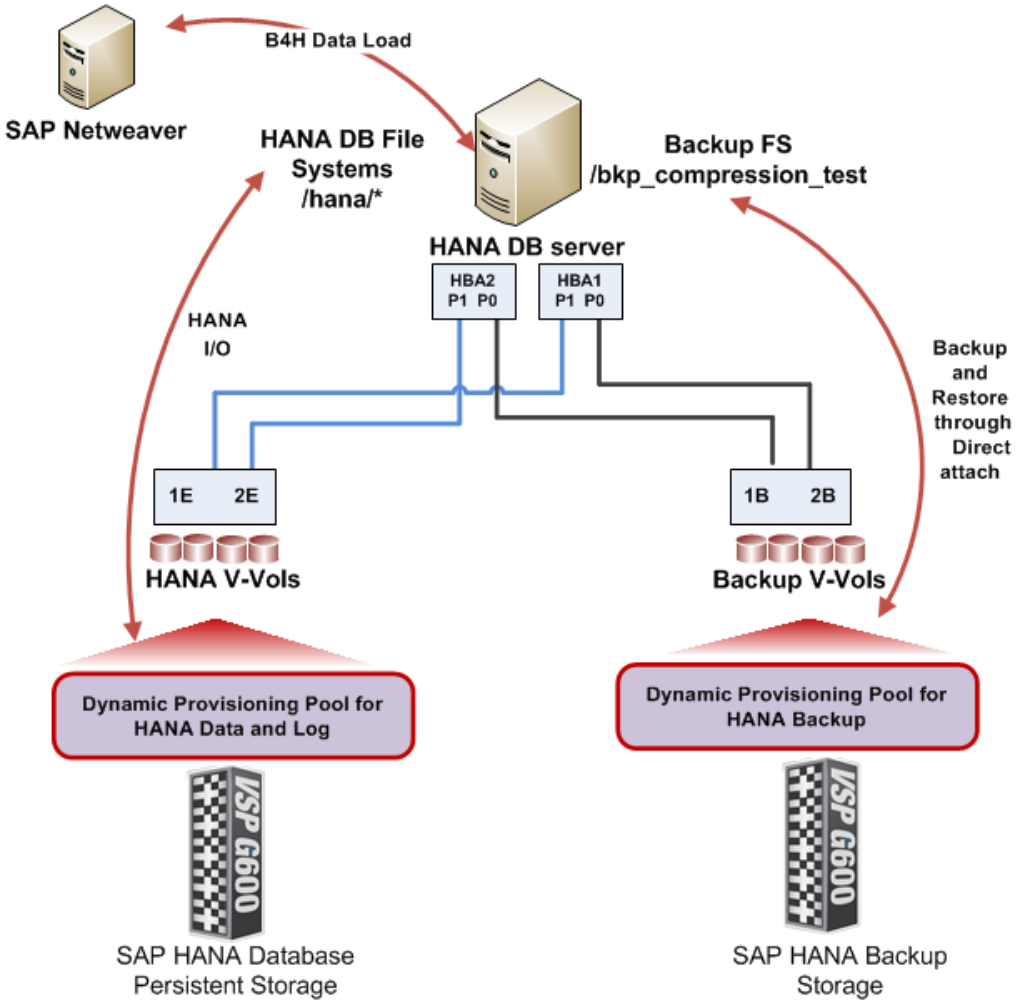
TABLE 1. SOFTWARE PRODUCTS USED IN THE TEST ENVIRONMENT

Software		Version
Operating System	SUSE Linux Enterprise Server for SAP Applications	SLES12
SAP HANA		1.0 SPS12
Hitachi Storage Virtualization Operating System, which includes the following: ■ Hitachi Dynamic Provisioning ■ Hitachi Storage Navigator		83-04-20-40/00

Test Environment Configuration

Figure 1 shows the test environment configuration with Hitachi Compute Blade 2500 using 520X B3 server blades and Hitachi Virtual Storage Platform G600 hosting SAP HANA.

Figure 1



Server Configuration

The test configuration of the 520X B3 server blades used in Hitachi Compute Blade 2500 was the following:

- 2 Intel Xeon E7-8880 processors
- 2.20 GHz processor frequency
- 44 processor cores
- 2 × 2-port 10GBASE-SR LAN PCIe adapters on two I/O board modules
- 2 Hitachi 16 Gb/sec 2-port Fibre Channel adapters on two I/O board modules

There were two 520X B3 server blades used for the following when testing:

- First server blade for the SAP B4H benchmarking server, for loading data into the SAP HANA database
- Second server blade for the SAP HANA database server

SAP HANA database server connected separately to two Hitachi Virtual Storage Platform G600 storage arrays, as shown in Figure 1 on page 4.

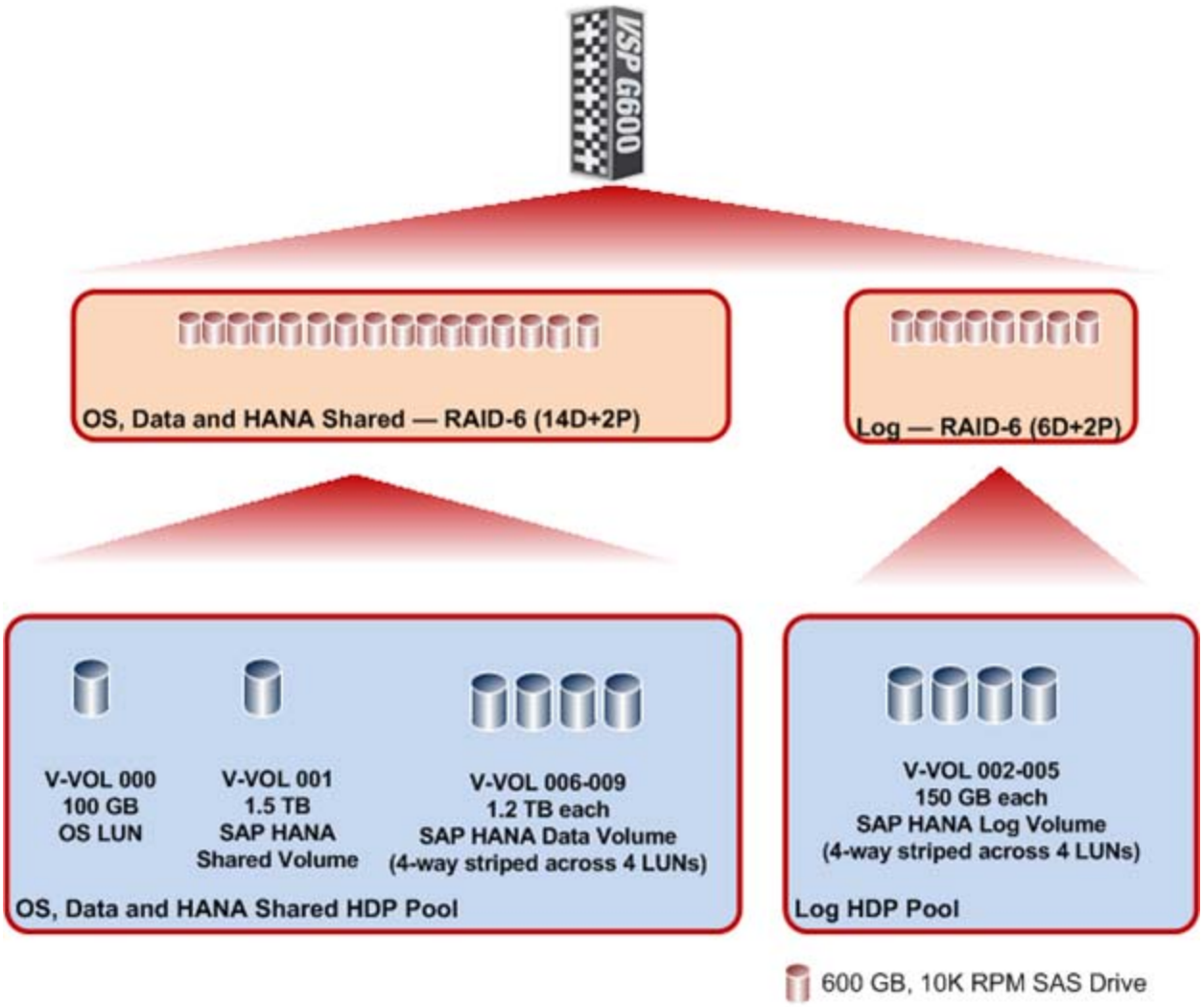
Storage Configuration

The test environment used two different Hitachi Virtual Storage Platform G600 storage arrays for SAP HANA database persistent storage and backup storage.

Using a dynamic provisioning pool design in the configuration is a prerequisite for capacity-saving with Virtual Storage Platform Gx00.

Figure 2 on page 6 shows the configuration of SAP HANA database persistent storage.

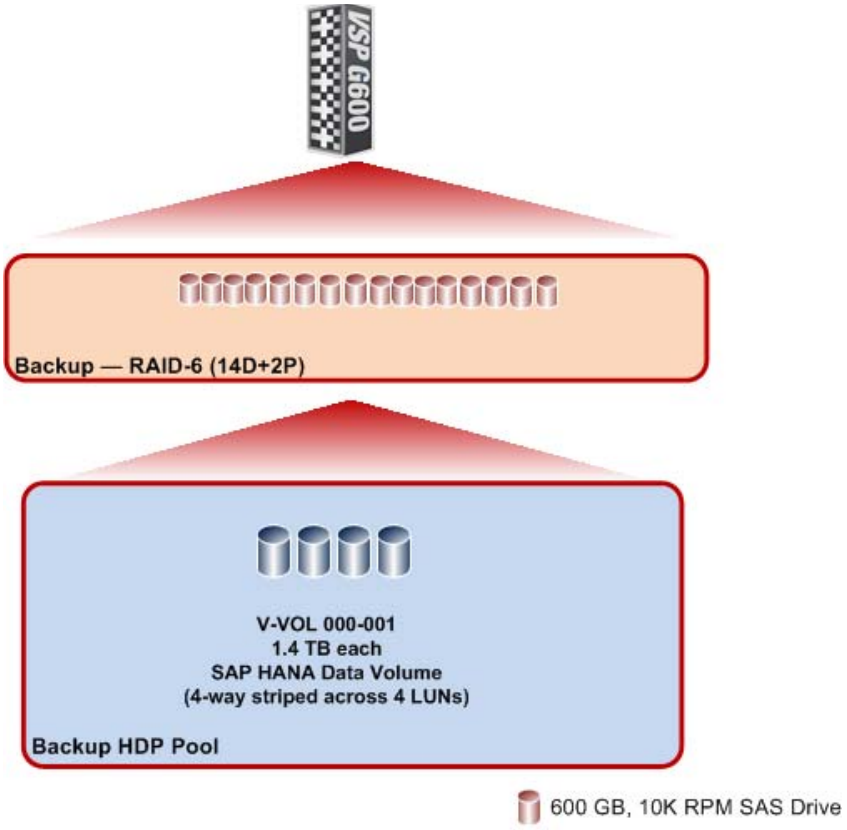
Figure 2



SAP HANA Persistent Storage

Figure 3 shows the configuration of SAP HANA database backup storage.

Figure 3



Backup Storage

Test Methodology

The objective of this test was to identify the following:

- The capacity savings on a SAP HANA database backup
- Any performance impact of enabling capacity savings on backup and restore

The following test procedure was used:

1. Install a SAP HANA scale-up system on Hitachi Compute Blade 2500 using 520X B3 server blades. Follow the TDI storage design using dynamic provisioning pools created with Hitachi Dynamic Provisioning. More details can be found in SAP HANA Tailored Data Center Integration on Hitachi Virtual Storage Platform Gx00 using Hitachi Dynamic Provisioning Pools Reference Architecture Guide (AS-490-00, PDF).
2. Setup SAP BW edition for SAP HANA Standard Application Benchmark on the other 520X B3 server blade in Hitachi Compute Blade 2500.
3. Establish benchmark values. The benchmark consists of 3 phases:
   (1) Data load phase
   (2) Query throughput phase
   (3) Query runtime phase

   Allowed data volumes are a multiple of 1.3 billion initial records.
4. Load data into the SAP HANA database using the benchmarking tool. For this test, 725 GB of data was loaded into the database.
5. Create a dynamic provisioning pool and V-VOLs on the backup Hitachi Virtual Storage Platform G600 and attach the pools to a SAP HANA database server. Create the backup file system using these V-VOLs.
6. Perform backup and restore from the SAP HANA database server in three phases:
   (1) Before enabling compression on backup V-VOLs (baseline values)
   (2) During the post-processing of compression while it is in progress
   (3) After the compression post-processing finishes

   Measure the time taken for each operation and compare the results in Step After the compression post-processing finishes with the baseline measurements taken in in Step Before enabling compression on backup V-VOLs (baseline values).
7. Observe the capacity savings after compression post-processing finishes.
8. Create a deduplication system data volume in the dynamic provisioning pool of the backup storage.
9. Enable deduplication on the backup storage V-VOLs, also.
10. Observe the time taken for deduplication post-processing.
11. Observe the overall capacity savings after post-processing for compression and deduplication is finished.

Analysis

Based on the tests performed on the scale-up SAP HANA database system containing data loaded using the SAP B4H benchmarking tool, following observations were made:

1. There is no impact on storage level post-processing on timelines for the SAP HANA database backup and restore.
2. Restore time was about 10% higher when the database was restored from compressed backup.
3. Storage performance monitor showed about 10% MPU utilization during the compression post-processing and in the range of 10-30% utilization during the dedupe post-processing.
4. Capacity saving after enabling compression on backup storage was observed about 19% saving and 965 GB. This was for 7 × 725 GB backups on a 5.5 TB file system and 5 TB used space.
5. Capacity saving after enabling compression and dedupe on backup storage were observed about 75% saving and 3.76 TB. This was for 7 × 725 GB backups on a 5.5 TB file system and 5 TB used space.

Compression and dedupe savings depend on the dataset present in the database. It cannot be assured that the results achieved during testing in the Hitachi Data Systems lab will be consistent at any other implementation. For example, the dedupe savings demonstrated here was high because the same database was used to create seven backups (high duplications).

Dedupe savings depend on a change between all the backups retained on backup storage at that point of time. Fewer changes result in higher dedupe savings.

Storage resources and time required for post-processing also depends on how much compression or deduplication can be dome on the given dataset. With more compression and dedupe, more storage resources and time are required.

Test Results

These are the test results from testing by Hitachi Data Systems.

Backup and Restore Operation Timelines

Using the procedure described in “Test Methodology” on page 8, seven full database backups and six restores were performed at different times. The same data was used for all the database backups.

Table 2, “Test Results (Before Compression),” on page 9, Table 3, “Test Results (During Compression),” on page 10, and Table 4, “Test Results (After Compression),” on page 10 provide the time taken for backup and restore operations during different phases of compression. Backup throughout has been measured using SAP HANA Studio.

TABLE 2. TEST RESULTS (BEFORE COMPRESSION)

	Backup 1	Backup 2	Backup 3	Restore 1	Restore 2
Time Taken	17 mins	21 mins	18 mins	20 mins	21 mins
Backup Throughput	689 MB/s	565 MB/s	681 MB/s		

TABLE 3. TEST RESULTS (DURING COMPRESSION)

	Backup 1	Backup 2	Restore 1
Time Taken	17 mins	18 mins	20 mins
Backup Throughput	690 MB/s	676 MB/s	

TABLE 4. TEST RESULTS (AFTER COMPRESSION)

	Backup 1	Backup 2	Restore 1	Restore 2	Restore 3
Time Taken	17 mins	18 mins	23 mins	23 mins	24 mins
Backup Throughput	690 MB/s	676 MB/s			

Capacity Savings

There were three tests run to compare capacity savings.

Initial Statistics for Backup Storage Dynamic Provisioning Pool

Table 5 shows the initial state of the backup storage dynamic provisioning pool created with Hitachi Dynamic Provisioning.

TABLE 5. INITIAL STATE OF COMPRESSION BACKUP POOL 1

Status	Tier Management
Pool Capacity (Used/Total)	5.46 TB/5.85 TB [93%]
Physical Pool Capacity (Used/Total)	- / - [-%]
Saving Effect	0% (0.00 MB)
V-VOL Capacity (Used/Total)	5.46 TB/5.46 TB [100%]
Subscription (Current/Limit)	93%/Unlimited
User-Defined Threshold (Warning/Depletion)	100%/100%

Capacity Savings after Enabling Compression

Table 6 shows the capacity savings after enabling compression on V-VOLs of backup storage dynamic provisioning pool. Capacity saving on backup storage after 7 × 725 GB backups on a 5.5 TB file system (5 TB used space) were observed about 19% saving and 965 GB.

TABLE 6. STATE OF COMPRESSION BACKUP POOL 1 AFTER ENABLING COMPRESSION

Status	Tier Management
Pool Capacity (Used/Total)	3.98 TB/5.85 TB [68%]
Physical Pool Capacity (Used/Total)	- / - [-%]
Saving Effect	19% (965.20 GB)
V-VOL Capacity (Used/Total)	4.95 TB/5.46 TB [90%]
Subscription (Current/Limit)	93%/Unlimited
User-Defined Threshold (Warning/Depletion)	100%/100%

Capacity Savings After Enabling Compression and Dedupe

Table 7 shows the capacity savings after enabling compression and dedupe on the V-VOLs of backup storage dynamic provisioning pool. Capacity savings on backup storage after 7 × 725 GB backups on a 5.5 TB file system (5 TB used space) were observed at about 75% savings and 3.76 TB.

TABLE 7. STATE OF COMPRESSION BACKUP POOL 1 AFTER ENABLING COMPRESSION AND DEDUPE

Status	Tier Management
Pool Capacity (Used/Total)	1.98 TB/5.85 TB [20%]
Physical Pool Capacity (Used/Total)	- / - [-%]
Saving Effect	75% (3.76 TB)
V-VOL Capacity (Used/Total)	4.96 TB/45.46 TB [10%]
Subscription (Current/Limit)	776%/Unlimited
User-Defined Threshold (Warning/Depletion)	100%/100%

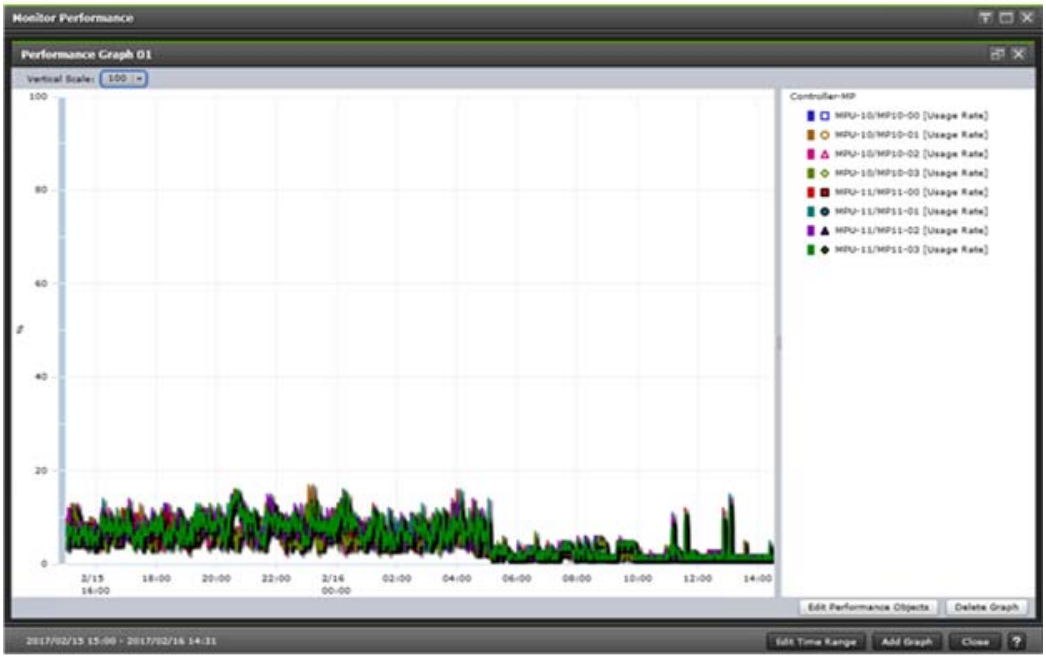
Storage Performance

Figure 4 on page 12 and Figure 5 on page 12 shows sample performance data for MPU utilization during the post-processing of compression (Figure 4) and dedupe (Figure 5). When the post-processing for compression or dedupe was going on, no other activities were in progress on the storage. Refer to Performance Guide for VSP Gx00 and Fx00 to understand more about Virtual Storage Platform Gx00 storage performance.

Note — Performance data captured in the Figure 4 and Figure 5 is only sample data from specific MPUs of the complete data set. It does not cover all MPUs of the storage. However, similar utilization was seen on other MPUs, as well.

- MPU utilization during compression post-processing

Figure 4



- MPU utilization during dedupe post-processing

Figure 5

For More Information

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

Live and recorded product demonstrations are available for many Hitachi products. To schedule a live demonstration, contact a sales representative. To view a recorded demonstration, see the Resources website.

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

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

Hitachi Data Systems

Corporate Headquarters
2845 Lafayette Street
Santa Clara, CA 95050-2639 USA
www.HDS.com community.HDS.com

Regional Contact Information
Americas: +1 866 374 5822 or info@hds.com
Europe, Middle East and Africa: +44 (0) 1753 618000 or info.emea@hds.com
Asia Pacific: +852 3189 7900 or hds.marketing.apac@hds.com

AS-591-00, March 2017.