HITACHI
Inspire the Next

Testing 250 Virtual Machines Performance During Dedupe Operations on Hitachi NAS Platform

TECHNICAL PAPER

Tech Note

By Daniel Worden; Jeff Chen

August 2015

Hitachi Data Systems

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.

Contents

Testing 250 Virtual Machines Performance During Dedupe Operations on Hitachi NAS Platform

Tech Note

This paper has the following for Hitachi NAS Platform used in a VMware environment:

- Recommended configurations for certain virtual machine counts and expected performance (IOPS and latency) using current best practice deployments. There were 250 virtual machines per cluster tested.
- Virtual machines performance results during the dedupe operation testing on Hitachi NAS Platform at scale through the virtual machine life cycle, including the following:
  - Performance measurement of the virtual machines application performance before starting the dedupe operations
  - Performance measurement of the virtual machines application performance after starting and running dedupe

Hitachi NAS Platform is an enterprise NFS solution for VMware that delivers a level of scalable and predictable performance (IOPS and latency). In addition to other benefits, NAS Platform also delivers capacity efficiency and data protection for VMware cloud environments.

During testing, all workloads performed as expected during the dedupe operation, without loss of I/O on NAS Platform. Workloads included virtual machines running the following:

- Microsoft® SQL Server® and Microsoft Exchange on Microsoft Windows Server®.
- Web server and OLTP transactions on SUSE Linux Enterprise Server 11

Table 1 describes the test cases.

Table 1. Test Cases Verified in This Document

Test Case	Pass/fail Criteria	Result
Mixed workload application performance without a dedupe operation on Hitachi NAS Platform running (establish a baseline)	Virtual machines reach and maintain the configured I/O target	Pass
Mixed workload application performance was not impacted during the dedupe operation on Hitachi NAS Platform	Virtual machines reach and maintain the configured I/O target configured during the dedupe operation on NAS Platform	Pass

These are the tests that were run:

- Test execution against 250 virtual machines without dedupe running for a 6 hour test to establish a baseline.
- Test execution against 250 virtual machines with dedupe running for a 12 hour test to compare to the baseline.

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.

Use Case Overview

During testing, 250 virtual machines ran a variety of workload profiles. The workloads were divided into tiles of 25 virtual machines each. Each workload profile had a light, medium, and heavy component. The virtual machines in each tile were configured for an average of 25 IOPS. The workload profile within each tile is described Table 2.

Table 2. Virtual Machine Workload Profile for Each Tile

Workload	Weight	IOPS	VMDK Size	Number of Virtual Machines	Total IOPS	Total VMDK Size
Microsoft SQL Server	Light	10	60 GB	3	30	180 GB
Web Server	Light	10	50 GB	6	60	300 GB
Microsoft Exchange	Light	10	60 GB	3	30	180 GB
OLTP	Light	10	60 GB	3	30	180 GB
Microsoft SQL Server	Medium	25	200 GB	1	25	200 GB
Web Server	Medium	25	80 GB	2	50	160 GB
Microsoft Exchange	Medium	25	100 GB	1	25	100 GB
OLTP	Medium	25	100 GB	1	25	100 GB
Microsoft SQL Server	Heavy	70	40 GB	1	70	40 GB
Web Server	Heavy	70	40 GB	2	140	80 GB
Microsoft Exchange	Heavy	70	40 GB	1	70	40 GB
OLTP	Heavy	70	40 GB	1	70	40 GB
Total		**25 per virtual machine**		**25**	**625**	**1600 GB**

The total VMDK size includes a 20 GB VMDK for the operating system. The remainder of the VMDK size was used as a VMDK for a data volume. Each virtual machine had the following resources assigned:

- 1 vCPU
- 4 GB of RAM.

Table 3 shows the I/O profile for each workload.

Table 3. Workload I/O Profiles

Workload	I/O Size	Percent Random	Percent Read
Microsoft® SQL Server®	64 KB	100%	66%
Web Server	8 KB	75%	95%
Microsoft Exchange	8 KB	80%	55%
OLTP	8 KB	100%	70%

Tested Components

Table 4 lists the specific hardware components used during testing.

Table 4. Tested Hardware Components

Hardware	Description	Version	Quantity
Hitachi Unified Storage VM	■ Dual controllers ■ 16 × 8 Gb/sec Fibre Channel ports ■ 64 GB cache memory ■ 128 × 600 GB 10k RPM SAS disks, 2.5 inch SFF ■ 4 x 2.5 TB flash module drives	73-03-06-00/00	1
Hitachi NAS Platform 4100	■ 2 × 10 Gb/sec cluster ports ■ 4 × 10 Gb/sec Ethernet ports ■ 4 × 8 Gb/sec Fibre Channel ports	12.3.3826.03	1
Hitachi Compute Blade 500 chassis	■ 8-blade chassis ■ 2 Brocade 5460 Fibre Channel switch modules, each with 6 × 8 Gb/sec uplink ports ■ 2 Brocade VDX 6746 Ethernet switch modules, each with 8 × 10 Gb/sec uplink ports ■ 2 management modules ■ 6 cooling fan modules ■ 4 power supply modules	SVP: A0170-D-8920 5460: FOS 7.0.2C VDX6746: NOS 4.1.2	1
520H B2 server blade	■ Half blade ■ 2 × 12-core Intel Xeon E5-2697 processors, 2.70 GHz ■ 256 GB RAM ■ 16 × 16 GB DIMMs	BMC/EFI:01-29	2
Brocade 6510 switch	■ SAN switch 48 × 8 Gb/sec Fiber Channel ports	FOS 7.0.1a	2
Brocade VDX 6740 switch	■ Ethernet switch with 40 × 10 Gb/sec ports	NOS 4.1.2	2

Table 5 lists the specific software components used during testing.

Table 5. Tested Software Components

Software	Version
Hitachi Storage Navigator Modular	Microcode Dependent
VMware vCenter server	5.5.0 U2
VMware Virtual Infrastructure Client	5.5.0 U2
VMware ESXi	5.5.0 U2
VDbench	5.04
Microsoft Windows Server 2008 R2 (Microsoft SQL Server virtual machine operating system)	
Microsoft Windows Server 2012 R2 (Microsoft Exchange Server virtual machines operating system)	
SUSE Linux Enterprise Server 11 SP2 (OLTP and Web Server virtual machines operating system)	

High Level Test Infrastructure

Figure 1 illustrates the storage configuration for NFS storage testing on Hitachi NAS Platform.

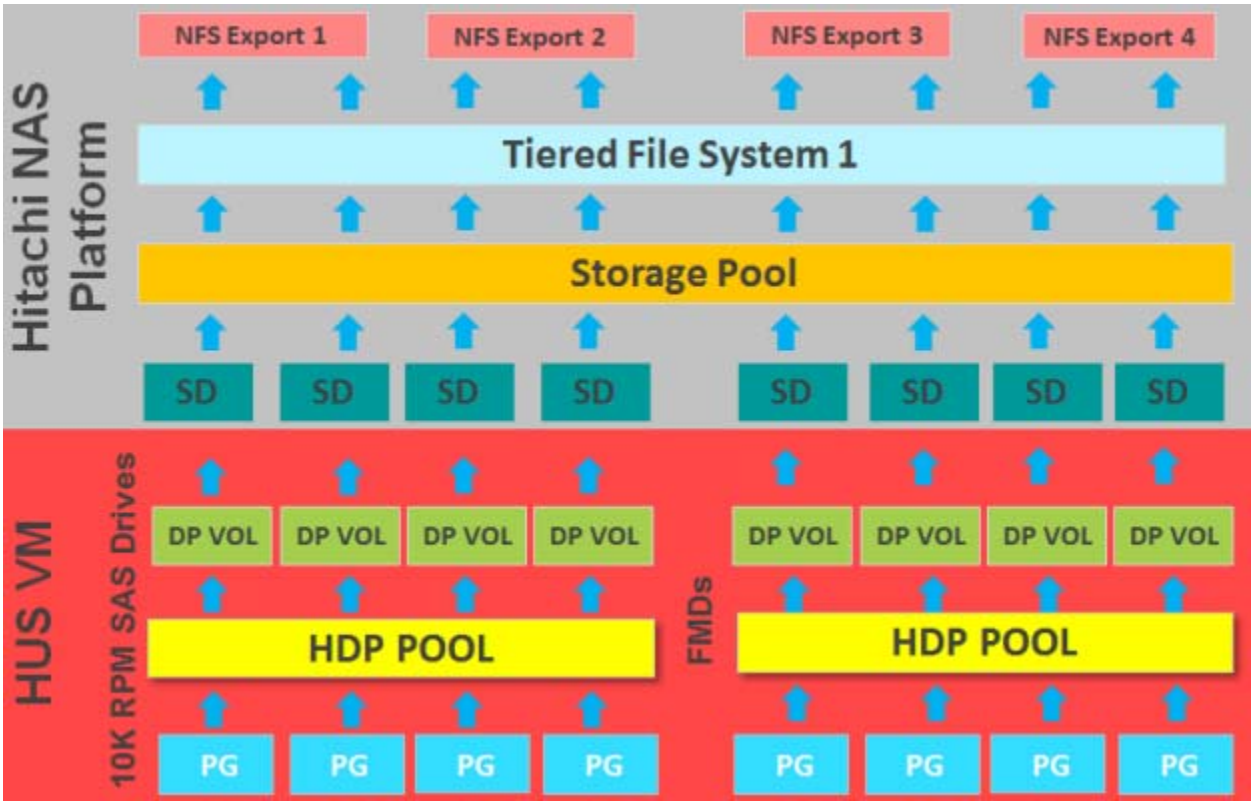

Figure 1

The Hitachi Unified Storage VM storage system was configured with 10k RPM SAS drives, grouped and presented as the following:

- 13 RAID-6 (6D+2P) parity groups
- 1 dynamic provisioning pool was created from the 13 parity groups
- 16 × 8 TB thin provisioned DP-VOLs were created from the single dynamic provisioning pool

The Unified Storage VM was configured with flash module drives grouped and presented as the following:

- 1 RAID-5 (3D+1P) parity group
- 1 dynamic provisioning pool was created from the single parity group
- 8 × 4 TB thin provisioned DP-VOLs were created from the single dynamic provisioning pool

The following configuration was used for Hitachi NAS Platform:

- 16 system drives were created from the 16 DP-VOLs that were provisioned from the 10k RPM SAS drives
- 8 system drives were created from the 8 DP-VOLs from the flash module drives
- One storage pool with a tiered file system was created from the 24 system drives as listed below.
  - **Stripeset 0 (Tier 0 - metadata)** — 8 flash module drives with 4 TB system drives
  - **Stripeset 1 (Tier 1)** — 8 × 10k SAS 8 TB SDs
  - **Stripeset 2 (Tier 1)** — 8 × 10k SAS 8 TB SDs

  This storage pool creation pattern takes future file system expansion into consideration. The file system should be expanded with eight SDs at a time.
- One file system was created from the storage pool.
- 4 NFS exports were created as mount points for the VMware ESXi datastore.

The ability to use thin provisioned DP-VOLs as Hitachi NAS Platform system drives was introduced in version 12.1.3613.10 software. Since this version, thin provisioned DP-VOLs are recommended for the following reasons:

- It is quicker and easier to expand a NAS Platform file system.
- Data rebalancing is automatically performed by the dynamic provisioning pool more efficiently
- Prior to the support of thin-provisioned system drives, new system drives had to be presented to NAS Platform and the storage pool had to be expanded by a stripe set. Then, the file system had to be rebalanced.

When using dynamically provisioned thin-provisioned volumes with Hitachi NAS Platform, it is a best practice to not over-provision more than three times the actual available space.

Eight DP-VOLs were created from a parity group of flash module drives, because flash module drives are fast enough to handle the additional I/O threads when compared to mechanical spindle-based disks.

Super flush was configured on all system drives on Hitachi NAS Platform with the setting of 3 wide by 128 KB (3 × 128). These settings are the best practice when using super flush on NAS Platform with Hitachi Unified Storage VM.

All best practices were followed when configuring Hitachi Unified Storage VM and Hitachi NAS Platform.

For information on Ethernet networking configuration recommendations, see Deploy Hitachi Unified Compute Platform Select for VMware vSphere using Hitachi NAS Platform with Hitachi Unified Storage VM.

Test Result

These are the test results.

NFS Datastore Baseline on Hitachi NAS Platform

The first test established a baseline on Hitachi NAS Platform without dedupe running.

- 10 tiles (250 virtual machines) were distributed evenly across four datastores.
- The 10 tiles were provisioned as thin VMDKs, which used 8.4 TB of disk space before the dedupe operation was used.

The 10 tiles ran for 6 hours to establish the baseline for comparison (Table 6).

Table 6. First Test Case Results

Test Case	Description	Result
Mixed workload application performance without a dedupe operation running on Hitachi NAS Platform to establish a baseline	Run workload on 10 tiles of 250 virtual machines for 6 hours	Passed
	Expected result: Virtual machines reach and maintain the I/O target configured for them.	

The tiles were distributed across the ESXi host as shown in Table 7.

Table 7. Tile Distribution Across Hosts in the First Test Case

Host	Tiles
Host1	1, 2, 3, 4, 5
Host2	6, 7, 8, 9, 10

The tiles were distributed across the four NFS datastores, as shown in Table 8.

Table 8. Tile Distribution Across Datastores in the First Test Case

NFS Datastore	Tiles
NFS_Datastore_001	1, 5, 9
NFS_Datastore_002	2, 6, 10
NFS_Datastore_003	3, 7
NFS_Datastore_004	4, 8

VMDKs were thin provisioned, which are the best practice and default when using Hitachi NAS Platform with VMware ESXi hosts. Thin provisioned is the VMDK type when migrating it to or creating it on an NFS datastore.

All VMDKs were 50 percent filled. During testing the environment was limited to 20 percent of the VMDK blocks.

During testing of 10 tiles on the NFS datastores on NAS Platform, 100 percent of the target I/O was achieved. The 10k RPM SAS drive parity groups that made up tier 1 storage averaged 45 percent busy.

The CPU controllers on NAS Platform average 10 percent busy. The FPGA percent busy was 16 percent during testing. This leaves plenty of room for an increase in workload on the NAS Platform node.

NFS Dedupe Testing on Hitachi NAS Platform

The second test was on Hitachi NAS Platform with dedupe running.

- The 10 tiles (250 virtual machines) were distributed across the four datastores and two VMware ESXi hosts, as listed in Table 7 on page 8 and Table 8 on page 8.

The 10 tiles ran for 12 hours (Table 9).

Table 9. Second Test Case Results

Test Case	Description	Result
Mixed workload application performance with an dedupe operation running on Hitachi NAS Platform to establish a baseline	Run workload on 10 tiles of 250 virtual machines for 12 hours	Passed
	Expected result: Virtual machines reach and maintain the I/O target configured for them.	

Figure 2 shows the target IOPS of each workload and the IOPS reached. Each workload reached its target IOPS during the dedupe test.

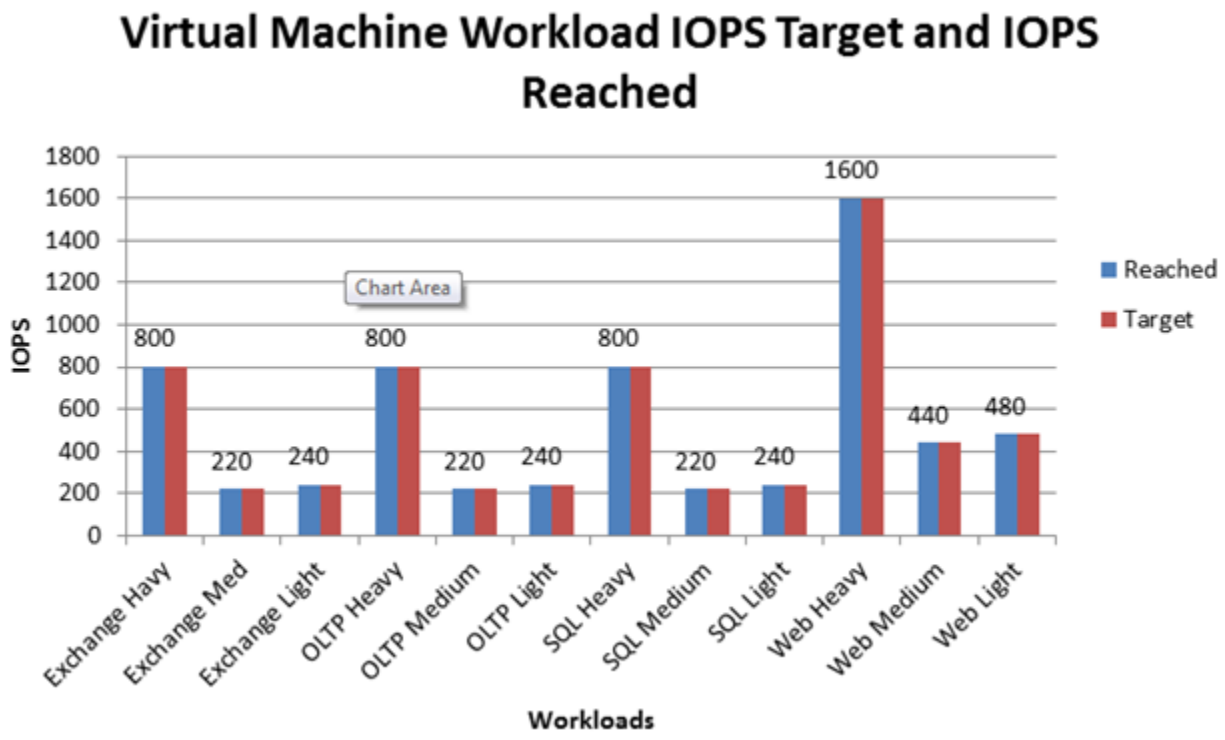

Figure 2

The dedupe operation did create additional load on Hitachi NAS Platform and Hitachi Unified Storage VM. Each application workload reached its target IOPS, but there was an impact in the application latency. Figure 3 shows how the application latency increased an average of 28.5 percent during the dedupe operation.

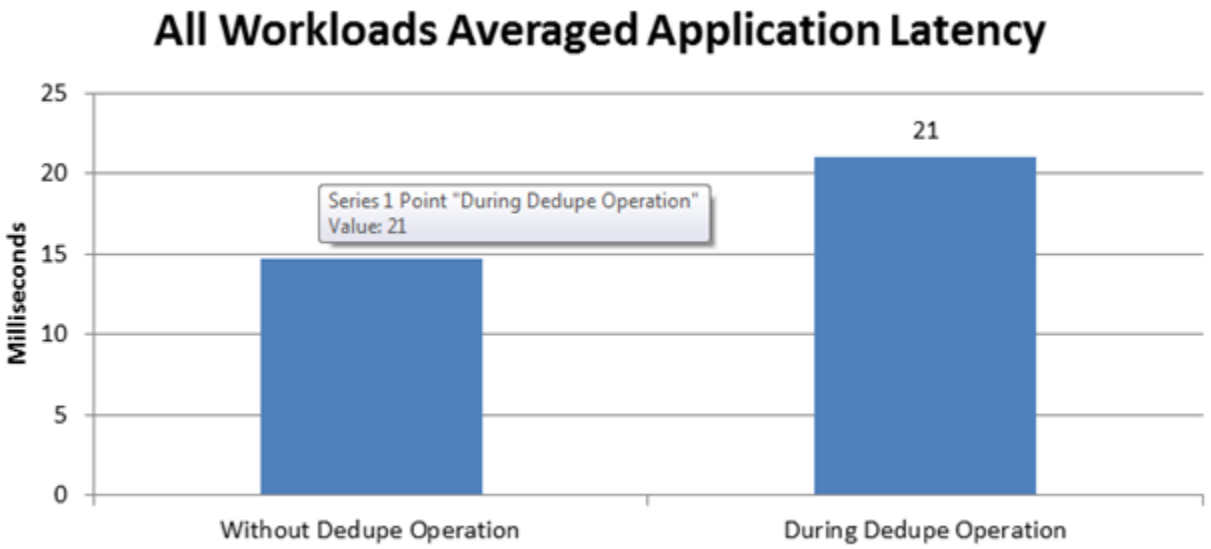

Figure 3

Figure 4 shows the increase in the percent busy of the parity group in Hitachi Universal Storage VM from Test 1, without dedupe running on Hitachi NAS Platform, to the percent busy of the parity group in Test 2, with dedupe running on NAS Platform. The increase in the dedupe operation workload resulted in the parity groups load increasing on Unified Storage VM.

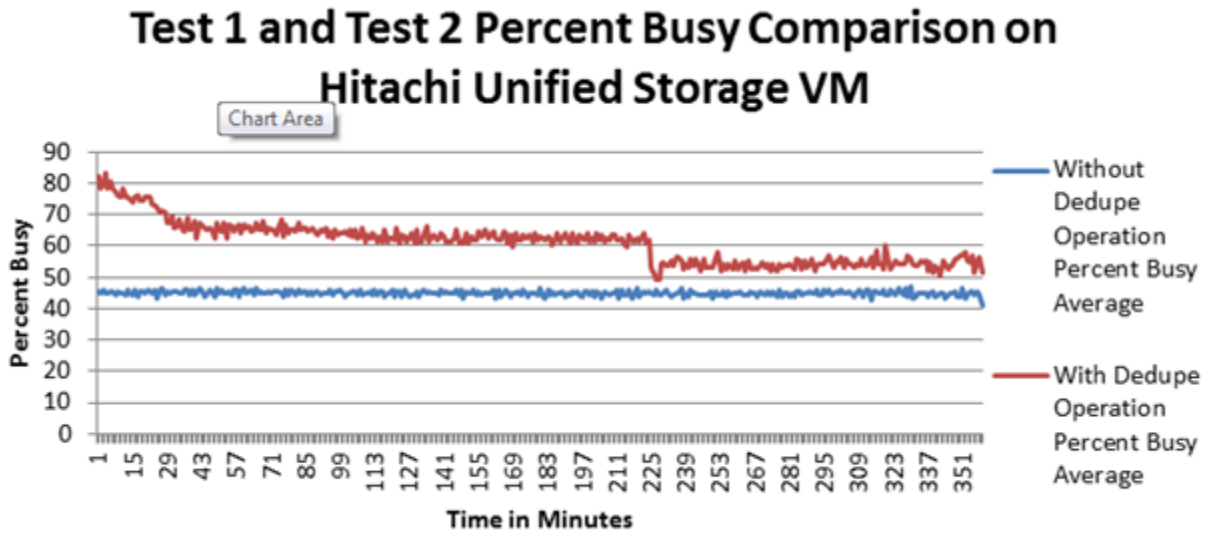

Figure 4

During the dedupe operation in Test 2, the CPU and FPGA utilization doubled on Hitachi NAS Platform when compared to Test 1 without dedupe running. Figure 5 illustrates the increase in CPU and FPGA utilization.

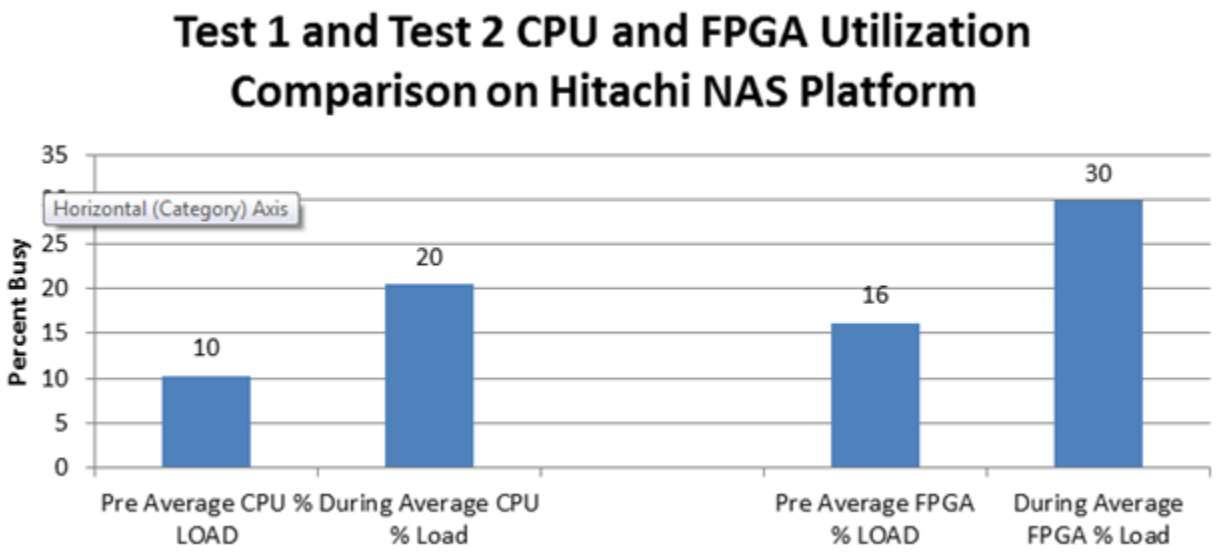

Figure 5

Testing showed that, for this particular workload, Hitachi NAS Platform maintained application I/O requirements during testing, including with dedupe running. The increased workload caused an increase in application I/O latency.

Conclusion

During dedupe testing with Hitachi NAS Platform, application I/O targets were consistently reached. Dedupe increased application latency by 28 percent during testing.

The impact of dedupe of the workloads may be able to be mitigated by configuring various dedupe and snapshot deletion parameters, but this was not tested during this round of testing.

Dedupe on Hitachi NAS Platform

This provides information on dedupe operations on Hitachi NAS Platform.

Types of Hitachi NAS Platform Dedupe Operations

There are three types of Hitachi NAS Platform dedupe operations. Table 10 describes all three.

Table 10. Tile Distribution Across Datastores on Hitachi NAS Platform

Type of Dedupe Operation	Description
Full	Runs if a file system is formatted initially without dedupe enabled, and then later enabling dedupe on that file system.
Triggered Incremental	Triggered after a configured amount of changes has occurred on the file system. The default is 1TB.
Daily Incremental	Triggered at a set time each day, as long as there are at least 20 GB of changes on the file system.

Platform Dedupe Operation Process on Hitachi NAS Platform

A dedupe operation on Hitachi NAS Platform has four steps.

1. The dedupe operation is initiated.
2. A block-based snapshot is taken.
3. The dedupe operation runs until completed or aborted.
4. The snapshot is deleted.

During testing there was an increase in NFS datastore and application latency. During the snapshot deletion process, the datastore and application latency was more pronounced. The duration of the snapshot deletion impact depends on the length of the dedupe operation and snapshot size.

Dedupe Troubleshooting on Hitachi NAS Platform

During the day-to-day operation of Hitachi NAS Platform with dedupe enabled, use the information in Table 11 to know the status of dedupe jobs, snapshot creation, and snapshot deletion.

Table 11. Commands to Determine Status of Dedupe Jobs, Snapshot Creation, and Snapshot Deletion

Command	Command Syntax	Description
fs-dedupe-history	`fs-dedupe-history -a <fsname>`	List any current dedupe operation and up to the last 5 previously running dedupe operations.
dedupe-queue-add	`dedupe-queue-add --full -f <fsname>`	Manually trigger a full dedupe operation.
dedupe-queue-add	`dedupe-queue-add --incremental -f <fsname>`	Manually trigger an incremental dedupe operation.
snapshot-list	`snapshot-list -a --file-system <fsname>`	List current snapshot on the specified file system, if there is any.

Table 11. Commands to Determine Status of Dedupe Jobs, Snapshot Creation, and Snapshot Deletion

Command	Command Syntax	Description
event-log-show	`event-log-show -o | grep -i dedupe`	When combined with the grep command, list all event log entries containing the dedupe key word.
event-log-show	`cn all event-log-show -o | grep -i dedupe`	When working in a clustered environment, list event log entries on all cluster nodes by adding "cn all" before the event-log-show command.
dedupe-queue-status	`dedupe-queue-status`	List all 5 dedupe queues and the status of any jobs in those queues.

Dedupe Customization on Hitachi NAS Platform

There are some customizable Hitachi NAS Platform parameters to fit dedupe operations better in your environment. Table 12 lists some of these parameters and how to use them.

Table 12. Tile Distribution Across Datastores

Command	Description
dedupe-daily-run-scheduler-frequency-in-seconds	The default value is 24 hours × 60 minutes per hour × 60 seconds per minute, or 24 hours.
dedupe-daily-run-scheduler-poll-interval-in-second	Set the frequency to check a file system to see if a dedupe should be ran. The Default is 1 hour.
dedupe-threshold-max-factor	Set the threshold to run an incremental dedupe. The trigger is set at 2 TB/<value of the variable>. The default variable value is 2, so the default trigger is 1 TB.
snapshot-max-unlink-blocks-per-checkpoint	Default is 100,000

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 Hitachi Data Systems Global 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 Hitachi Data Systems Corporate Resources website. Click the **Product Demos** tab for a list of available recorded demonstrations.

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 Education 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 96050-2639 USA
www.HDS.com community.HDS.com

Regional Contact Information
Americas: +1 408 970 1000 or info@hds.com
Europe, Middle East and Africa: +44 (0) 1753 618000 or info.emea@hds.com
Asia Pacific: +852 3189 7900 or hds.marketing.apac@hds.com

AS-411-00, August 2015