

Storage Subsystem Administration Guide

Hitachi NAS Platform

Release 14.2

Legal Notices

© 2020 Hitachi, Ltd. All rights reserved.

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including copying and recording, or stored in a database or retrieval system for commercial purposes without the express written permission of Hitachi, Ltd., or Hitachi Vantara LLC (collectively "Hitachi"). Licensee may make copies of the Materials provided that any such copy is: (i) created as an essential step in utilization of the Software as licensed and is used in no other manner; or (ii) used for archival purposes. Licensee may not make any other copies of the Materials. "Materials" mean text, data, photographs, graphics, audio, video and documents.

Hitachi reserves the right to make changes to this Material at any time without notice and assumes no responsibility for its use. The Materials contain the most current information available at the time of publication.

Some of the features described in the Materials might not be currently available. Refer to the most recent product announcement for information about feature and product availability, or contact Hitachi Vantara LLC at https://support.hitachivantara.com/en_us/contact-us.html.

Notice: Hitachi products and services can be ordered only under the terms and conditions of the applicable Hitachi agreements. The use of Hitachi products is governed by the terms of your agreements with Hitachi Vantara LLC.

By using this software, you agree that you are responsible for:

1. Acquiring the relevant consents as may be required under local privacy laws or otherwise from authorized employees and other individuals; and
2. Verifying that your data continues to be held, retrieved, deleted, or otherwise processed in accordance with relevant laws.

Notice on Export Controls. The technical data and technology inherent in this Document may be subject to U.S. export control laws, including the U.S. Export Administration Act and its associated regulations, and may be subject to export or import regulations in other countries. Reader agrees to comply strictly with all such regulations and acknowledges that Reader has the responsibility to obtain licenses to export, re-export, or import the Document and any Compliant Products.

Hitachi and Lumada are trademarks or registered trademarks of Hitachi, Ltd., in the United States and other countries.

AIX, AS/400e, DB2, Domino, DS6000, DS8000, Enterprise Storage Server, eServer, FICON, FlashCopy, GDPS, HyperSwap, IBM, Lotus, MVS, OS/390, PowerHA, PowerPC, RS/6000, S/390, System z9, System z10, Tivoli, z/OS, z9, z10, z13, z14, z/VM, and z/VSE are registered trademarks or trademarks of International Business Machines Corporation.

Active Directory, ActiveX, Bing, Edge, Excel, Hyper-V, Internet Explorer, the Internet Explorer logo, Microsoft, the Microsoft corporate logo, the Microsoft Edge logo, MS-DOS, Outlook, PowerPoint, SharePoint, Silverlight, SmartScreen, SQL Server, Visual Basic, Visual C++, Visual Studio, Windows, the Windows logo, Windows Azure, Windows PowerShell, Windows Server, the Windows start button, and Windows Vista are registered trademarks or trademarks of Microsoft Corporation. Microsoft product screen shots are reprinted with permission from Microsoft Corporation.

All other trademarks, service marks, and company names in this document or website are properties of their respective owners.

Copyright and license information for third-party and open source software used in Hitachi Vantara products can be found at <https://www.hitachivantara.com/en-us/company/legal.html>.

Contents

Legal Notices.....	2
Preface	6
Accessing product documentation.....	6
Getting help.....	6
Comments.....	6
Chapter 1: Understanding storage and tiering.....	7
Understanding tiered storage.....	7
Storage management components.....	8
System drives.....	8
Storage pools.....	9
Tiered storage pools.....	10
Dynamically provisioned volumes.....	10
File system types.....	11
Fibre Channel connections.....	11
About FC paths.....	12
Load balancing and failure recovery.....	14
Fibre Channel statistics.....	15
RAID controllers (HNAS Gateway only).....	15
Chapter 2: Managing the storage system.....	16
Supported Hitachi storage systems.....	16
Hitachi Dynamic Provisioning.....	18
HDP high-level process.....	19
Understanding HDP thin provisioning.....	19
Deciding how far to over-provision storage.....	21
Configuration considerations when using HDP with NAS	22
Using HDP to expand a span.....	22
Configuration guidelines for HNAS with HDP.....	23
Disable zero page reclaim.....	24
Storage pool chunks.....	24
Storage pool pre-allocation.....	25
Using HDP pools with a NAS server.....	26
Performing large file system expansions.....	27
Upgrading from older HNAS systems.....	27

Using HDP storage.....	28
Considerations when using HDP pools.....	28
Creating an HDP pool with untiered storage.....	28
Creating HDP pools with tiered storage.....	29
Storage pool naming.....	29
Creating a storage pool using the CLI.....	31
Creating a storage pool using the GUI.....	32
Adding the metadata tier.....	34
Allowing access to a storage pool.....	35
Creating storage pools with DP pools from HDP storage.....	36
Moving free space between storage pools.....	36
Important unmapper considerations.....	37
Using the unmapper to move space between storage pools.....	38
Thin provisioning using the NAS server filesystem-thin command.....	39
System drive groups and dynamic write balancing.....	39
Chapter 3: Using a storage pool.....	41
Expanding storage pools.....	41
Expanding space in a thinly provisioned HDP storage pool.....	41
Expanding storage space using DP-Vols.....	42
Expanding a non-HDP storage pool or tier.....	42
Configuring automatic file system expansion for an entire storage pool.....	44
Renaming a storage pool.....	45
Reducing the size of a storage pool.....	45
Denying access to a storage pool.....	45
Deleting a storage pool.....	47
Chapter 4: Managing flash storage.....	49
Using compression with flash module drives (FMDs)	49
Expanding file systems on flash storage.....	50
Detecting free space on flash storage.....	51
Using compression statistics.....	51
Low physical space.....	53
Using SVOS RF Data Reduction (ADR).....	53
Chapter 5: Supporting external storage.....	56
Using UVM.....	56
Taking over a UVM span.....	57
Using multiple HDP pools.....	58
Expanding a UVM-based filesystem.....	58
Migrating data to a local cluster.....	62
Making efficient use of disk space.....	62

When to avoid block-level migration.....	63
--	----

Preface

This guide provides information about managing the supported storage systems attached to the server/cluster. The guide includes information about tiered storage, storage pools, system drives (SDs), using Hitachi Dynamic Provisioning (HDP) with the server, and other storage related configuration and management features and functions. Note that some features apply only to individual platforms and may not be applicable to your configuration.

Virtual Storage Platform G400, G600, G800 and Virtual Storage Platform F400, F600, F800 storage systems can be configured with NAS modules to deliver native NAS functionality in a unified storage platform. The term 'NAS module' in this document also applies to VSP F series, VSP G series, and VSP N series. The unified VSP Gx00 models, VSP Fx00 models, and VSP N series models automatically form a two-node cluster in a single chassis upon installation, with no external cabling required.

Accessing product documentation

Product user documentation is available on the Hitachi Vantara Support Website: <https://knowledge.hitachivantara.com/Documents>. Check this site for the most current documentation, including important updates that may have been made after the release of the product.

Getting help

The [Hitachi Vantara Support Website](https://support.hitachivantara.com/en_us/contact-us.html) is the destination for technical support of products and solutions sold by Hitachi Vantara. To contact technical support, log on to the Hitachi Vantara Support Website for contact information: https://support.hitachivantara.com/en_us/contact-us.html.

[Hitachi Vantara Community](https://community.hitachivantara.com) is a global online community for Hitachi Vantara customers, partners, independent software vendors, employees, and prospects. It is the destination to get answers, discover insights, and make connections. **Join the conversation today!** Go to community.hitachivantara.com, register, and complete your profile.

Comments

Please send us your comments on this document to doc.comments@hitachivantara.com. Include the document title and number, including the revision level (for example, -07), and refer to specific sections and paragraphs whenever possible. All comments become the property of Hitachi Vantara LLC.

Thank you!

Chapter 1: Understanding storage and tiering

Understanding tiered storage

Tiered storage allows you to connect multiple diverse storage systems behind a single server (or cluster) and match application storage requirements (in terms of performance and scaling) to your storage systems. This section describes the concept of tiered storage, and explains how to configure the storage server to work with your storage systems to create a tiered storage architecture.

Based on a storage system's performance characteristics, it is classified as belonging to a certain tier, and each tier is used differently in the enterprise storage architecture. The currently supported storage systems fit into the tiered storage model as follows:

Tier	Performance	Drive Type	RPM (if disk)
0	Extremely high	Flash or solid-state memory; not disk	N/A
1	Very high	SAS	15,000
2	High	SAS	10,000
3	Nearline	Nearline (NL) SAS	7,200
4	Archival	Nearline (NL) SAS	7,200
5	Long-term storage (Tape)	N/A (Tape)	N/A

The NAS server supports tiers of storage, where each tier is made up of devices with different performance characteristics or technologies. The NAS server also supports storage virtualization through applicable Hitachi RAID storage systems using UVM (Universal Volume Manager).



Note: Utilizing UVM with HNAS architecture is not common practice and can lead to performance anomalies, particularly if it is a non-Hitachi storage system. Always use local disks as first priority in system design.

Tiers of storage and storage virtualization are fully supported by Data Migrator, an optional feature which allows you to optimize the usage of tiered storage and remote NFSv3 servers. Note, however, that Data Migrator does not support migration to or from tape storage devices or tape library systems. For detailed information about Data Migrator, refer to the *Data Migrator Administration Guide*.

Storage management components

The storage server architecture includes system drives, storage pools, file systems, and virtual servers (EVSs), supplemented by a flexible quota management system for managing utilization, and Data Migrator to Cloud (DM2C) along with classic Data Migrator, both of which optimize available storage by moving cold data to less expensive storage. This section describes each of these storage components and functions in detail.

System drives

System drives (SDs) are the basic logical storage element used by the server. Storage systems use RAID controllers to aggregate multiple physical disks into SDs (also known as LUs). An SD is a logical unit made up of a group of physical disks or flash drives. The size of the SD depends on factors such as the RAID level, the number of drives, and their capacity.

Hitachi Enterprise RAID storage systems have a limit of 3 TiB for standard LUs or 4 TiB for virtualized LUs (HUVL). When using legacy storage systems, it is a common practice for system administrators to build large RAID groups (often called parity groups or volume groups) and then divide them into SDs (LUs) of 2 TiB or less. With today's large physical disks, RAID groups must be considerably larger than 2 TiB to make efficient use of space.

When you create SDs:

- Use the Hitachi storage management application appropriate for your storage system. You cannot create SDs using NAS Manager or the NAS server command line.
- You may need to specify storage system-specific settings in the storage management application.

For more information about the settings required and the firmware that is installed for each type of storage system, contact customer support.

Storage pools

A NAS server storage pool (known as a "span" in the command line interface) is the logical container for a collection of four or more system drives (SDs). There are two types of NAS server storage pools:

- An untiered storage pool is made up of system drives (SDs) created on one or more storage systems within the same tier of storage (storage systems with comparable performance characteristics). To create an untiered storage pool, there must be at least four available and unused system drives on the storage system from which the SDs in the storage pool are taken.
- A tiered storage pool is made up of system drives (SDs) created on storage systems with different performance characteristics. Typically, a tiered storage pool is made up of SDs from high-performance storage such as flash memory, and SDs from lower-performance storage such as SAS (preferably) or NL SAS (near line SAS). You can, however, create a tiered storage pool from SDs on storage systems using any storage technology, and you can create both tiers on the same storage system.

NAS server storage pools:

- Can be expanded as additional SDs are created in the storage system, and a storage pool can grow to a maximum of 1 PiB or 256 SDs. Expanding a NAS server storage pool does not interrupt network client access to storage resources. SDs may be based on parity groups, or on HDP DP-Vols (preferably).
- Support two types of thin provisioning:
 - NAS server storage pools can be thinly provisioned when created using SDs based on HDP DP-Vols.
 - File system thin provisioning, through the use of the NAS server **filesystem-thin** command and file system confinement and auto-expansion. This type of thin provisioning allows you to create a small file system, which can then automatically expand when necessary, and that ability saves the overhead associated with sustaining unnecessary storage.

When file system thin provisioning is enabled, the server reports to protocol clients (though not at the CLI or in the GUI) that the file system is larger than it really is: either the capacity to which you have confined it or the maximum capacity to which it can ever grow, whichever is smaller.

Refer to the *Command Line Reference* for more information on the **filesystem-thin** command.

- Contain a single stripeset on initial creation. Each time the storage pool is expanded, another stripeset is added, up to a maximum of 64 stripesets (after creation, a storage pool can be expanded a maximum of 63 times). As HDP is the preferred method to provision DP Vols to the NAS, Hitachi recommends thin provisioning the pool 200 - 300 percent, which lessens the likelihood of stripeset expansions.
- Contain the file systems and enable the user to manage the file system settings that are common to all file systems in the storage pool. For example, the settings applied to a storage pool can either allow or constrain the expansion of all file systems in the storage pool.



Note: By default, there is a limit of 32 file systems per storage pool. Recently deleted file systems that are still in the recycle bin do not count towards this number. It is possible to increase this limit using the `filesystem-create` CLI command with the `--exceed-safe-count` option. See the command man page for details.

Tiered storage pools

Currently, a tiered storage pool must have two tiers:

- Tier 0 is used for metadata, and the best-performing storage should be designated as Tier 0.
- Tier 1 is used for user data.

When creating a tiered storage pool, at least four unused SDs must be available for each tier. When you create a tiered storage pool, you first create the user data tier (Tier 1), then you create the metadata tier (Tier 0).

During normal operation, one tier of a tiered storage pool might become filled before the other tier. In such a case, you can expand one tier of the storage pool without expanding the other tier. When expanding a tier, you must:

- Make sure that the SDs being added to the tier have the same performance characteristics as the SDs already in the tier. For example, do not add NL SAS (near line SAS) based SDs to a tier already made up of flash drives.
- Add SDs to the tier. See the `span-create` man page for more information about minimum SD counts and creating storage pools.

Dynamically provisioned volumes

A dynamically provisioned volume (DP-Vol) is a virtualized logical unit (LU) that is used with Hitachi Dynamic Provisioning (HDP). You create DP-Vols in a dynamically provisioned pool (a DP pool), which is an expandable collection of physical storage. The maximum capacity of an SD is 64TiB.

The total capacity of a DP-Vol can exceed that of the underlying parity groups or pool volumes (called thin provisioning). Every DP-Vol can draw space from any of the underlying parity groups or pool volumes, so the system performs well even if the load on the SDs is unbalanced.

Hitachi Dynamic Provisioning (HDP) thin provisioning enables granular span-expansion without loss of performance, which is impossible without dynamic provisioning. The NAS server is aware of thin provisioning, and does not use more space than actually exists, making thin provisioning safe with a NAS server. With this server, unlike other server platforms, it is not necessary to expand a thickly provisioned HDP pool as soon as it becomes 70 percent full.

File system types

A file system typically consists of files and directories. Data about the files and directories (as well as many other attributes) is the metadata. The data within the file system (both user data and metadata) is stored in a storage pool.

Like storage pools, file system data (metadata and user data) may be stored in a single tier, or in multiple tiers.

- When file system metadata and user data are stored on storage systems of a single storage tier, the file system is called an *untiered* file system. An untiered file system must be created in an untiered storage pool, it cannot be created in a tiered storage pool.
- When file system metadata and user data are stored on storage systems of different storage tiers, the file system is called a *tiered* file system.

In a tiered file system, metadata is stored on the highest performance tier of storage, and user data is stored on a lower-performance tier. Storing metadata on the higher-performance tier provides performance and cost benefits.

A tiered file system must be created in a tiered storage pool; it cannot be created in an untiered storage pool.

Fibre Channel connections



Note:

- The number and operational speed of Fibre Channel (FC) ports on a NAS server are dependent on the server model. Refer to the hardware manual for your server model for more information on the number, operational speed, and location of FC ports on your NAS server.
- The ports on the HNAS 5000 series are in reverse order from prior gateways.
 - Left to right: FC 4-3-2-1
 - Earlier generations were FC 1-2-3-4

Hitachi NAS Platform servers

Each HNAS server supports up to four independently configurable FC ports. Independent configuration allows you to connect to a range of storage systems, which allows you to choose the configuration that best meets the application requirements. The server manages all back-end storage as a single system, through an integrated network management interface.

Hitachi NAS Platform server model	Supported FC port operational speeds
3080, 3090, 3100, and 4040	1, 2, or 4 Gbps
4060, 4080, and 4100	2, 4, or 8 Gbps
5200, 5300	4, 8, or 16 Gbps

The server supports connecting to storage systems either through direct-attached FC connections to the storage system (also called DAS connections) or FC switches connected to the storage system (also called SAN configurations):

- In direct-attached (DAS) configurations, you can connect up to two storage systems directly to a server or a two-node cluster. Clusters of more than two nodes must use a FC switch configuration.
- In configurations using FC switches (SAN configurations), the server must be configured for N_Port operation (when not using the 5000 series).
 - N_Port is the only operational mode on the 5000 series.
 - Contact customer support for more information on supported FC switch interoperability.

You can manage the FC interface on the server/cluster through the command line interface (CLI), using the following commands:

- `fc-link` to enable or disable the FC link.
- `fc-link-type` to change the FC link type. (This command is not available on the 5000 series because only N_Port is supported).
- `fc-link-speed` to change the FC interface speed. (Only the HNAS 5000 series supports Auto Negotiation, which is the default setting.)

For more information about these commands, refer to the *Command Line Reference*.

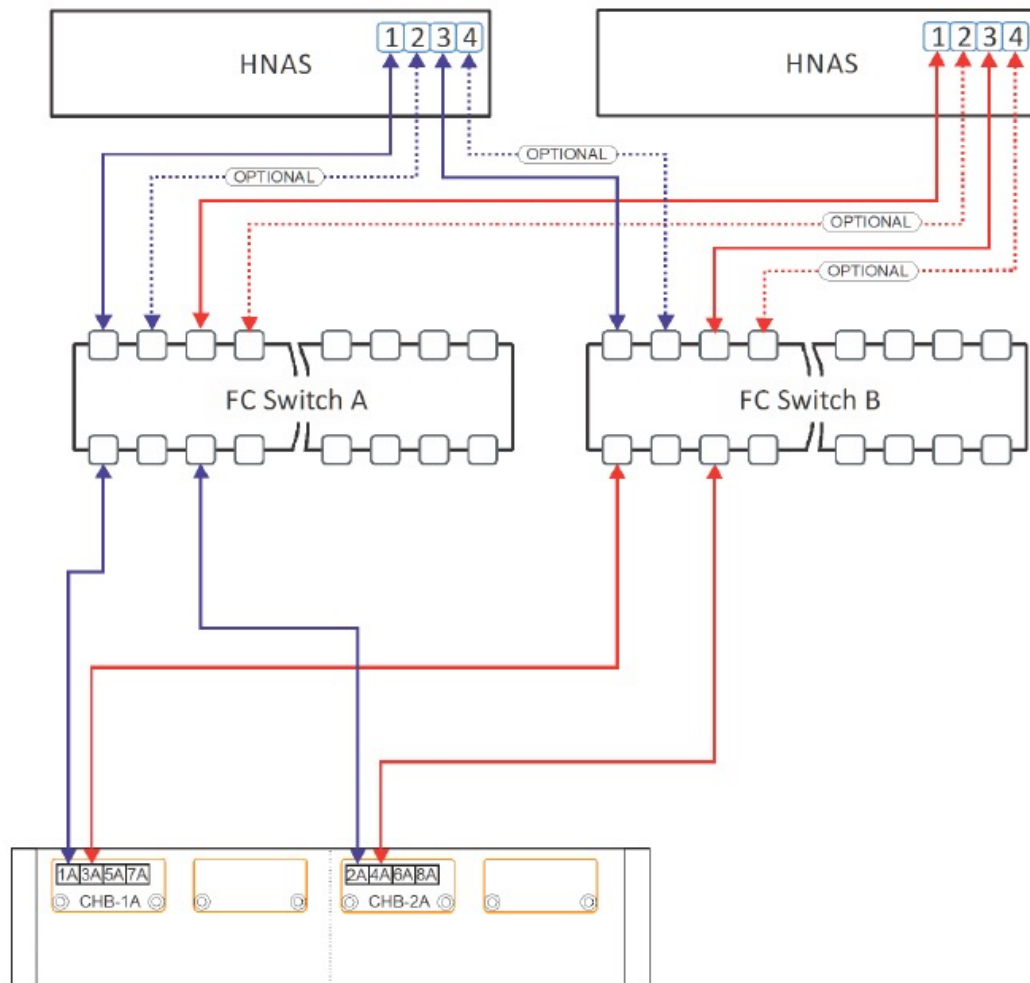
VSP F/G/Nx00 servers with NAS modules

Depending on the model, your NAS module server may contain 8 Gbps and/or 16 Gbps FC ports for the block connectivity. NAS module connectivity is by FC protocol over PCIe. Refer to the *Hardware Reference Guide* for your VSP Gx00 or Fx00 model server for more information about the FC ports for block connectivity, and contact Hitachi Vantara Customer Support for information about using the ports.

About FC paths

The NAS server accesses the storage system through a minimum of two Fibre Channel (FC) paths (at least one from each of the FC switches). Unless otherwise stated, the recommended number of paths per SD should be limited to 16 (2 HNAS host FC ports per switch x 2 switches x 4 target ports per switch). An FC path is made up of the server's host port ID, the storage system port WWN (worldwide name), and the SD identifier (ID).

The following illustration shows a complete path from the server to each of the SDs on the storage system:



You can display information about the FC paths on the server/cluster through the command line interface (CLI), using the `fc-host-port-load`, `fc-target-port-load`, and the `sdpath` commands.

Fibre channel interface

The ports are in reverse order from prior gateways (due to mounting the HBA within the chassis):

- Left to right: FC 4-3-2-1 (earlier generations were FC 1-2-3-4)
- 2-node maximum for DAS storage (which is the same for prior gateways)



Figure 1 FC interface ports

Load balancing and failure recovery

Load balancing on a storage server is a matter of balancing the loads to the system drives (SDs) on the storage systems to which the storage server is connected. A logical unit (LU), known to the server as an SD, is a piece of disk space managed by the block storage, spread across several physical disks.

The server routes FC traffic to individual SDs over a single FC path, distributing the load across two FC switches and, when possible, across dual active/active or multi-port RAID controllers.

Following the failure of a preferred path, disk I/O is redistributed among other (non-optimal) paths. When the server detects reactivation of the preferred FC path, it once again redistributes disk I/O to use the preferred FC path.

Default load balancing (load balancing automatically performed by the storage server) is performed based on the following criteria:

- “Load” is defined as the number of open SDs, regardless of the level of I/O on each SD. An SD is open if it is in a span that has a file system that is mounted or is being formatted, checked or fixed. SDs count towards load at the target if they are open on at least one cluster node; the number of nodes (normally all nodes in a cluster, after boot) is not considered.
- Balancing load on RAID controller target ports takes precedence over balancing load on server FC host ports.
- Balancing load among a system's RAID controllers takes precedence over balancing among ports on those controllers.
- In a cluster, choice of RAID controller target port is coordinated between cluster nodes, so that I/O requests for a given SD do not simultaneously go to multiple target ports on the same RAID controller.

You can manually configure load distribution from the CLI (overriding the default load balancing performed by the server), using the `sdpath` command. When manually configuring load balancing using the `sdpath` command:

- You can configure a preferred server host port and/or a RAID controller target port for an SD. If both are set, the RAID controller target port preference takes precedence over the server host port preference. When a specified port preference cannot be satisfied, port selection falls back to automatic selection.
- For the SDs visible on the same target port of a RAID controller, you should either set a preferred RAID controller target port for **all SDs** or for **none of the SDs**. Setting the preferred RAID controller target port for only some of the SDs visible on any given RAID controller target port may create a situation where load distribution is suboptimal.



Note:

Manually setting a preferred path is not necessary or recommended.

The `sdpath` command can also be used to query the current FC path being used to communicate with each SD. For more information on the `sdpath` command, enter `man sdpath` command.

To see information about the preferred path, navigate to Home > Storage Management > System Drives, then select the SD and click details to display the **System Drive Details** page. If available, the FC Path section provides information about the path, port, and controller.

Fibre Channel statistics

The server provides per-port and overall statistics, in real time, at 10-second intervals. Historical statistics cover the period since previous server start or statistics reset. The **Fibre Channel Statistics** page of the NAS Manager displays the number of bytes/second received and transmitted during the past few minutes.

RAID controllers (HNAS Gateway only)

The RAID controllers operate as an Active/Active (A/A) pair within the same storage system. Both RAID controllers can actively process disk I/O requests. If one of the two RAID controllers fails, the storage server reroutes the I/O transparently to the other controller, which starts processing disk I/O requests for both controllers.

Chapter 2: Managing the storage system

NAS Platform storage systems can be managed using NAS Manager. Common operations are:

- Changing the storage system name, password, or media scan period
- Checking the status of media scan and other operations
- Reviewing events logged by the RAID rack
- Determining the status of physical disks

Supported Hitachi storage systems

All NAS Platform Series 3000, NAS Platform Series 4000, and NAS Platform Series 5000 storage servers support storage systems manufactured exclusively by Hitachi Vantara. Supported storage systems are dependent on server series and models, which are listed as follows:



Note:

- If a storage system is listed in the Currently Supported Offerings column for the specified Hitachi NAS Platform server model, it is properly recognized and effectively managed by the servers.
- Systems not appearing in the supported list may be recognized by the server, but normal server functionality is limited and may result in data availability or system stability issues. Refer to the system event log or trouble report for information.
- Unsupported storage systems must never be connected to Hitachi NAS Platform servers.

Server Series	Server Model	Currently Supported Offerings	Notes
NAS Platform Series 4000	For the NAS Platform Series 4000: 4060 4080 4100 (End of Sale)	VSP G200, VSP G350, VSP G370, VSP G400, VSP G600, VSP G700, VSP G800, VSP G900, VSP G1000, VSP G1500, VSP F1500, VSP F200, VSP F350, VSP F370, VSP F400, VSP F600, VSP F700, VSP F800, VSP F900, VSP 5100, VSP 5200, VSP 5500, VSP 5600, VSP 5100H, VSP 5200H, VSP 5500H, VSP 5600H, VSP E590, VSP E790, VSP E590H, VSP E790H, and VSP E990	
NAS Platform Series 5000	For the NAS Platform Series 5000: 5200 5300	same as above, plus VSP E1090, VSP E1090H	
NAS Platform Series 4000	4040 (End of Sale)	VSP G200, VSP G400, VSP G600, VSP G800, VSP G1000, VSP G1500, VSP F1500, VSP F200, VSP F400, VSP F600, VSP F800	

Server Series	Server Model	Currently Supported Offerings	Notes
NAS Platform Series 3000	3080 and 3090 (End of Service Life)	VSP G200, VSP G400, VSP G600, VSP G800, VSP G1000, VSP G1500, VSP F1500, VSP F200, VSP F400, VSP F600, VSP F800	

**Note:**

- For external storage virtualized using the Universal Volume Manager (UVM), consult the current UVM matrix for supported storage systems.
- All currently supported Hitachi Vantara storage systems support RAID levels 1, 5, 6, and 10.

Hitachi Dynamic Provisioning

You can use Hitachi Dynamic Provisioning (HDP) software to improve your storage utilization. The HDP software uses storage-based virtualization layered on top of RAID technology (RAID on RAID) to enable virtual LUNs (dynamically provisioned volumes, DP-Vols) to draw space from multiple pool volumes. This aggregated space widens the storage bottleneck by distributing the I/O to more disks. The greater distribution insulates the server from the realities of the pool volumes (small capacities of individual disks).

HDP with a NAS server provides the following benefits:

- Improves performance by striping I/O across all available disks
- Supports larger LUs, up to 64 TiB
- Reduces the need to use the **span-expand** command. When HDP thin provisioning is used, an HDP pool can be expanded in small increments any number of times. However, if you expand a storage pool, make the increments as large as the initial size of the storage pool to avoid performance problems.
- File system creation and expansion are safe, even on thinly provisioned HDP pools, because they check the amount of available space.

If you are using HDP, see the *Hitachi NAS Platform HDP Best Practices* (MK-92HNAS063) for recommendations.



Note: When using a storage system, there are commonly used Host Mode Options (HMOs) and System Option Modes (SOMs) which should be set correctly. For example, on Hitachi Enterprise RAID systems always enable HMO 7 and 68. When using an HNAS Gateway in a stretched cluster with GAD, enable HMO 78 on the host group containing the HNAS WWPN for the NAS that is considered remote to the system. Contact customer support for more information.

HDP high-level process

The following flow chart shows the high-level process for provisioning storage with HDP:

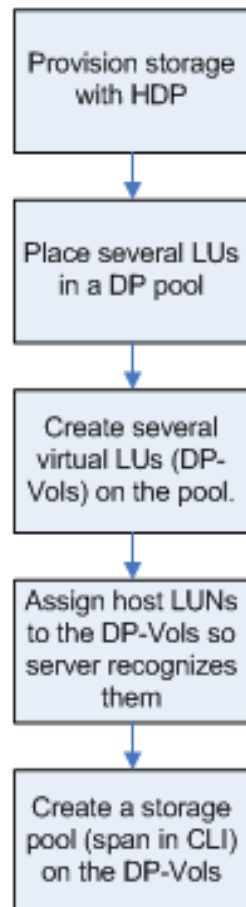


Figure 2 High-level process for HDP provisioning

Understanding HDP thin provisioning

Thin provisioning allows space to be allocated to an application without it being physically mapped on the storage system until it is actually used. Thin provisioning also decouples the logical provisioning of storage to an application from the physical addition of storage capacity to the storage system.

For example, given 30 TiB of physical storage, you can create an HDP pool with 80 TiB of DP-Vols and create an 80 TiB NAS storage pool on those DP-Vols. None of the available space is allocated until you create and expand file systems. Because only 30 TiB of real space is available, the NAS server will not create more than 30 TiB of file systems in the storage pool. If you later add more parity groups or pool volumes to the HDP pool, you can expand the file systems in the storage pool beyond 30 TiB without creating additional DP-Vols or expanding the NAS storage pool.

The ability to expand file systems instead of the storage pool is advantageous because it enhances performance by spreading the storage chunks used to expand a file system across all the SDs and physical disks in the DP pool. In contrast, a storage pool expansion limits performance by restricting the individual chunks to a small number of SDs and physical disks.



Note: It is strongly recommended that you always use thin provisioning with HDP.

The NAS server reads the real space available in a DP pool. When you create or expand a file system, the NAS server checks for available space, then pre-allocates the space needed for that operation. If the DP pool has too little free space for the operation to complete successfully, the NAS server safely aborts the creation or expansion of the file system.

Every new storage pool should use a single stripeset that resides on a thinly provisioned HDP pool. This way, storage can be expanded in small increments without loss of performance, and all I/O will use all DP-Vols (and all their queue depth) and all physical storage media. For more about queue depth, see the `sd-queue-depth` man page. To use a single stripeset, follow the instructions below.

The process is as follows:

1. When provisioning a new NAS server storage pool, use just enough real disk space to meet your immediate needs for performance and capacity.
2. Place all your parity groups into a single HDP pool, then create DP-Vols whose total capacity roughly meets your expected needs for the next 18 to 24 months.



Note: It does not matter if you over-estimate or under-estimate your capacity needs, because you can easily expand the storage pool beyond the total capacity of the original DP-Vols by adding another set of DP-Vols.

3. Create a NAS server storage pool on these DP-Vols, placing all the DP-Vols into a single stripeset.

Use enough DP-Vols to provide adequate queue depth in the future, after you have added enough parity groups to match the total capacity of the DP-Vols. Four DP-Vols is the bare minimum, but eight DP-Vols will provide better performance than four, and sixteen DP-Vols will be faster than 8 DP-Vols. In practice, a storage pool usually contains an even number of DP-Vols, and the capacity of each DP-Vol is 8 TiB.



Note: If using the CLI `span-create` command, list all the SDs in the initial `span-create` command. Do not run a single `span-create` command, then a series of `span-expand` commands.



Note: When using an application to create a storage pool, specify all the available SDs when creating the storage pool; do not create a single storage pool on a subset of the available SDs, then expand that storage pool onto the rest of the available SDs.

If there are more than 32 available DP-Vols, create the minimum possible number of NAS server stripesets consistent with making all stripesets identical, even if this means creating slightly more or slightly fewer DP-Vols than would otherwise have been created. For example, if you initially estimate that, in two years, you will need 50 8 TiB DP-Vols, you should now create 48 DP-Vols and make 2 stripesets of 24 DP-Vols each.

4. To expand the NAS server storage pool beyond the total capacity of the original DP-Vols, simply add another, identical set of DP-Vols (refer to the **span-expand** man page for more information).



Note: Every new storage pool contains one stripeset, and every expansion (other than by adding storage to the underlying HDP pool) adds a further stripeset.

Deciding how far to over-provision storage

When using HDP, you should make a rough forecast of how much data storage capacity will be needed in the next 12 to 24 months, and then configure your DP-Vols to be larger than your estimate. If you overestimated your data storage requirements, not too much space will have been wasted.

The total capacity of the DP-Vols should exceed the total capacity of the parity groups or pool volumes by a factor of 2:1 or 3:1, depending on how far you expect the storage pool to expand. The total capacity of the DP-Vols created when the storage pool was initially set up does *not* constrain the eventual size of the storage pool.

For example, if you have 20 TiB of storage and the storage pool may need to expand to 50 TiB later on, you should set up 50 TiB of DP-Vols. If you ever need to grow the storage pool beyond 50 TiB, you can always add a second stripeset using the **span-expand** command, then continue to expand the DP pool in increments as small as required.

Limits on thin provisioning:

- You can make the storage pool capacity larger than the total capacity of the DP-Vols that you created at the outset by adding more DP-Vols later.
- Some storage systems and systems do not over-commit by more than a factor of ten to one.
- For HDP, the storage requires an amount of memory that is proportional to the capacity of the large, virtual DP-Vols, rather than to the smaller, real parity groups or pool volumes. Therefore, consider the following:
 - Massive over-commitment causes storage to run out of memory prematurely.
 - Enterprise storage uses separate boards called *shared memory*.

Configuration considerations when using HDP with NAS

When your system must work with HDP-based storage, there are things to remember that affect both the storage and the NAS server. For detailed information on using HDP with your particular storage, consult the HDP software documentation.

When using HDP pools with NAS servers, consider the following:

- All current Hitachi Vantara storage supports HDP as a licensed option.
- All NAS servers support HDP without requiring a server-side license.
- The HDP software has no effect on protocol clients, such as NFS, and SMB.
- Each DP-Vol draws space from multiple pool volumes, which helps to relieve the storage bottleneck by distributing I/O to more disks.
- If you attempt to create or expand a file system and either there are not enough free chunks on the span or there is not enough free disk space on the DP pool, the server will make space by recycling one or more deleted file systems, if the span contains any.



Note: When you recycle or delete a file system, the amount of free space shown in a storage configurator such as Hitachi Ops Center Administrator or Hitachi Storage Navigator does not reflect the new space until after you have run the **span-unmap-vacated-chunks** command. Do not run this command unnecessarily because performance may be reduced.

- Recycling a file system causes the chunks that stored the file system data to be moved into the vacated-chunks list, which contains records of which freed chunks were used by which file system.
- Creating or expanding a file system draws space from the vacated chunks list, if any is available, without using new space from the HDP pool. Any further space is pre-allocated at once.
- Writing to a file system costs no space because that space was pre-allocated.
- When creating or expanding a storage pool on HDP pools, you must use the DP-Vols from a single DP pool. This rule applies whether you are using the CLI or the NAS Manager. Later storage pool expansions can be done using storage from different DP pools.

Using HDP to expand a span

Using HDP to add space provides additional benefits.

These benefits include:

- You can add disks in small increments, even just a single pool volume.
- Data gets restriped.
- Span gets faster and performance remains almost even.

Consider the following use case for using HDP to expand a span:

If you originally created a pool containing 10 TiB of real storage and eight DP-Vols of 2.5 TiB each, totalling 20 TiB, the pool is over-committed by a ratio of 2:1. As always, a storage pool (span on the CLI) resides on the DP-Vols. As time goes by, you make a series of expansions of 4 TiB each by adding new parity groups or pool volumes. The first expansion increases the amount of real storage in the pool to 14 TiB and the second expansion takes it to 18 TiB.

After each of these expansions, no further action is necessary. However, after a third 4 TiB expansion, the pool contains 22 TiB of real storage, but its DP-Vols total only 20 TiB. As a result, 2 TiB of the disk space that you have installed are inaccessible to the server.

More DP-Vols are needed, but any expansion should always add at least as many DP-Vols as were provided when the span was created. You must therefore create a further eight DP-Vols, preferably of the same 2.5 TiB as the original ones, and add them to the storage pool by using the **span-expand** command or NAS Manager equivalent. This addition brings the total DP-Vol capacity to 40 TiB. No further DP-Vols will be necessary unless the real disk space in the pool is expanded beyond 40 TiB.

Configuration guidelines for HNAS with HDP

Follow these configuration guidelines for best results.

On the storage:

- Make every new HDP pool thinly provisioned.
- Create enough DP-Vols to meet the expected size of the storage pool and to provide enough queue depth (minimum of four).
- Limit each HDP pool to hosting only a single NAS server storage pool, particularly with low performing HDD's; when using flash media using multiple spans per HDP pool is more acceptable. With the exception of tiered file systems, if you need fifty NAS server storage pools, create fifty HDP pools.
- Do not share an HDP pool between two or more clusters, as disparate clusters are unaware of the other's storage needs; additionally, troubleshooting performance concerns is exacerbated due to unknown workloads.
- Do not share an HDP pool between an HNAS system and a foreign server, as disparate clusters are unaware of the other's storage needs; additionally, troubleshooting performance concerns is exacerbated due to unknown workloads.

On the NAS server:

- To avoid server timeouts when creating a new NAS server storage pool/span, wait for the HDP pool to finish formatting before creating the NAS server storage pool/span using NAS Manager or the **span-create** command.



Note: If the HDP pool has finished formatting, but the NAS server does not detect the new DP-Vols, run the **scsi-refresh** CLI command so the NAS server will detect the new DP-Vols, , though this should not be necessary when HMO 7 is enabled on the target port host group.

- For tiered file systems, you may either use two HDP pools when creating the NAS server storage pool that will host the tiered file systems or one can use a single HDP pool and pin DP Vols to the highest performing disks and use those as HNAS Tier 0.
- Do not mix HDP DP-Vols and plain parity groups in a single NAS server storage pool/span. However, it is acceptable for some NAS server storage pools/spans to use HDP DP-Vols while other NAS server storage pools use parity groups.
- For best performance, when creating a new NAS server storage pool based on HDP DP-Vols, specify all the DP-Vols in a single **span-create** command or NAS Manager equivalent. Do not create the storage pool on just a few of the DP-Vols and then make a series of small expansions.
- Create as many file systems as needed and expand them as often as necessary (or allow them to auto-expand). For maximum flexibility and responsiveness, create small file systems and allow them to auto-expand as data is written to them.



Note: The maximum size of a newly created file system on an HDP pool is 1 TiB, but file system expansions can increase the file system size to the maximum limit supported by your NAS server and underlying storage.

Disable zero page reclaim

HDP offers the ability to unmap individual pages within a file system. This capability is called zero page reclaim.

Consult the Hitachi Dynamic Provisioning software documentation for more information about zero page reclaim.



Important: Zero page reclaim should not be executed when HDP is used with HNAS servers.

Storage pool chunks

Storage pools are made up of multiple small allocations of storage called “chunks.”

The size of the chunks in a storage pool is defined when the storage pool is created, and a guideline chunk size is set. The guideline chunk size is between 500 MiB and 18 GiB. It determines the maximum size that the storage pool (span) can ever reach, because a storage pool can contain up to a maximum of 60,000 chunks.

Chunk size is an important consideration when creating storage pools. The size of the chunks in a storage pool is defined when the storage pool is created, and a guideline chunk size is set. The guideline chunk size is between 500 MiB and 18 GiB. It determines the maximum size that the storage pool (span) or file system can ever reach, because a storage pool or file system can contain up to a maximum of 60,000 chunks.

Larger chunks maximize scalability, but smaller chunks provide more granular file system expansion, because a file system always expands by a whole number of chunks.

If you create a storage pool using NAS Manager, the guideline chunk size is 18 GiB (the maximum allowable size). The default chunk size set by NAS Manager can be larger than the guideline chunk size calculated and suggested by the server if you created the storage pool using the CLI.



Note: When creating a storage pool using the HNAS server CLI, you can override the default guideline chunk size. When creating a storage pool using NAS Manager, you cannot change the guideline chunk size.

Storage pool pre-allocation

In normal operation, on HDP, the NAS Server prevents you from running out of real disk space. If creating or expanding a file system would use more space than is available on the parity groups or pool volumes in the DP pool, the operation fails safely. Without this check, if the DP pool ran out of space, write operations would fail and file systems would be forcibly unmounted.

An important part of this strategy is to pre-allocate real disk space for all the HDP pages in a chunk when that chunk is allocated to a file system. The server achieves this pre-allocation by writing one non-zero block to each HDP page. Without this pre-allocation, the free space on the DP pool would fall some time later, when data was first written to the HDP pages in the chunk; meanwhile, the server would overestimate the amount of free space on the DP pool, and would be in danger of running out of space.

Although pre-allocation protects your system from running out of disk space on a DP pool, it has a number of disadvantages:

- It slows down filesystem creation and expansion.
- Owing to the amount of time that it takes, the system limits the size of a new filesystem or a single manual expansion.
- Pre-allocation writes places stress on the storage, and flushes data out of the cache.
- Pre-allocation is not a sufficient safeguard if your storage system uses FMDs that perform data compression. On such systems, deleting compressible data from a file system and writing incompressible data in its place can cause the system to run out of space. The Administrator is therefore responsible for monitoring free space, expanding file systems and adding physical media as required.

It is possible to disable pre-allocation and enable the Administrator to manage free space by using the **span-hdp-preallocation** command. For further information, see the **span-hdp-preallocation** man page in the *Command Line Reference*.



Caution: When using this command, it is essential to ensure that no DP pool ever runs out of space. Otherwise, the file systems would be forcibly unmounted and would be unable to remount.

You can re-enable pre-allocation at any time using the same command. However, because some chunks may recently have been allocated without the pre-allocations taking place, some HDP pages in the file systems may not yet be mapped to real storage, and the server may overestimate the amount of free space available on one or more DP pools. It is therefore necessary to prevent filesystem auto-expansion until the server has written to all of the HDP pages in allocated chunks and obtained an accurate assessment of the amount of free space on DP pools.

Using HDP pools with a NAS server

You must configure the NAS server so that the HNAS software and HDP software can work together.



Important: No configuration changes are required on the NAS server to work with HDP pools.

The high-level process for using HDP pools with a NAS server is as follows:

Procedure

1. Configure the HDP pools and DP-Vols on the storage:
 - a. Use your storage configurator to create an HDP pool containing sufficient pool volumes to meet your immediate requirements for capacity and performance.
 - b. Using your storage configurator, create DP-Vols on the HDP pool. The total capacity of the DP-Vols should significantly exceed that of the pool volumes in order to fulfil future storage requirements. There should be enough DP-Vols to provide enough queue depth for good performance. As a guideline, DP-Vols usually have a capacity of 8 TiB.
 - c. Place the new DP-Vols into host groups and assign host LUNs to them. Before assigning host LUNs, enable Host Mode Option 7 and 68 on every host group so that the server can detect the new DP-Vols automatically. Alternatively, assign host LUNs to the DP-Vols and then run the server's `scsi-refresh` command.
2. Use the HDP-based storage on the NAS server:
 - a. Allow access to the DP-Vols (SDs) on the NAS server. You can allow access using NAS Manager or the command line interface (see the `sd-allow-access` command).
 - b. Create the NAS server storage pool from the HDP DP-Vols.



Note: If you plan to create tiered file systems, you must create a tiered storage pool using DP-Vols from two HDP pools.

To display a table that relates server device identifiers to DP-Vol internal LUNs, use the `sd-list --hdp` command. This information is useful when you create the storage pool. See the `span-create` man page.

- c. Create the file system on the storage pool.
- d. Format the file system.
- e. Mount the file system.

If you are using the command line interface, the **filesystem-create -b4** or **filesystem-create -b32** syntax can create, format, and mount a file system in a single step. See the *Command Line Reference* for your system.

Performing large file system expansions

Large file system expansions should not be performed when the system is heavily loaded.



Note: The following warning against large expansions applies only when you have not used **span-hdp-preallocation** to disable pre-alloc writes.

The system is heavily loaded when:

- experiencing heavy I/O load, especially write load
- rebuilding a parity group or pool volume after a disk failure
- formatting new disk space
- zero-initializing space after recent use of the **span-unmap-vacated-chunks** command.

Therefore, if a large expansion is requested, the **filesystem-expand** command prompts you to supply the **--storage-is-lightly-loaded** switch and will not proceed without it.

Upgrading from older HNAS systems

Any pre-existing storage pool (span in the CLI) should be left thickly provisioned after a recent upgrade.



Note: Run **span-unmap-vacated-chunks --unused-chunks** to reclaim space from any deleted file systems and wait for the zero initialization to finish. Thin provisioning can safely be used once the total space used on the pool equals the total size of the file systems on the storage pool.

When upgrading from an older version of an HNAS system, be aware that certain conditions produce the results and restrictions described in this section.

The conditions:

- A storage pool (span in the CLI) that was created in an earlier release violates the restriction that *SDs in each stripeset must come from one HDP pool*.
- A storage pool (span) contains a mixture of HDP DP-Vols and plain parity groups. This configuration is unsupported.

The following results are produced:

- Events are logged at boot time
- The `span-list` and trouble span will issue warnings
- Some operations fail cleanly, for example:
 - You cannot create a file system
 - You cannot expand a file system
 - You cannot delete a file system
- You can still load Cod
- You can still mount file systems



Note: `span-hdp-thickly-provisioned` will waive the usual rules if you will never use thin provisioning on the DP pool that hosts the span.

Using HDP storage

When working with HNAS systems, the HDP software supports up to two levels of tiered storage (Tier 0 and Tier 1).

See the *Hitachi NAS Platform HDP Best Practices* (MK-92HNAS063) for recommendations.

Considerations when using HDP pools

Consider the following when using the HDP pools:

As with storage pools based on parity groups:

- Deleting a file system is not always permanent. Sometimes file systems are recoverable from the recycle bin by issuing the `filesystem-undelete` command.
- Recycling a file system is permanent.

Unlike storage pools based on parity groups, on HDP-based storage pools:

- Freed chunks move to the vacated-chunks list, which is stored in Cod.
- Vacated chunks are reused when you create or expand file systems in the same storage pool.
- By reusing the same chunks, the server avoids exhausting space prematurely. Reusing chunks from recycled file systems allows the server to avoid wasting real disk space on deleted data. Instead, the server reallocates chunks from recycled file systems to other file systems.

Creating an HDP pool with untiered storage

Create the HDP pool and DP volumes for NAS server.

With untiered storage, tiers are not used. The metadata and the data reside on the same tier. The server has *no real perception of tiering* with untiered storage.

Creating HDP pools with tiered storage

Most storage pools reside on a single DP pool.



Important: The HNAS systems support up to two levels of tiered storage (Tier 0 and Tier 1).

With tiered storage, the metadata is kept on Tier 0 and the data on Tier 1. Tier 0 should be smaller than Tier 1, but should consist of faster storage. In general, the amount of real storage allocated for Tier 0 should be about 10% the size of Tier 1. The metadata Tier 0 contains is more compact than user data but is accessed more often. A single DP pool can support both the metadata and data tiers.

Storage pool naming

Storage pool naming rules and conventions.

A storage pool label may be one of the following types:

- A base name, which is the name that the storage pool or file system was given when it was created. All copies of a storage pool have the same base name. All copies of a file system have the same base name unless you decide to rename one or more copies. Storage pool labels are not case sensitive, but they do preserve case (labels will be kept as entered, in any combination of upper and lower case characters). Also, storage pool labels may not contain spaces or any of the following special characters:
 - Double quote (")
 - Single quote (')
 - Ampersand (&)
 - Asterisk (*)
 - Slash (/)
 - Backslash (\)
 - Colon (:)
 - Semicolon (;)
 - Greater than (>)
 - Less than (<)
 - Question mark (?)
 - Vertical bar/pipe (|)

Guidelines for choosing a good storage pool label include:

- The label should reflect the contents of the storage pool. Reasonably short, but distinctive and descriptive labels will help to guard against mistakes.
- Storage pool labels should be unique across the entire site, not just on the local cluster. If you move storage between servers or clusters, duplicate names will cause needless difficulty. Also, generic labels such as 'SAS_POOL_0' are best avoided, because they are not mnemonic and they are more likely to be duplicated among the clusters at a site. See the man page or the Command Line Reference for the storage-based-snapshot-label command for an explanation of rules and the various name types and interactions.
- A storage pool label should not be the same as a file system label.
- The storage pool label should not resemble a device ID (it should not be just a sequence of 1-4 digits).
- A snapshot name (or snap name), which identifies a single copy of the data. When you place a storage pool into snapshot mode, and every time you add a new snapshot, you specify a new snap name. Every snapshot of a given storage pool must have a unique snap name, although snapshots of different storage pools may have the same snapshot name.



Note: A snapshot name follows the same rules for special characters as a storage pool or filesystem base name, but snap name also cannot contain a dash (-).

- An instance name, which is automatically constructed from the base name and the snapshot name, if snapshots are used. The instance name identifies a single copy of a storage pool or file system.

For example, if a storage pool labelled 'Accounts' has a snapshot called 'Wednesday', the instance name of this snapshot is 'Accounts-Wednesday'. If the storage pool has a file system called 'External' then the instance name of this snapshot of the filesystem is 'External-Wednesday'. Most storage pool commands and file system commands expect instance names.



Note: When the server creates or loads a file system, file system details are stored in the server's registry, where they can be displayed by the **filesystem-list-stored** command. The server compares these labels as follows:

- Case-insensitive: a storage pool called 'AAA' cannot be loaded at the same time as a storage pool or file system called 'aaa'.
- Objects of the same type. For example, you cannot have two file systems with the same name, even if they are in different storage pools.
- Objects of different types. For example, if you have snapshotted storage pools with instance names 'Accounts-Main' and 'Accounts-DataMine', you cannot then create a new unsnapshotted storage pool or a file system with a label of 'Accounts'.

When labelling a storage pool, no storage pool may have the same "base name" or "instance name" as the base or instance name of any other loaded storage pool or file system.

Creating a storage pool using the CLI

You can use the CLI to create storage pools.



Note: For detailed information about the **span-create** command, see the CLI man pages. To create smaller filesystems, use the CLI instead of the GUI as it enables the use of smaller chunks.

Procedure

1. On the HNAS system, use the **span-create** command to create a storage pool using the SDs from the DP-Vols (on storage). For more information about the **span-create** command, refer to the *Command Line Reference*.

**Note:**

If you are using HDP:

- To avoid server timeouts when creating a new NAS server storage pool, wait for the HDP pool to finish formatting before creating the NAS server storage pool.

If the HDP pool has finished formatting, but the NAS server does not detect the new DP-Vols, run the `scsi-refresh` command so the NAS server will detect the new DP-Vols, though this should not be necessary when HMO 7 is enabled on the target port host group.

- If you are using HDP thin provisioning, list all the SDs in the initial `span-create` command. Do not run a single `span-create` command, then a series of `span-expand` commands.

Creating a storage pool using the GUI

With available SDs, administrators can create a storage pool at any time. After being created, a storage pool can be expanded until it contains up to 256 SDs.

When creating a tiered storage pool, to attain optimal performance, make sure that the SDs of the metadata tier (Tier 0) are on the highest performance storage type.

Procedure

1. Navigate to **Home > Storage Management > Storage Pools**, and click **create** to launch the **Storage Pool Wizard**.

The screenshot shows the 'Storage Pools' management interface. At the top, there's a breadcrumb trail: 'Storage Management > Home > Storage Management > Storage Pools'. Below this is a 'Filter' section with 'No Filtering Applied'. A table lists storage pools with columns: Label, Capacity, Used (%), Used, Free, and Status. One pool, 'sp1', is listed with a capacity of 2.04 TiB, 12% used (256.00 GiB), and 1.79 TiB free. Its status is 'Healthy'. Below the table are 'Actions' (allow access, deny access, create) and 'Shortcuts' (File Systems, System Drives).

Label	Capacity	Used (%)	Used	Free	Status
sp1	2.04 TiB	12 %	256.00 GiB	1.79 TiB	Healthy

2. Select either **Untiered Storage Pool** or **Tiered Storage Pool**.



Note: If you are creating a tiered storage pool, this will create **Tier 1 User-data**, of the tiered storage pool).

3. From the list of available SDs, select the SDs for the storage pool/tier.

Select at least four SDs for use in building the new storage pool/tier. To select an SD, select the check box next to the ID (Label).

An untiered storage pool cannot contain SDs on RAID systems with different disk types or RAID levels. Any attempt to create a storage pool from such dissimilar SDs will be refused.

A tiered storage pool can contain SDs on RAID systems with different disk types, as long as they are in different tiers. A tiered storage pool cannot, however, contain SDs with different RAID levels. Any attempt to create a storage pool with SDs that have different RAID levels will be refused.

For the most efficient use of storage capacity in an untiered storage pool or in a tier of a tiered storage pool, best practice is for all SDs be of the same capacity, width, stripe size, and disk size. However, after first acknowledging a warning prompt, you can create a storage pool with SDs that are not identically configured.

4. Specify the storage pool label.
5. Verify your settings, and click **next** to display a summary page.

The summary page displays the settings that will be used to create the storage pool/tier.

If you have already set up mirrored SDs for disaster preparedness or replication purposes, and you want the server to be aware of the mirror relationship, select the **Look For Synchronously Mirrored System Drives** check box.



Note: Before selecting the **Look For Synchronously Mirrored System Drives** check box, you must have finished configuring the mirrored SDs using the RAID tools appropriate for the system hosting the mirrored SDs. For example, you would use True Copy to create the mirrored SDs.



Note: The **Look For Synchronously Mirrored System Drives** check box is used only when setting up mirrored SD relationships using Hitachi TrueCopy. The **Look For Synchronously Mirrored System Drives** check box is not used with Hitachi Universal Replicator (HUR) or global-active device (GAD) software.

6. After you have reviewed the information, click **create** to create the storage pool/tier.
7. If you are creating an untiered storage pool, you can now either:
 - Click **yes** to create file systems (refer to the *File Services Administration Guide* for information on creating file systems).
 - Click **no** to return to the **Storage Pools** page without creating file systems.

8. If you are creating a tiered storage pool, you can now either:
 - Click **no** to return to the **Storage Pools** page if you do not want to create the metadata tier (Tier 0) of a tiered storage pool.
 - Click **yes** to display the next page of the wizard, which you use to create the metadata tier (Tier 0) of a tiered storage pool.
 - a. Specify which SDs to use in the tier by selecting the check box next to the SD label.
 - b. Click **next** to display the next page of the wizard, which is a summary page.
 - c. If you have mirrored SDs, for disaster preparedness or replication purposes, and you want the server to be aware of the mirror relationship, select the **Look For Synchronously Mirrored System Drives** check box.
 - d. After you have reviewed the information, click **add** to create the metadata (Tier 0) tier of the storage pool. A confirmation dialog appears, and you can now choose to create file systems in the storage pool, or you can return to the **Storage Pools** page.
 - Click **yes** to create file systems (refer to the *File Services Administration Guide* for information on creating file systems).
 - Click **no** to return to the **Storage Pools** page.



Note: After the storage pool has been created, it can be filled with file systems. For more information, see the *File Services Administration Guide*.

Adding the metadata tier

If you created a tiered storage pool, but only defined the SDs for the user data tier (Tier 1), you must now create the metadata tier (Tier 0).



Note: You can convert an untiered storage pool to a tiered storage pool using the `span-tier` command. For more information about this command, refer to the *Command Line Reference*.

To add a tier to a storage pool:

Procedure

1. Navigate to **Home > Storage Management > Storage Pools**.

Storage Management [Home](#) > [Storage Management](#) > Storage Pools

Storage Pools

Filter
No Filtering Applied
[filter...](#)

Label	Capacity	Used (%)	Used	Free	Status
sp1	2.04 TiB	12 %	256.00 GiB	1.79 TiB	Healthy

[Check All](#) | [Clear All](#)

Actions: [allow access](#) [deny access](#) [create](#)

Shortcuts: [File Systems](#) [System Drives](#)

2. Select the storage pool to which you want to add the tier. Click **details** to display the **Storage Pool Details** page.
3. Click the **Add a Tier** link to display the **Storage Pool Wizard** page.
4. Select the SDs to make up the metadata tier.
Using the **Storage Pool Wizard** page, select the SDs for the tier from the list of available SDs on the page. To select an SD for the tier, select the check box next to the **SD ID Label** in the first column. Verify your settings, and then click **next** to display a summary page.
5. Review and apply settings.
The summary page displays the settings that will be used to create the storage pool/tier.

If you have already created mirrored SDs for disaster preparedness or replication purposes, and you want the server to be aware of the mirror relationship, select the **Look For Synchronously Mirrored System Drives** check box.



Note: Before selecting the **Look For Synchronously Mirrored System Drives** check box, you must have finished configuring the mirrored SDs using the RAID tools appropriate for the system hosting the mirrored SDs. For example, for Hitachi Vantara storage systems, you would use True Copy to create the mirrored SDs.



Note: The **Look For Synchronously Mirrored System Drives** check box is used only when setting up mirrored SD relationships using Hitachi TrueCopy. The **Look For Synchronously Mirrored System Drives** check box is not used with Hitachi Universal Replicator (HUR) or global-active device (GAD) software.

Once you have reviewed the information, click **add** to create the second tier of the storage pool.



Note: After the storage pool has been created, it can be filled with file systems.

6. Complete the creation of the storage pool or tier.
After clicking **add** (in the last step), you will see a confirmation dialog.
You can now click **yes** to create a file system, or click **no** to return to the **Storage Pools** page. If you click **yes** to create a file system, the **Create File System** page will appear.

Allowing access to a storage pool

This procedure allows server access to an existing storage pool, but can also be used when a storage system previously owned by a server is physically relocated to be used by another server. The process restores access to the SDs that belong to the storage pool, and then restores access to the pool itself.



Note: Before moving the storage pool from one NAS server or cluster to another, refer to the *Command Line Reference* for the `span-assign-to-cluster` command, or view the `span-assign-to-cluster` man page for information on migrating a storage pool safely.

To allow access to a storage pool:

Procedure

1. Navigate to **Home > Storage Management > System Drives**.
2. Select one of the SDs belonging to the storage pool, and click **Allow Access**.
3. Select a pool, and click **details**. In the **Details** page for that storage pool, click **Allow Access**; then, in the **Confirmation** page, click **OK**.



Note: To become mountable, each file system in the storage pool must be associated with an EVS. To do this, navigate to the **Details** page for each file system in the storage pool and assign it to an EVS.

Creating storage pools with DP pools from HDP storage

After you have created an HDP pool with tiered or untiered storage, you can use DP-Vols to create storage pools.

See the CLI man pages for detailed information about commands.

Procedure

1. Use the command **span-create** or the NAS Manager equivalent to create the storage pool on the first HDP pool's DP-Vols.
2. Use the command **span-expand** to expand the storage pool on to the second HDP pool's DP-Vols.

Expanding the storage pool at the outset avoids the disadvantages of expanding it on a mature span. This is the only recommended exception to the rule of one pool per storage pool and one storage pool per pool.

3. When necessary, add new pool volumes to whichever pool needs them. Use the following steps:
 - a. Add parity groups or pool volumes.
 - b. If the amount of storage in the affected pool exceeds the total size of its DP-Vols, add more DP-Vols and use **span-expand**.

Moving free space between storage pools

You can move free space between storage pools that reside on the same HDP pool; however, before starting, consider the strong performance impacts.

The `span-list -s` command shows:

- The amount of vacated space on each HNAS stripeset
- The HDP pool that hosts the stripeset
- Other spans that share the HDP pool

Use the `span-list -s` command to determine how much space can be moved to other spans on the same HDP pool. If you recently deleted a file system and see less vacated space than expected, use the `span-list-recycle-bin` command to identify any file systems that have recently been deleted and still occupy space. To recycle these file systems, use the `filesystem-recycle` command, which makes it impossible to un-delete the file systems you specify on the command line.

The `span-unmap-vacated-chunks` command launches a background thread that may run for seconds, minutes, hours, days, or even months. See the man page for commands that monitor and manage its progress.

The free space on the DP pool continues increasing while this background thread runs.

In configurations where the storage has to zero-initialize (overwrite with zeros) HDP pages before they can be re-used, the free space on the pool may continue to increase even after the unmapping thread terminates.

The performance of all DP pools on the affected system are lower than usual until free space has finished increasing, but DP pools on other storage systems are unaffected.

Important unmapper considerations

Although not recommended, if multiple storage pools (spans) exist on a single pool, use the unmapper feature to move space between the storage pools on that pool.



Important: Using the unmapper commands can have serious consequences. It is strongly recommended that you read the CLI man page for each command.

Considerations:

- Unmapping vacated chunks does free space, but the performance of the storage is reduced until the server has zero-initialized all the space that you unmap. Never unmap chunks just to affect the appearance of available storage.
- You can unmap space on any number of spans at one time.
- The server has no commands for monitoring or managing the HDP zero-init process. Once the process starts, you must wait until it completes. The time can exceed many hours, even weeks in some cases.

Further reasons to avoid using the unmapper:

- In most storage configurations, an HDP page cannot be reused immediately after being unmapped. For security reasons, the page must first be zero-initialized to overwrite the previous page with zeros. This process occurs inside the storage, and cannot be monitored or managed by commands on the server.

The unmapper feature uses the following commands:

- **span-vacated-chunks** displays the number of vacated chunks in a storage pool and the progress of any unmapper.
- **span-stop-unmapping** cancels an unmapper without losing completed work.
- **span-throttle-unmapping** helps you avoid long queues of pages waiting to be zero-initialized.

Using the unmapper to move space between storage pools

If, after considering the consequences associated with use of the unmapper, you decide it is worth the significant performance impact, you can use the following steps to move space between storage pools.

Unmapping chunks (using the unmapper) does not increase the amount of free space in a storage pool; instead, the unmapper:

1. Takes chunks out of the VC list and returns the underlying HDP pages to the DP pool.
2. Removes the capacity from the chunks listed in the vacated chunks list by unmapping the space.
3. Zero-initializes the space.
4. Returns the freed capacity to the underlying DP pool.

After being added to the DP pool, the freed space can be used by other storage pools that reside on the same DP pool.



Note: See the CLI man pages, or the *Command Line Reference* for detailed information about the commands mentioned below.

The following procedure describes how to move space from storage pool S to storage pool T. This procedure is based on a configuration where both storage pools are based on DP-Vols from the same DP pool, and that DP pool is thinly provisioned.

Procedure

1. Delete and recycle a file system from storage pool S (Span S).
2. Run **span-unmap-vacated-chunks** on storage pool S.
3. Run the **span-list --sds T** command, and look at the amount of free space in the DP pool.

When the output from the **span-list --sds T** command shows that the DP pool has enough free space, create a new file system in storage pool T and/or expand one or more file systems in storage pool T.



Note: It may take a significant amount of time for the zero-initialization process inside the storage to complete and the amount of space free in the DP pool to increase. The amount of time it takes for this process is dependent on the type of storage, the amount of space being initialized, and utilization of the storage system.

If it takes too long to add the freed space to the DP pool, expand storage pool T onto a different DP pool (one that has available space).

Thin provisioning using the NAS server `filesystem-thin` command

HDP thin provisioning differs from thin provisioning achieved by the NAS server's `filesystem-thin` command.

- The `filesystem-thin` command instructs the server to report an artificially large file system to NFS and SMB clients. It bypasses free-space checks in programs that are not aware that file systems can expand on demand and reports large file system capacities to tenants even when the file systems are newly created and are still small.
- HDP thin provisioning instructs the block storage to report artificially large capacities for system drives. It enables spans to be expanded in smaller increments without loss of performance, and it widens the storage bottleneck by helping to spread I/O across all available storage media.

The two types of thin provisioning are commonly used together. They both enable Administrators to delay the purchase of new storage until it is actually needed.

To enable thin provisioning using the `filesystem-thin` command:

1. Create a small file system.
2. Confine it (using the `filesystem-confine` command) to a larger capacity.
3. Use the `filesystem-thin` command to enable thin provisioning on that file system.

Now, if applications or NFS and SMB clients ask for the capacity of the file system, they are told that the capacity is the confined amount, even though the actual amount of disk capacity allocated to the file system is less. As data is written to the file system, the file system grows, but only up to the maximum amount to which the file system has been confined, and no further.



Note: A NAS server storage pool can contain a mixture of thinly and thickly provisioned file systems.

System drive groups and dynamic write balancing

When used with a storage pool based on parity groups (RAID groups), dynamic write balancing ensures that the NAS server writes to all SDs in parallel, improving performance over older releases. Dynamic write balancing also improves flexibility by letting the server reflect the physical characteristics of the storage without the need to reconfigure spans.

Dynamic write balancing is enabled by default for storage pools. To implement dynamic write balancing, the NAS server requires some knowledge of the physical configuration of the storage. After SD groups have been configured (when using multiple logical devices from a single parity group of spinning disk), write operations are associated with SD groups rather than with SDs. Within each SD group, the NAS server scans the whole of one SD for free space before moving on to the next SD. For more information on SD Groups, see the **sd-group** man page.



Note: Dynamic write balancing can also be used with HDP. With two or more HDP pools, it helps to balance load across all available storage. Even with two or more stripesets on a single HDP pool, it maximizes the use of the available queue depth.

Optimizing dynamic write balancing performance

Although dynamic write balancing removes many of the restrictions of older allocation schemes, a few important guidelines still apply:

-  **Note:** Some software user interfaces have their own algorithms, usually creating even numbers of SDs in a stripe set
- Never divide storage into dozens of tiny SDs, and then create a storage pool from the many small SDs.

All the SDs in a RAID group or an HDP pool should be used in the same storage pool. If multiple SDs in an SD group are shared between storage pools, the SD group mechanism will not prevent the server from writing to several of the shared SDs at once. Writing to several of the shared SDs at once can cause a performance reduction, because one HDP pool may be driving the storage pool very hard, causing a slow-down on the other storage pools using the same resources.

Troubleshooting system drive groups

To help with server to storage troubleshooting, keep the following in mind:

- The server logs an event when a system drive becomes degraded on an HUS 100 family system. For example:

```
Warning: Device 0 (span "SPAN" ID 5B95587A30A2A328) : Device
reporting : SD 0: SCSI Lan Sense LU status reports DEGRADED"
```
- A **trouble** command reporter will identify an SD that may have higher than average response times.
- The output of the **scsi-devices** command includes the internal LUN value of any SD.
- For solutions with Hitachi Universal Replicator and TrueCopy, if the primary SDs are visible only to one server (or cluster) and the secondaries only to another server (or cluster), the **sd-mirror-remotely** command must be used to add SD mirror relationship information to the server's internal database.

Chapter 3: Using a storage pool

Storage pools contain one or more file systems, which consume space from the storage pool upon creation or expansion. A storage pool can also be used to control the auto-expansion policy for all of the file systems created in the storage pool. The following procedures describe how to create, delete, expand, remove from service, and rename a storage pool.

Once access is allowed to one system drive (SD) in a storage pool, that storage pool becomes visible in NAS Manager. If access is denied to all SDs in a storage pool, the storage pool is not visible in NAS Manager.

Expanding storage pools

To expand a storage pool without using Hitachi Dynamic Provisioning, add SDs to the storage pool. If the storage pool is tiered, you can add SDs to one tier without adding SDs to the other tier. To see how full each tier is, use the `span-list --sds` command.

If you are using Hitachi Dynamic Provisioning, see the *Hitachi NAS Platform HDP Best Practices* (MK-92HNAS063) for recommendations.

Expanding space in a thinly provisioned HDP storage pool

You can easily add space to a storage pool that uses thin-provisioned HDP.

The pool formatting process is non-disruptive, so the file systems stay mounted during the process.



Note: For detailed information about specific commands and how they are used, see the CLI man pages or the *Command Line Reference*.

Procedure

1. Create the pool volumes.
2. Use the `span-confine` command to confine the span.
3. Add the pool volumes to the HDP pool.
Adding the pool volumes automatically enables the **Optimize** check box on Hitachi Midrange storage systems.
4. Wait for the pool to finish formatting.
5. If required, use the `span-release` command to release the span on the HNAS system.
The HNAS system auto-detects the new space and lets you use it in new or existing file systems.
6. Check that the real disk space in the pool still does not exceed the total capacity of the pool's DP-Vols.

Expanding storage space using DP-Vols

Eventually, the total size of the HDP pool volumes reaches the total size of the DP-Vols. If the NAS server storage pool (span) needs more space, you can add space to it.

You can add as many HDP pool volumes as you want; however, you typically only need to add a small amount of space.



Note: See the CLI man pages or the *Command Line Reference* for detailed information about commands.

Procedure

1. Add the new pool volumes to the original pool.
2. If the new total capacity of the pool volumes exceeds that of the DP-vols, add more DP-vols to the same HDP pool.



Note: Make the new DP-Vols the same size and number as you originally created. All stripesets must be the same.

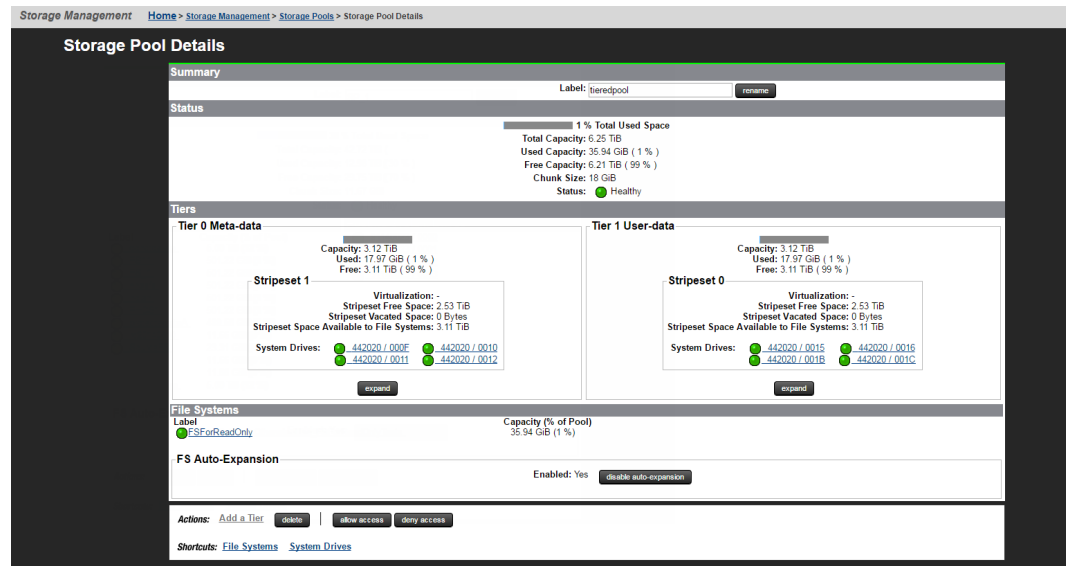
3. Wait for formatting to finish.
Otherwise, the file systems may auto-expand onto the new storage and find it so slow that the entire span fails.
4. If new DP-vols were created, use the `span-expand` command to expand the span on to the new DP-Vols.

Expanding a non-HDP storage pool or tier

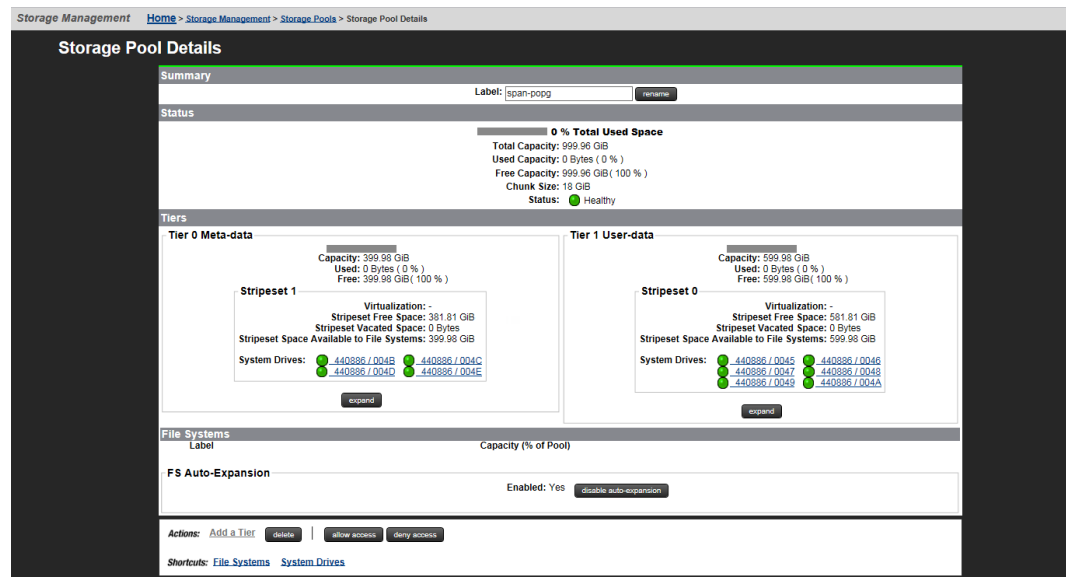
Procedure

1. Navigate to **Home > Storage Management > Storage Pools** to display the **Storage Pools** page.
2. Select the check box next to the label of the storage pool you want to expand, and click **details**.

For an untiered storage pool, the **Storage Pool Details** page shows File System and Hosting System Drive details:



For a tiered storage pool, the **Storage Pool Details** page shows the stripeset information and the specific System Drives in each tier. Clicking on a drive shows System Drive details:



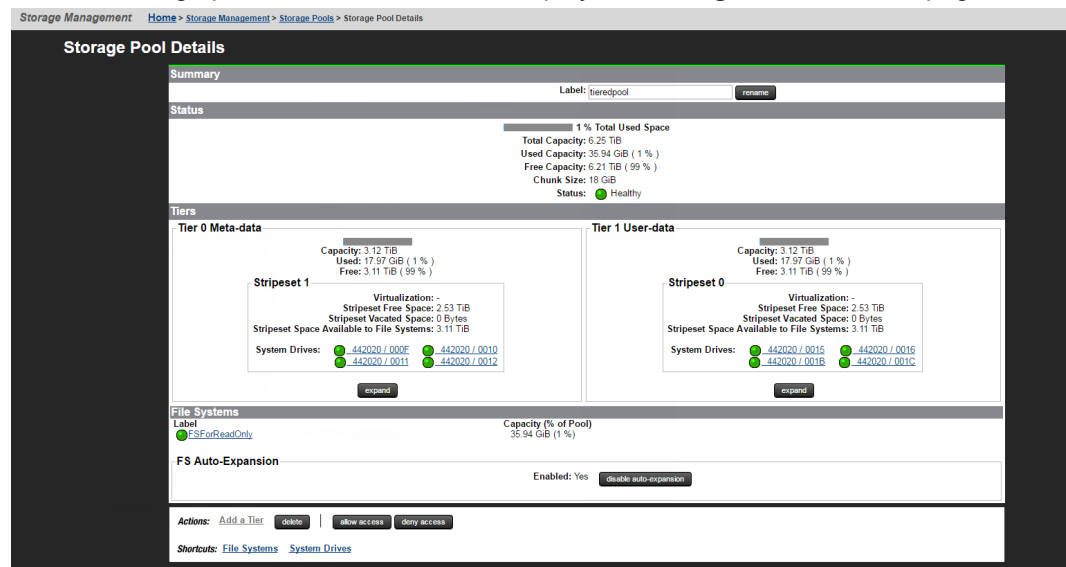
- To display the available system drives to add to the storage pool, click **expand**. To display the available system drives to add to a tier, select the tier you want to expand, and click **expand**. The **Storage Pool Wizard** page is displayed.
- Select the check box next to the label of the system drive you want to add, then click **next** to display the next **Storage Pool Wizard** page.
- Click **expand** to add the SDs to the storage pool/tier.

Configuring automatic file system expansion for an entire storage pool

Use this procedure to allow or prohibit automatic expansion of all file systems in the specified storage pool. This setting only affects auto-expansion; manual expansion of file systems in the storage pool is not affected by this setting.

Procedure

1. Navigate to **Home > Storage Management > Storage Pools**.
2. Select a storage pool, and click **details** to display the **Storage Pools Details** page.



3. Configure auto-expansion.
You can configure file system auto-expansion at the storage pool level as follows:

- Enable auto-expansion

Even if the storage pool is configured to allow its file systems to automatically expand, the file systems must also be configured to support automatic expansion. After a file system has expanded, its size cannot be reduced.

If file system auto-expansion is currently disabled, you can enable it by clicking **enable auto-expansion** in the **FS Auto-Expansion** option box.

- Disable auto-expansion

When automatic expansion of a file system is disabled, manual expansion of file systems in the storage pool is still possible.

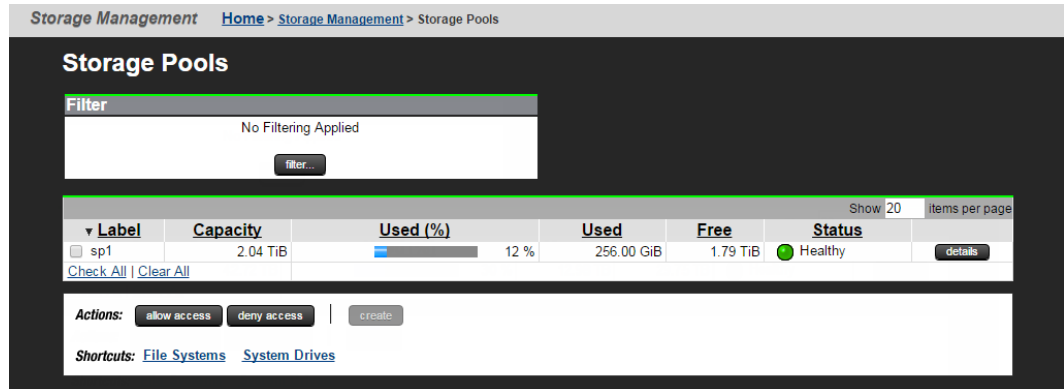
If file system auto-expansion is currently enabled, you can disable it by clicking **disable auto-expansion** in the **FS Auto-Expansion** option box.

Renaming a storage pool

The name for a storage pool can be changed at any time, without affecting any clients.

Procedure

1. Navigate to **Home > Storage Management > Storage Pools** to display the **Storage Pools** page.



2. Select a storage pool, and click **details**.
3. Enter a new name in the **Label** text box, and click **rename**.

Reducing the size of a storage pool

On an HDP pool, use the `span-unmap-vacated-chunks` command to return unused space to the pool. The actual capacity of the SD cannot be changed/shrunk; however, you can shrink the HDP pool to remove capacity and allocate elsewhere. Under no circumstances should the HDP pool be shrunk below the confined capacity of all associated file systems.

On a non-HDP pool, the size of the pool cannot be reduced.

Denying access to a storage pool

Typically, you deny access to a storage pool when you no longer want to use the storage pool with the local NAS server or cluster. If you want to use the storage pool with another NAS server or cluster, refer to the *Command Line Reference* for the `span-assign-to-cluster` command, or view the `span-assign-to-cluster` man page for information on migrating a storage pool safely.

Denying access to a storage pool:

- Unloads the storage pool and its file systems from memory, so that they do not appear in the NAS Manager list of storage pools or in the results of the `span-list -f` command.
- Unlicenses the host SDs.
- Deletes all the non-volatile memory contents for all file systems in the storage pool.



Caution: If a file system did not unmount cleanly before access to the storage pool is denied, all data written by clients and acknowledged by the server, but not yet been written to the storage, will be lost.

- Removes all file system to EVS bindings. If you later allow access to the storage pool, you will need to assign each file system to an EVS, either using NAS Manager or using the `evsfs add` command.



Note: Denying access to a storage pool does not delete the storage pool, the file systems, or the SDs.

Procedure

1. Navigate to **Home > Storage Management > File Systems** to display the **File Systems** page.

Storage Management [Home > Storage Management > File Systems](#)

File Systems

Filter
No Filtering Applied
[filter](#)

Label	Total	Used (%)	Used	Free	Storage Pool	Status	EVS
☒ data2	3.97 GiB	53%	2.12 GiB	1.85 GiB	sp1	Mounted	g5-evs2 details
☐ FileSystem2	3.97 GiB	54%	2.14 GiB	1.83 GiB	sp1	Mounted	g5-evs1 details
☐ FS	3.98 GiB	54%	2.15 GiB	1.83 GiB	sp1	Mounted	g5-evs1 details
☐ FS2	39.82 GiB	6%	2.48 GiB	37.34 GiB	sp1	Mounted	g5-evs1 details
☐ g5-fs1	200 GiB	2%	3.50 GiB	196.50 GiB	sp1	Mounted	g5-evs1 details
☐ TestFileSystem	3.75 GiB	64%	2.41 GiB	1.34 GiB	sp1	Mounted as Object Replication target	g5-evs1 details

[Check All](#) | [Clear All](#)

Actions: [mount](#) [unmount](#) | [create](#) [Download File Systems](#)

Shortcuts: [System Drives](#) [Quotas by File System](#) [Storage Pools](#) [Active Tasks](#) [Dedupe File Systems](#)

2. Identify the file systems in the storage pool to which you want to deny access, then select those file systems.
To select a file system, select the check box next to the file system label.
3. Unmount the selected file systems.
Click **unmount**, and in the confirmation dialog, click **OK**.
4. Click the **Storage Pools** shortcut to display a list of all pools, select a particular storage pool, and click **Deny Access**; in the confirmation dialog, click **OK**.



Note: This will also remove the pool from the storage pools list, but it will not be deleted.

Deleting a storage pool

A storage pool that does not contain file systems can be deleted at any time. If the storage pool contains file systems, you must unmount and then delete the file systems before you can delete the storage pool. After the pool has been deleted, its SDs become free and available for use by other storage pools.



Note: For detailed information about specific commands, see the CLI man pages or the *Command Line Reference*.

If you are using HDP:

1. Unmount the file systems in the storage pool using the `unmount` command.
2. Delete the file systems in the storage pool by issuing the `filesystem-delete` command or by using the NAS Manager.

If you delete an HDP-resident span and then reuse its DP-vols in a new span, the NAS server automatically launches an unmapper (as if you had run `span-unmap-vacated-chunks`). This unmapper recovers the leaked space that used to be occupied by the filesystem on the old span, so that no space is wasted if the new span's filesystems initially occupy less space than the old span's filesystems. However, the performance of the affected storage systems is lower than usual until they have reinitialized the space freed up by the unmapper.

To minimize the duration of this slowdown, do not delete and recreate HDP-resident spans unless you really need to, and as soon as you can, create and pre-expand file systems on the new span, reducing the amount of space that needs to be unmapped and reinitialized.

3. Delete the storage pool using the `span-delete` command.

You can use NAS Manager to delete a storage pool.

Procedure

1. Navigate to **Home > Storage Management > Storage Pools** to display the **Storage Pools** page.

The screenshot shows the 'Storage Pools' page in the NAS Manager. At the top, there's a breadcrumb trail: 'Storage Management > Home > Storage Management > Storage Pools'. Below this is a 'Filter' section with a text input and a 'filter...' button. The main content is a table with the following data:

Label	Capacity	Used (%)	Used	Free	Status
sp1	2.04 TiB	12 %	256.00 GiB	1.79 TiB	Healthy

Below the table, there are buttons for 'allow access', 'deny access', and 'create'. There are also links for 'Check All' and 'Clear All'. At the bottom, there are links for 'File Systems' and 'System Drives'.

2. Click **details** for the storage pool you want to delete.
The **Storage Pool Details** page opens.

3. Click **delete**, then click **OK** to confirm.

Chapter 4: Managing flash storage

The Unified NAS Server (NAS Modules) supports flash storage that performs transparent data-compression in hardware - i.e. FMD.

Using compression with flash module drives (FMDs)

Some storage contains flash module drives (FMDs) that transparently compress the data stored on them. Compression is a powerful way to reduce storage costs, but it does require extra care and monitoring by the administrator.

Not all data is equally compressible. For example, plain English text is readily compressible, whereas media files or archives that have already been compressed cannot be compressed again. Therefore, the amount of data that an FMD can store depends on the nature of that data. Even in the absence of file system snapshots, if a client deletes a large directory of plain text files and replaces it with a directory of media files of the same total size, more physical space will be needed, and the HDP pool may run out of space.

If the HDP pool runs out of space, it will be blocked, and all spans that use DP-Vols on that HDP pool will fail. The server issues warnings when physical space is low. However, the administrator must monitor physical space on the HDP pool and prevent it from running out of space.

Avoiding over-committing space on FMDs

To set up an HDP pool on FMDs, follow these steps in your storage configurator (such as Storage Navigator):

1. Create one or more parity groups on FMDs.
2. Enable compression on the parity groups. The storage configurator now treats each parity group as if it were 8 times its actual size.
3. Create LDEVs on these parity groups. Avoid filling up the available space with LDEVs. For example, if the physical capacity of a parity group is 10 TiB, switching on compression will enable you to create 80 TiB of LDEVs on the parity group. Never do this, because the FMDs cannot achieve an 8:1 compression ratio.
4. Create an HDP pool on the LDEVs. The LDEVs are now known as pool volumes.
5. Create DP-Vols on the HDP pool.
6. Assign the DP-Vols to host paths, and give them host LUNs. Use HDP thin provisioning; make the total capacity for the DP-Vols roughly three times that of all the underlying pool volumes.
7. On the server, license the DP-Vols (perhaps using `sd-allow-access`) and create a span (perhaps using `span-create`).

Ideally, the ratio of LDEV space to parity group space will be very slightly smaller than the compression ratio achieved by the FMDs. However, actually, it is impossible to make an accurate prediction of the compression ratio that FMDs will achieve before the data has been written. Indeed, the ratio will change as old data is deleted and new data is written. When setting up a new system, therefore, it is wise to assume that the FMDs will achieve no compression at all. On each 10 TiB parity group, set up only 10 TiB of LDEVs.

Once you have a mature span containing a good deal of data of the kind you wish to store, determine the compression ratio, create more LDEVs, and add them to the HDP pool. For example, if you find that the FMDs have achieved a 1.5:1 compression ratio, create a further 4 TiB of LDEVs for each 10 TiB parity group, or an extra 8 TiB for each 20 TiB parity group, and so forth. Assuming that the FMDs will achieve 1.4:1 when they have proved themselves capable of achieving 1.5:1 gives you a margin of error and ensures safety, provided that the nature of the data stored on the HDP pool does not change.

For further information on compression and how to avoid over-committing space, see the `span-fmd-compression` man page.

Expanding file systems on flash storage

If an untiered span (or a single tier of a tiered span) has stripesets on two or more compressed HDP pools, the server cannot determine which stripeset, and, therefore, which HDP pool can most safely supply space for a filesystem-expansion.

In this scenario, file systems cannot be auto-expanded. If all compressed stripesets are on the same HDP pool, or if two compressed stripesets are in different tiers, this condition does not apply, and auto-expansion works normally.

To help identify spans with multiple compressed pools, the `sd-list` command's `Con` column, which displays information about filesystem auto-expansion, displays an MCP symbol (Multiple Compressed Pools).

For example:

Span instance name	OK?	Free	Cap/GiB	System drives	Con
Span1	Yes	100%	19440	0,1,2,3;4,5,6,7	MCP

If an untiered span (or one tier of a tiered span) resides on two or more compressed HDP pools, `filesystem auto-expansion` will be disabled on that span, because the server cannot determine which HDP pool can most safely provide new space. The `span-expand` command will not place a span into this state without an override. If you try to add a second HDP pool to a span or a tier, `span-expand` requires you to supply the `--permanently-disable-filesystem-auto-expansion` switch to show that you understand the consequences. This switch is not accepted in any other circumstances.

Using manual file system expansion

During manual expansion of a span that resides on multiple compressed HDP pools, you must explicitly specify a stripeset by running the following CLI command:

```
filesystem-expand --on-stripeset
```

Alternatively, in NAS Manager, the **Expand File System** page shows a list of stripesets from which you can select if the server cannot choose:

Detecting free space on flash storage

The server performs a regular scan for spans on compressed HDP pools which are in danger of failing because free physical space is very low.

When free space is low, the server can take the following actions:

- On spans with low space where deleted filesystems exist but vacated chunks do not, the server recycles one or more deleted filesystems.
- On spans which have vacated chunks, the server launches a process inside the storage that slowly hands space back to the HDP pool and eventually increases physical free space.

To instruct the server to perform the next scan immediately instead of waiting up to three hours for the next scheduled scan, use the following command:

span-expedite-compression-space-check



Note: Use this command sparingly. It is possible that there are HDP pages waiting to be initialised inside the storage and, given time, the system can recover enough free space without taking action.

Using compression statistics

It is important to monitor compression-based storage to avoid running out of physical space.



Note: Only VSP Unified F/G/N Series (with NAS modules) that are configured using FMD compression will report compression statistics. FMD compression information is not available from the system when used with HNAS gateways.

The following commands display and control statistics for compression-based storage:

- **span-list**
- **span-dump**
- **sd-list**
- **span-set-cap-warn-thresh**

The **span-list -s**, **span-dump** and **sd-list --scsi** commands all display the following statistics for compressed storage:

- The amount of physical space that isn't in use
- The total amount of physical space
- The percentage saving by which compression has reduced the amount of physical space required to store data on the span

For example:

```
Compressed: 200GiB (0TiB) (3%) of 6843GiB (7TiB) is free; saving is 95%
```

Using the span-list command

For all storage types except UVM, the **span-list -s** command also displays an 'available' figure, which is a conservative estimate of the amount of space that could be allocated from the stripeset.

For example:

```
Set 0: 4 x 2430GiB = 9720GiB, of which 9119GiB is free, 299GiB is vacated,
900GiB is available
```

Using the sd-list command

The **sd-list** command also contains the following options for compressed storage:

- **--compression**
- **--is-compressed**
- **--is-uncompressed**

The **--compression** parameter reports the compression information in tabular form and can be combined with **--hdp** switch.

To view compressed or uncompressed SDs only, use the **--is-compressed** and **--is-uncompressed** switches (-e and -E), respectively.

To see which SDs are most urgently in need of new physical storage, a sort criterion is available:

```
sd-list --compression --sort e
```

This command sorts SDs by the percentage of physical space that is still free. Users can reverse the sort order by using an upper case E , and can combine this sort criterion with others.

Using the `span-set-cap-warn-thresh` command

Use the `span-set-cap-warn-thresh` command to set the capacity warning threshold for spans and compressed HDP pools.

This threshold controls:

- The point at which `span-list`, `trouble` and the event log warn about low physical space.
- The safety margin that `span-list -s` uses when calculating the available free space on a stripeset.
- The point at which the server takes precautionary measures such as disabling filesystem auto-expansion and unmapping vacated chunks when physical space is low.

Low physical space

If the physical space on an HDP pool is exhausted, the HDP pool will fail, along with every span that uses it. Therefore, when physical space is low, the server automatically disables filesystem auto-expansion in order to help keep the system serving data.

The span enters this state when the percentage of free physical space plus the percentage threshold set by the `span-set-cap-warn-thresh` command is below 100. For example, with the capacity warning threshold at 90%, filesystem auto-expansion will be disabled when less than 10% of the physical space is free, and will be re-enabled when more physical free space is provided.

When a span is in this state, the `sd-list` command's Con column displays LPS (Low Physical Space), to show that auto-expansion is disabled.

Using SVOS RF Data Reduction (ADR)

Only the HNAS Unified VSP F/G/Nx00 support FMD-based compression statistics. When using HNAS gateway models with any system that supports FMD compression, the statistics are not provided to the server.

The Hitachi NAS Platform currently offers hardware accelerated, fixed block de-duplication as a base feature of the NAS Operating System. This is a commonly deployed feature of the NAS OS that has been adopted by many customers. HNAS does not currently offer data compression. As such some customers, have expressed an interest in leveraging the native compression capability offered in SVOS RF ADR in conjunction with HNAS de-duplication to extend capacity savings and lower storage costs.

Customers commonly deploy HNAS with Thin Provisioned HDP Pools. This has been a “best practice” since the introduction of HNAS HDP Integration Support in HNAS OS Server release 12.1. HNAS customers typically deploy NAS spans or pools with HDP Pools that are overprovisioned at a factor of 200-300%. This mechanism allows customers to easily scale NAS capacity incrementally with zero disruption and automatic load-levelling. HNAS is able to obtain available physical free space from the storage system using SCSI based commands, such that the NAS will never allow capacity to be provisioned beyond the currently apportioned capacity to the associated HDP Pool, thereby safeguarding the file systems from unexpectedly running out of capacity.

Contrary to Hitachi Accelerated Flash module drives (FMDs) physical capacity statistics reporting with Hitachi VSP Unified or VSP N, SVOS RF-based data reduction does not currently report in-band the data reduction levels or physical pool utilization level of the host HDP Pool. As such a HNAS system attached to an HDP Pool that is leveraging SVOS RF ADR cannot accurately measure the physical free space available to the HNAS file systems.

Since the HNAS is unable to measure actual physical free space when ADR is deployed, it cannot prevent a catastrophic capacity exhaustion event – an attempt made by the system to write to a mounted file system that fails due to lack of available free space (that was previously provided by HDP Integration).

In such an event, the file system and its host span will fail and remain offline until additional capacity can be acquired and provisioned to the underlying pool. If this is the only HDP pool used by the NAS system, all NAS resource would be rendered offline, until the situation is rectified (by acquisition of additional capacity). While the use of SVOS RF ADR with HDP Pools attached to HNAS is a supported configuration, for these reasons it is not recommended.

Customers who choose to leverage SVOS RF ADR with HDP Pools used by HNAS, must acknowledge and understand that free space management must be performed by consulting the storage system itself for free space information. The storage system is able to report data reduction statistics and available free space, but it will not be able to account for HNAS capacity associated with vacated span chunks or file system free space. Vacated chunks are NAS storage pool chunks that were allocated from the HDP Pool to the NAS span, but whose host file system was deleted – the NAS considers the capacity as free space that is re-usable by other file systems in that span, but the system considers this as allocated.

Customers should also understand that since Data Reduction occurs at the block level (below the HNAS file system), that the amount of free space available to the HNAS will vary according to the changes in the data type and content and the system’s ability to compress the information. Consider that a file system into which was previously written GBs of highly compressible text will consume a much lower amount of physical capacity than an equivalent amount of uncompressible data (e.g. video, rich media), while the file system itself may not have changed its logical size. Therefore, changes in data type written to the system may not lower or raise the utilization level of NAS file systems, but they can alter the utilization level of underlying physical capacity.



Note:

Please use caution when using both HNAS deduplication and SVOS RF compression at the same time.

When doing so, it is recommended to disable filesystem auto expansion and periodically, review the physical pool utilization using management tools such as Device Manager - Storage Navigator, Hitachi Ops Center Analyzer, or RAIDcom/CCI command ("get system -key efficiency").

Chapter 5: Supporting external storage

The NAS server can support data on an external server using Hitachi Universal Volume Manager (UVM). UVM permits storage on external storage systems to be presented to the server as if the storage is local. To subsequently migrate data from the external storage onto the local storage, the server also supports Hitachi Storage Administrator Migrator functionality.

Using UVM

UVM enables storage that is accessed by one cluster to be taken over by another cluster. Using UVM instead of Universal Migrator enables the NAS server to preserve snapshots, quotas and ACLs. It also has the ability to replicate a whole span in a single operation.

Detecting UVM

The server detects UVM separately for each span. On a single system, some spans can reside on plain storage, some on HDP and some on UVM internal LUs.

UVM cannot be used with spans or file systems that are not supported on the new cluster. These include:

- Spans with only one admin area per system drive.
- File systems with object-based snapshots.

In these circumstances, use Universal Migrator instead of UVM.

Virtualizing LUs

UVM works with external and internal LUs. External LUs are LUs presented by other storage systems; the local storage system virtualizes these external LUs and presents them to the server as internal LUs, which appear to the server as if they resided on the storage system's own disks.

If the LUs to be virtualized are relatively large, the administrator must create a pool on the local storage with data direct mapping (DDM) enabled. The administrator uses the external LUs as the underlying storage for this DDM pool, and each DP-Vol in the DDM pool, which is presented to the server as an internal LU, corresponds exactly to one of these external LUs.

If the external LUs are relatively small, the administrator can still create a DDM pool, but may instead opt to virtualize the LUs directly.

In either case, the server can detect UVM internal LUs, and automatically treats the span as residing on UVM storage.

When using UVM, it is preferable to create the HDP pool on the system containing the external logical devices, therefore the HNAs will treat the HDP pool as if it were local disk in an HDP pool and not a pool of UVM disks.

Identifying a UVM span using the CLI

In the `span-list` command, for a UVM span, the 'Con' column displays the letters 'UVM'. However, the `span-list -s` command does not show how much free space is available at storage level, because that information is not available to the server.

Taking over a UVM span

A span which resides on external storage and is accessed by one cluster can be taken over by another cluster. For example, the span is on a storage system which is being accessed by Cluster A and you now want it to be accessed by Cluster B.

The following procedure also ensures that the two clusters never access the same span at the same time. This prohibition is necessary for data integrity.

To take over a span using UVM:

Procedure

1. On Cluster A:
 - a. Cleanly unmount all of the file systems.
 - b. Run the `span-assign-to-cluster <span-name> <UUID> <cluster-B-name>` command, specifying the name and UUID of Cluster B.
 - c. Run the `span-deny-access <span-name>` command.
2. Use LUN-mapping, switch-zoning or both to ensure that Cluster A can no longer access the LUs in the span.
3. Following the instructions in the *Hitachi Universal Volume Manager User Guide*, set up UVM internal LUs on Cluster B's local storage and make them visible to Cluster B by assigning host LUNs to them.
4. On Cluster B:
 - a. Depending on the configuration of your storage, it may be necessary to run the `scsi-refresh` command on the server to enable it to see the new internal LUs.
 - b. Run the `sd-list --hdp all:denied` command to identify the new UVM internal LUs.
 - c. Use the `sd-allow-access` command to license the UVM internal LUs. If there are no other unlicensed SDs, you can license the internal LUs more easily by running the `sd-allow-access all:denied` command.
 - d. Run the `span-wait-for-loading` command to give the span and filesystem time to load their disk configuration.
 - e. If the `span-list` command does not show the span, use the `sd-recover-from-changed-luids --span-base-name` and `span-rewrite-cod` commands to recover the span.
 - f. Use the `evsfs` command to bind each file system to an EVS.

- g. Mount each file system and create the necessary exports and shares.

Using multiple HDP pools

When you create or expand a filesystem, the server must find free chunks of disk space. If the span resides on two or more stripesets - that is, if the span has been expanded - then the server must select a stripeset to provide the space for each new chunk. If both stripesets' SDs are virtualizations of LUs on the same remote HDP pool, the server can select either stripeset with equal safety. However, if there are two or more remote HDP pools and they have different amounts of free space, the server has no way to tell which stripeset is the safest. That decision needs to be made by the Administrator.

In NAS Manager, if the server cannot select the stripeset, the **Expand File System** page shows a list of stripesets from which to select.

Through the command line, if a UVM-based span resides on two or more thinly provisioned remote HDP pools, the person provisioning the system must run the `span-uvm-thin-provisioning` command to warn the server. Each time a file system is expanded, the server will require the Administrator to select a stripeset using the `filesystem-expand --on-stripeset` command.

The person provisioning the system must run the `span-uvm-thin-provisioning` command once for each span that is in this configuration.

There is no `filesystem-create --on-stripeset` switch. Therefore, when a new filesystem is required, the administrator needs to:

- Check that all HDP pools contributing space to the span have enough space.
- Create the smallest possible filesystem.
- Use the `filesystem-expand --on-stripeset` command to expand it onto a stripeset that has enough disk space.

Expanding a UVM-based filesystem

With normal HDP pools, the server can detect how much disk space is available and it never allocates new chunks to a filesystem if no HDP pool in the span has enough space. The server also performs pre-allocation writes: when a chunk is allocated, the server writes a non-zero block to every HDP page in the chunk, so that the free space on the HDP pool falls immediately and the server has an accurate view of how much space is left. If an HDP pool were to run out of space, write operations would fail and file systems would unmount. It would be impossible to remount them until new space had been added to the HDP pool.

With UVM, the server has no way to determine:

- Whether the external LUs are thinly provisioned and, if so, how much free disk space is left.
- The page size of the HDP pool.

The server is, therefore, unable to guard against running out of disk space as effectively as it can with a local HDP pool.



Note: It is the Administrator's job to ensure that no external HDP pool ever runs out of space.

For a span residing on UVM LUs, the behavior of the NAS server changes as follows:

- **Auto-expansion**

- Because the server cannot determine that any filesystem expansion is safe, auto-expansion is disabled. Manual expansion is still permitted, but it is the Administrator's responsibility to ensure that any external HDP pools never run out of space.
- If a span resides on UVM storage, you can use Hitachi Storage Administrator Migrator functionality to migrate its LUs to a new, local HDP pool. While migration is in progress, the span is treated as still residing on UVM, even after some of its LUs have migrated to the new, local HDP pool. The server overlooks the fact that a stripeset is split between two pools (which would normally be a forbidden configuration), but filesystems do not auto-expand. As soon as all LUs have migrated, the server treats the span as residing on HDP: auto-expansion resumes, and the server performs its normal checks for free space before allocating chunks.

- **Pre-allocation writes**

- In order to help the Administrator form an accurate picture of the available free space, the server performs pre-allocation writes when a new chunk is allocated, just as it would on HDP. Because it cannot determine the external LUs' page size, it performs pre-allocation writes twice: once assuming 32 MiB pages (as used by HUS and AMS) and once assuming 42 MiB pages (as used by Enterprise storage platforms). However, if the external storage comes from a different vendor and uses a smaller page size, pre-allocation writes reduce the free space by less than the expected amount, and the rest of the reduction occurs later, when the Filesystem writes to the newly allocated chunks for the first time. The Administrator must take extra care when virtualizing non-Hitachi external storage over UVM.
- Performing two sets of pre-allocation writes does not take twice as long as performing one set, because HDP pages are mapped to real disk space only once. In rare cases where pre-allocation writes cause problems, they can be disabled for the system as a whole using the `span-hdp-preallocation` command. This makes filesystem-expansion faster, but also causes the server to over-estimate the available space on HDP and, therefore, can cause the Administrator to over-estimate the available space on UVM. Do not disable pre-allocation writes if there is any alternative solution.

- **Minimum SD rules**

- When any span is expanded, the server enforces minimum SD-counts to help maintain performance. With UVM, the server enforces the same minimum as for plain storage. However, whenever possible, we recommend adhering to the stricter HDP rules (see the `span-create` man page). Expanding on to more SDs helps to ensure that adequate queue depth is available.

- **Reuse of HDP pages**

- The server maintains a vacated-chunks list, just as on HDP, so that deleting and recycling one filesystem and creating or expanding another reuses the same chunks, instead of selecting new ones. If the external LUs are thinly provisioned, they benefit from the reuse of HDP pages that are already mapped to real disk space.
- The server cannot unmap HDP pages on the external LUs, so the `span-unmap-vacated-chunks` command does not run on a UVM-resident span.

- **DP-Vols**

- If a span resides on UVM internal LUs, you can expand it on to DP-Vols from a local HDP pool. You can also expand on to DDM LUs from any DDM pool.
- You cannot expand a span on to UVM internal LUs if it resides entirely on HDP DP-Vols.

Migrating data to a local cluster

It is possible to take a span that is virtualized by UVM and migrate it to the local storage, using Hitachi Storage Administrator Migrator functionality.

The two principal advantages of migration are:

- A span residing locally is faster than one accessed over UVM (especially if the local storage is newer).
- The server regains the ability to check for free space and protect the HDP pool from running out of space.

If the existing span violates the HDP rules introduced in release 12.1, for example, by spreading a single NAS stripeset across DP-Vols from multiple HDP pools, a migration can bring it into compliance. The best practice is to have one span per thinly provisioned HDP pool.

Making efficient use of disk space

To migrate data to a local HDP pool, that local HDP pool must be thickly provisioned, even if the external LUs are not. For example, for a remote HDP pool with six 10 TiB pool volumes and twenty 8 TiB DP-Vols, on migrating the data to local storage, the local HDP pool needs to provide 160 TiB of pool volumes or parity groups, instead of the 60 TiB that were present on the remote storage.

If the remote HDP pool contains multiple spans, it is possible to mitigate this inefficiency as follows:

Procedure

1. Identify the span with the largest amount of free space - that is, the largest difference (in TiB) between the total capacity of the filesystems and the capacity of the span.
2. If you have recently deleted filesystems from the span and you are confident that you will not need to undelete them, run the `filesystem-recycle --all-filesystems` command against the span.
3. Migrate this span using Hitachi Storage Administrator Migrator functionality. Wait for migration to complete.
4. When migration is complete, the free space on the local HDP pool falls by the capacity of the span (rather than just the capacity of the filesystems on it).

5. Run the **span-unmap-vacated-chunks --unused-chunks** command against the span. This command launches a background process that gives unused space back to the local HDP pool. Within a few minutes, the free space on the local HDP pool starts to rise slowly as the storage zero-initializes the HDP pages that the server has returned to the HDP pool. It keeps rising as the storage continues to zero-initialize HDP pages, even after the unmapping process logs an event on the Admin Service to say that it has finished. To reduce the performance impact on the system, use the **span-throttle-unmapping** command. You can check the free space on the HDP pool using the **span-list --sds** command, but the server has no way to monitor or manage the zero-initialization process.
6. As soon as the local HDP pool has space, migrate the second span, selected in the same way as before.
7. Continue with this process until all spans are migrated.

When to avoid block-level migration

For optimal performance, a span should have a single stripeset (it should never have been expanded). The stripeset should contain a large number of system drives, and it should reside on a thinly provisioned HDP pool. The more a span differs from this ideal configuration, the poorer the performance, even after migration to new storage.

The worst case is a span for which the **span-list -s** command shows a large number of stripesets of one or two SDs each and also where the **span-space-distribution** command shows that the filesystems are not evenly spread across those stripesets.

A second configuration that cannot be helped by migration is a filesystem that is much smaller than the current maximum size, but which, according to the **filesystem-scalability** command, cannot expand because it has used up most or all of the available chunk runs.

For these configurations, use Object Replication or Universal Migrator.

Hitachi Vantara



Corporate Headquarters

2535 Augustine Drive

Santa Clara, CA 95054 USA

HitachiVantara.com | community.HitachiVantara.com

Contact Information

USA: 1-800-446-0744

Global: 1-858-547-4526

HitachiVantara.com/contact