IBM Cloud Object Storage System™
Version 3.14.1

Storage Pool Expansion Guide

IBM

Contents

Document information

This guide explains how to add storage capacity using the Storage Pool Expansion feature.

Chapter 1. New capabilities

Storage Pool Expansion offers a new method to increase storage capacity of original Vaults.

A new set of Accesser® Nodes (a *Device Set*) is added to an original Storage Pool. After the new Device Set is added, a portion of the data stored on the original Device Set is reallocated to the new Device Set. This increases the storage and throughput capacity of that Storage Pool by an amount corresponding to the new Device Set.

Prior to version 3.6.0, storage capacity could be increased in three ways:
1. Adding drives to partially populated Slicestor® Nodes
2. Merging an original Storage Pool to a newly added Storage Pool
3. Creating additional Storage Pools

Note: Only the first two options increased Vault capacity.

Merging Storage Pools created *generations* of Storage Pools. This process worked with Simple Object deployments but not with Cloud Storage Object (CSO) deployments.

Simple Object methodology has the system create the object name. This name can be encoded to indicate deterministically on which generation an object resides. (A deterministic system can predict the output with 100 percent certainty.)

Named Object methodology uses an object name that the end user provided. It is not possible to encode generation information to this object. Without some other mechanism, Named Objects can store and access objects deterministically only on the first generation.

As metadata was always stored to the first generation of a Vault using Named Object, when more and more objects are stored, the first generation may fill eventually and be unable to accept new metadata, index, synced or embedded objects (making the system unable to accept new writes even if other generations have available capacity). Once the first generation is full, the Vault would effectively become read only.

Storage Pool Expansion introduces four changes:
1. Moving away from multi-generational storage to multiple Device Sets
2. Using a new algorithm to make deterministic actions in storing objects on one of many possible Device Sets
3. Older device sets can be replaced with a new storage pool set.
4. More capacity can be added to a partially populated device set

Replacement is intended to remove old device sets that contain outdated hardware. Existing device sets can be resized to add more capacity.

The weighting used in this new algorithm is based on the relative capacity of each Device Set. This has the added benefit of ensuring that all Device Sets fill at the same rate.

Vaults that use Named-Object interfaces such as CSO- or OpenStack-compatible APIs can now be scaled up by adding one or more Device Sets to increase capacity of an original Storage Pool.

Chapter 2. Benefits

Benefits of the Storage Pool Expansion.

Replacing an old device set with a new storage pool set.
- Allows old, outdated hardware to be removed from the Storage Pool and replaced with new, better, faster, larger capacity hardware.
- Multiple older sets can be replaced immediately with a new replacement set. The operation does not require a one-to-one replacement.

Resizing an existing device set.
- Allows the user to add capacity to partially populated Slicestors®. The current software will not properly use any additional capacity that is added after the creation of the storage pool set.

Expanding a Storage Pool by adding Device Sets.
- Eliminates the performance and metadata scalability bottleneck of named-object operations.
- Equalizes storage capacity and usage across all nodes and drives in the system which.
  - Keeps all nodes capable of accepting new data, which increases maximum ingest rate.
  - Spreads servicing metadata reads and writes across all nodes, which increases OPS capacity.

Chapter 3. Terminology and components

Important terms

These terms are used throughout this document.

Node

A generic term for a component that encompasses all types of that component: appliance, virtual appliance, application, or Docker container.

Device Set

A set of Slicestor® Nodes of a multiple of the Least Common Multiple of an IDA Width that either forms a new Storage Pool or expands an original Storage Pool. The amount of Slicestor® Nodes in a Device Set must be an integer multiple of the Least Common Multiple of all the IDA widths for Vaults the Device Set serves.

Storage Pool

A grouping of one or more Device Sets which together provide the physical storage resources for one or more vaults.

Source Slicestor® Node

A Slicestor® Node that sends Slices during reallocation to its destination peer in the new Device Set. During reallocation, this node serves as a proxy for all operations to its destination peer so that reallocated Slices remain visible and accessible to Accesser® Nodes and rebuild processes.

Destination Slicestor® Node

A Slicestor® Node that receives Slices that are sent during reallocation from its corresponding source peer in the current Device Set. During reallocation, this node handles operations that are proxied by source nodes.

Weighted Rendezvous Hash (WRH)

The algorithm that is used to select which Device Set manages a given name. This algorithm minimizes reallocation upon any change by deterministically mapping an object to its Device Set. The algorithm can handle changes to the underlying system efficiently.

Slicestor® Node Proxying

As Slices are reallocated, the source Slicestor® Node tracks which Slices are successfully moved to the destination Slicestor® Node. Since the source Slicestor® Node no longer holds the Slice after it is reallocated, and because Accesser® Nodes do not communicate to destination Slicestor® Nodes during reallocation, the source Slicestor® Node must proxy the read, write, and list requests it receives to be serviced by the destination Slicestor® Node. When the destination Slicestor® Node responds to its request, the source Slicestor® Node returns the destination Slicestor® Node response to the client.

Throughput

The total inbound and outbound rate that the Slicestor® Node can process, usually expressed in Megabytes per second (MBps).

OPS

The number of possible operations (reads / writes) per second.

Latency

The time to process an HTTP request.

Concentrated dispersal

A technique whereby each Slicestor device may be responsible for multiple slices of a given object stored in a system. For instance, a system of three Slicestor devices can have an IDA width of 36, using twelve slices for each Slicestor. To use a Concentrated Dispersal vault, a device set must be created using between three and six Slicestor devices.

Multiple of a Least Common Multiple

If Vaults exist in the original Storage Pool with IDA widths of 6, 9, and 24, then the new Storage Pool would need a multiple of 72 Slicestor® Nodes.

Table 1. IDA Values

IDA Width	2x	3x	4x	5x	6x	7x	8x	9x	10x	11x	12x	13x
6	12	18	24	30	36	42	48	54	60	66	72	78
9	18	27	36	45	54	63	72	81	90	99	108	117
24	48	72	96	120	144	168	192	216	240	264	288	312

Device sets

Device Sets add a layer of organization to the components.

These components are organized in a matrix:
- Each Vault resides on one Storage Pool.
- Each Storage Pool is composed of at least one Device Set.
- Each Device Set is composed of an integer multiple of an IDA width number of Slicestor nodes.
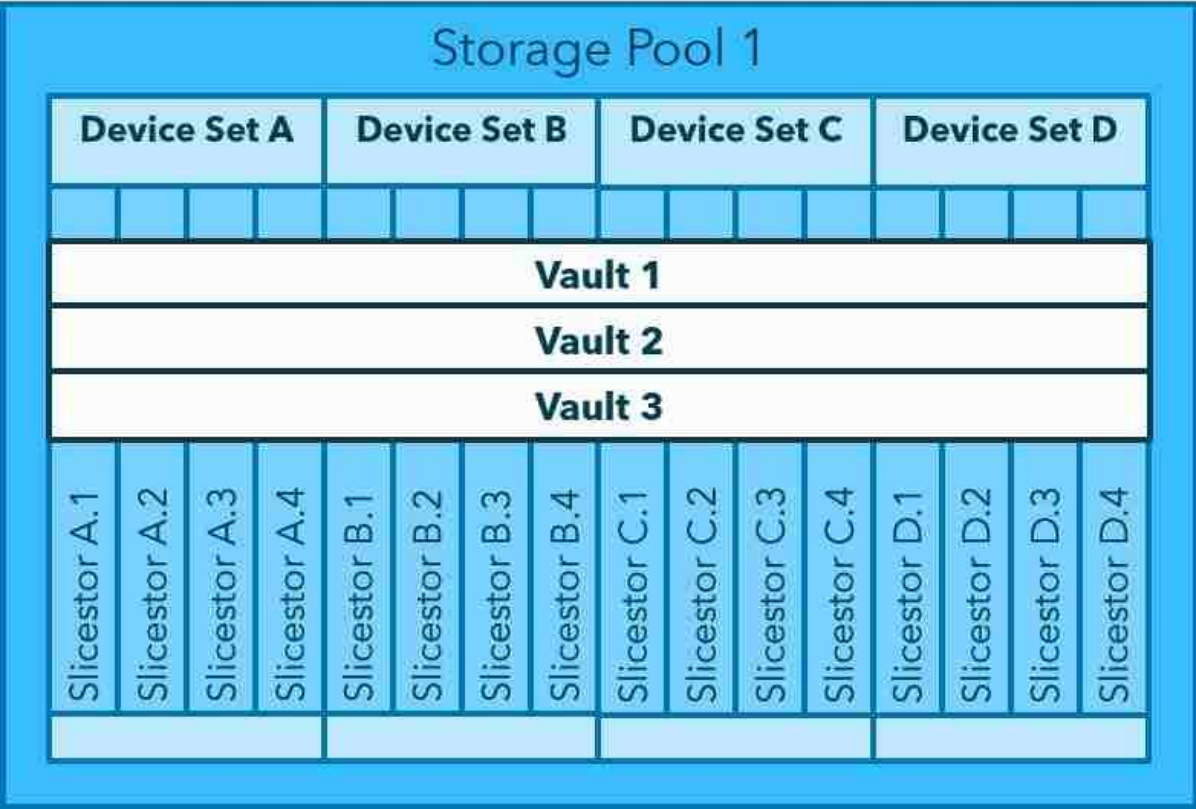
Figure 1. Matrix of storage resource organization

Chapter 4. Workflow

The following information covers the Storage Pool workflow.

When a Storage Pool is expanded with a new Device Set, the expected location of a subset of Slices changes due to the weighting of each set. The system determines which Slices need to be moved, or reallocated, to the new nodes. Slices are reallocated peer-to-peer between Slicestor® Nodes. Slices are not reconstituted first into objects then recalculated and sliced. Only those Slices for which the Algorithm calculates a new location are reallocated.

All reallocation occurs without impacting data availability.

When performing a device set replacement or resize, in addition to expansion, the system is in a reallocation state.

System behavior

When the system is expanded, it is considered to be in a reallocation state.

When all reallocations are completed, the system is in the stable state.

In the stable state, the Accesser® Nodes communicate directly with all Slicestor® Nodes and Slicestor® Nodes from different Device Sets do not communicate with each other. All Slicestor® Nodes manage storage operations. All Slices reside in the locations that are indicated by the Algorithm.

In the reallocation state, Slices do not yet reside in their proper Device Set. The Accesser® Node is unaware of the new Device Set and communicates with the source Slicestor® Nodes. The source Slicestor® Node accepts the Slice read or write request and requests the Slice from, or sends the Slice to, the destination Slicestor® Node if needed.

After all reallocation is complete, the system returns to the stable state.

Determine slice allocation

The Slicestor® Nodes use the Weighted Rendezvous Hash algorithm to determine where to move the Slices. It is determined by the utilization of each Device Set.

Example 1. How the WRH algorithm works

The WRH algorithm takes a name and the list of Device Sets and the capacities as inputs and returns the Device Set that should store that object.

Consider a Storage Pool with three Device Sets: A (10 PB), B (25 PB) and C (20 PB). The total of all the weights are 10 PB + 25 PB + 20 PB = 55 PB.

The probability that any data source name is mapped to the Device Sets is:
1. (10 PB / 55 PB) = 18.18%
2. (25 PB / 55 PB) = 45.45%
3. (20 PB / 55 PB) = 36.36%

Note: When new data is written, all the Device Sets fill up at a rate that equalizes their storage usage as a percentage of the total storage capacity. In the stable state, all Device Sets (and all Slicestor® Nodes within them) are used equally.

Any time the capacity of a Device Set changes or a Device Set is added, some Device Sets have more data than others. This algorithm is reapplied to all names to store the Slices in their proper locations and restore equilibrium.

Add new storage resources to a storage pool

Expanding a storage pool starts with one original device set.

In the figure below, a storage pool with one device set and a storage capacity of 10 PB adds one device set with 10 PB of capacity. It changes the algorithm that determines where the data is stored, triggering a reallocation. The IBM Cloud Object Storage Slicestor® nodes in each original Device Set reallocate a fraction of their slices directly to Slicestor® Nodes within the added Device Set. The fraction of slices each Slicestor® Node reallocates is equal to the ratio of the added storage that is compared to the new total of storage. In this case, the ratio equals [10 PB / 20 PB] or 1/2.
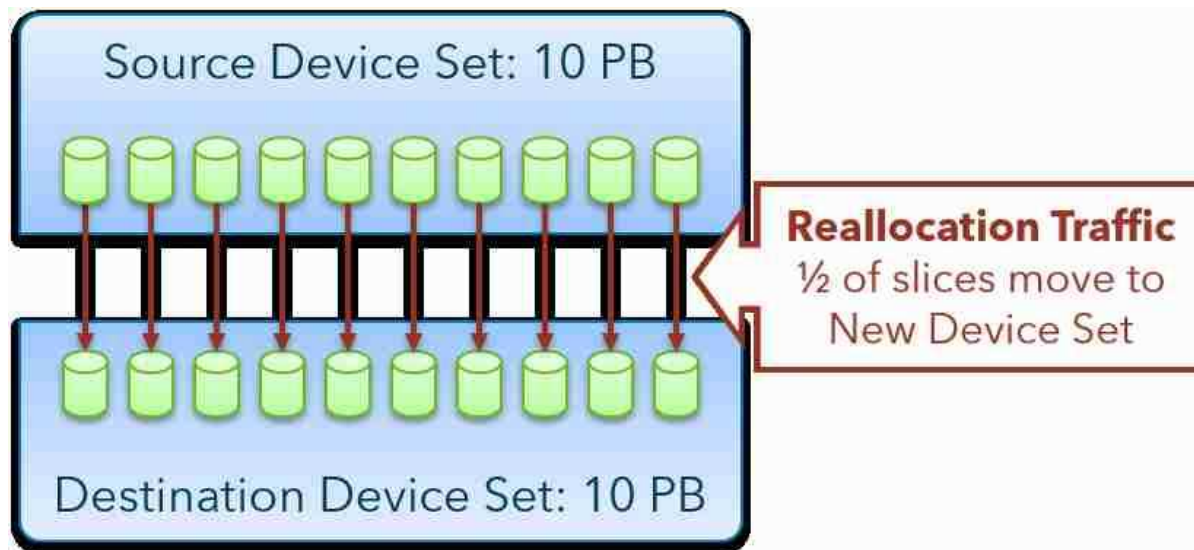
Figure 2. Adding storage resource to a storage pool

Reallocation progress within each Slicestor® Node

Each source Slicestor® Node tracks its progress in processing Slices for reallocation.
1. List Slices across an unprocessed portion of the data to determine what Slice names exist in that portion.
2. Apply the Algorithm to each Slice name returned by the listing to determine whether that Slice should be reallocated to the destination Slicestor® Node.
3. Reallocate the Slice to the new Device Set, if the Algorithm indicates that the move is needed.
4. Update the internal tracking to reflect the portions of the Slices that are processed.
5. Repeat the steps if unprocessed ranges remain.
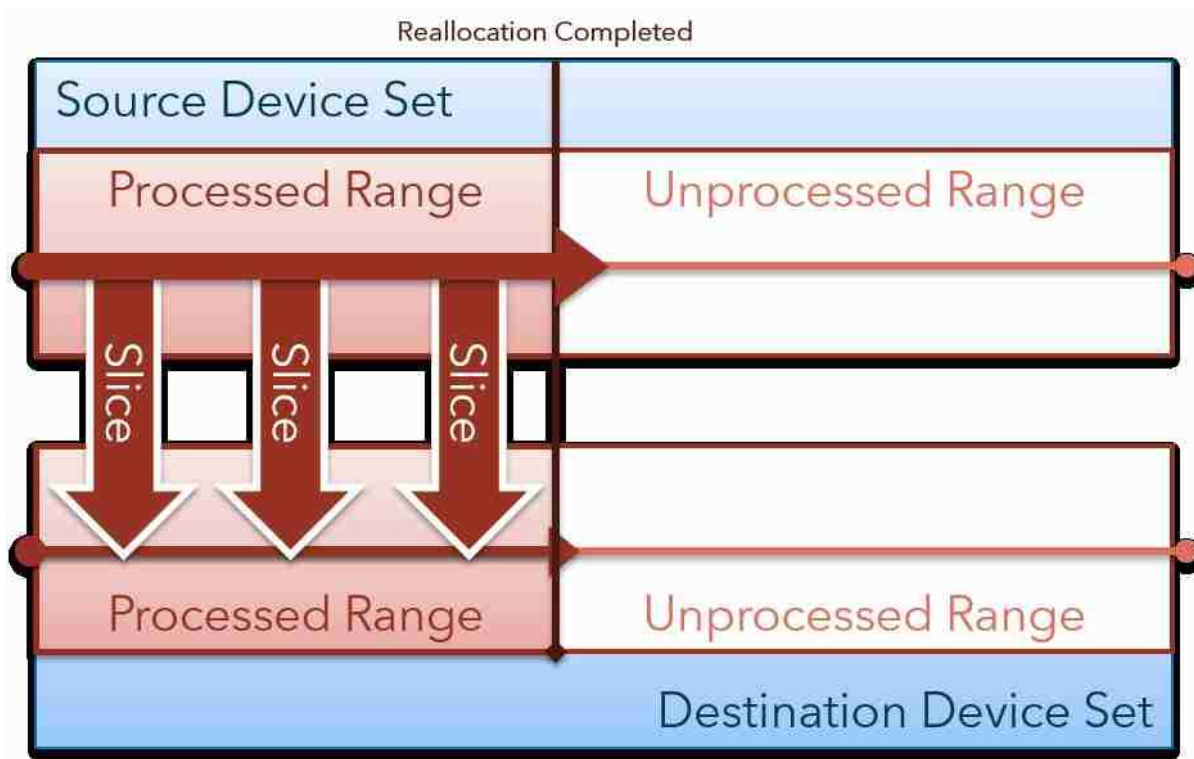
Figure 3. Simplified example of reallocation progress

During Reallocation, some Slices of an object can exist on the source Device Set while other Slices of that object exist on the destination Device Set.

Setting replacement slice

Slices are reallocated when storage resources are removed from the pool and replaced with a new resource.

At a high level, all the slice data on the replaced sets is moved to the remaining and new sets in the pool. The slices are reallocated according to the capacity of the remaining storage resources.

For example, a storage pool has three sets: 10 PB, 20 PB, and 25 PB for a total of 55 PB. The percentages of each set are:

1. 10 PB / 55 PB = 18.18%
2. 20 PB / 55 PB = 36.36%
3. 25 PB / 55 PB = 45.45%

The storage pool is storing 40 PB of data. Based on the set capacities and their respective percentage of the total capacity, this much data is stored on each set (40 PB / 55 PB = 72.72% full):

1. 10 PB * (40 PB / 55 PB) = 7.27 PB
2. 20 PB * (40 PB / 55 PB) = 14.54 PB
3. 25 PB * (40 PB / 55 PB) = 18.18 PB

The user decides to replace the 10 PB set with a new, larger device set of 30 PB. After initiating replacement, the percentages of each set are:

1. 30 PB / 75 PB = 40.00%
2. 20 PB / 75 PB = 26.67%

3. 25 PB / 75 PB = 33.33%

The 7.27 PB stored on the 1-PB set reallocates onto the replacement set. In addition, because the 20 PB and 25 PB sets now make up a lesser percentage of the total capacity, some of the slice data is reallocated to the 30 PB set. In the end, the three sets should have a percentage of stored data roughly equal to their percentage of the total capacity.

1. 30 PB * (40 PB / 75 PB) = 15.99 PB
2. 20 PB * (40 PB / 75 PB) = 10.67 PB
3. 25 PB * (40 PB / 75 PB) = 13.33 PB

Setting resize

Set Resize will reallocate slices between existing sets in a Storage Pool after more capacity is added.

For example, a storage pool has three sets: 10 PB, 20 PB, and 25 PB for a total of 55 PB. The percentages of each set are:

1. 10 PB / 55 PB = 18.18%
2. 20 PB / 55 PB = 36.36%
3. 25 PB / 55 PB = 45.45%

The storage pool is storing 40 PB of data. Based on the set capacities and their respective percentage of the total capacity, this much data is stored on each set (40 PB / 55 PB = 72.72% full):

1. 10 PB * (40 PB / 55 PB) = 7.27 PB
2. 20 PB * (40 PB / 55 PB) = 14.54 PB
3. 25 PB * (40 PB / 55 PB) = 18.18 PB

Both the 10 PB and 20 PB sets are partially populated with a max capacity of 25 PB each. The user adds the additional drives to the devices in these sets to reach 25 PB total capacity in each. It results in a storage pool with the following percentages:

1. 25 PB / 75 PB = 33.33%
2. 25 PB / 75 PB = 33.33%
3. 25 PB / 75 PB = 33.33%

Because the 10 PB and 20 PB now have more capacity, the data is reallocated to be roughly even across all sets. For the 20 PB set, it has more capacity but a lower percentage of the total capacity so it holds less data than before. In the end, the three sets should have a percentage of stored data roughly equal to their percentage of the total capacity:

1. 25 PB * (40 PB / 75 PB) = 13.33 PB
2. 25 PB * (40 PB / 75 PB) = 13.33 PB
3. 25 PB * (40 PB / 75 PB) = 13.33 PB

Setting removal

Slices are reallocated when storage resources are removed from the pool.

At a high level, all the slice data on the removed sets is moved to the remaining. The slices are reallocated according to the capacity of the remaining storage resources.

For example, a Storage Pool has three sets: 10 PB, 20 PB, and 25 PB for a total of 55 PB. The percentages of each set are:

1. 10 PB / 55 PB = 18.18%
2. 20 PB / 55 PB = 36.36%

3. 25 PB / 55 PB = 45.45%

The storage pool is storing 40 PB of data. Based on the set capacities and their respective percentage of the total capacity, this much data is stored on each set (40 PB / 55 PB = 72.72% full):

1. 10 PB * (40 PB / 55 PB) = 7.27 PB
2. 20 PB * (40 PB / 55 PB) = 14.54 PB
3. 25 PB * (40 PB / 55 PB) = 18.18 PB

When a 10 PB set is removed, the percentages of each set are:

1. 20 PB / 45 PB = 44.44%
2. 25 PB / 45 PB = 55.56%

The 7.27 PB stored on the 1-PB set reallocates onto the remaining sets. In the end, the two sets should have a percentage of stored data roughly equal to their percentage of the total capacity:

1. 20 PB * (40 PB / 45 PB) = 17.77 PB
2. 25 PB * (40 PB / 45 PB) = 22.22 PB

Changes to fulfilling requests

Where a Slice resides during the reallocation determines how the Slicestor® Node processes a request.

All Slices are available always both in the stable and reallocation states. Having the source Slicestor® Node serve as a proxy keeps operations transparent to Accesser® Nodes.

- If a Slice is reallocated and the system is still in the reallocation state, the sourceSlicestor® Node will proxy any request for that slice.

   Note: It could add latency due to the extra round-trip time between the source and destination Slicestor® Nodes. More Slices are proxied as reallocation progresses.

- All other requests are handled normally by the source Slicestor® Node.
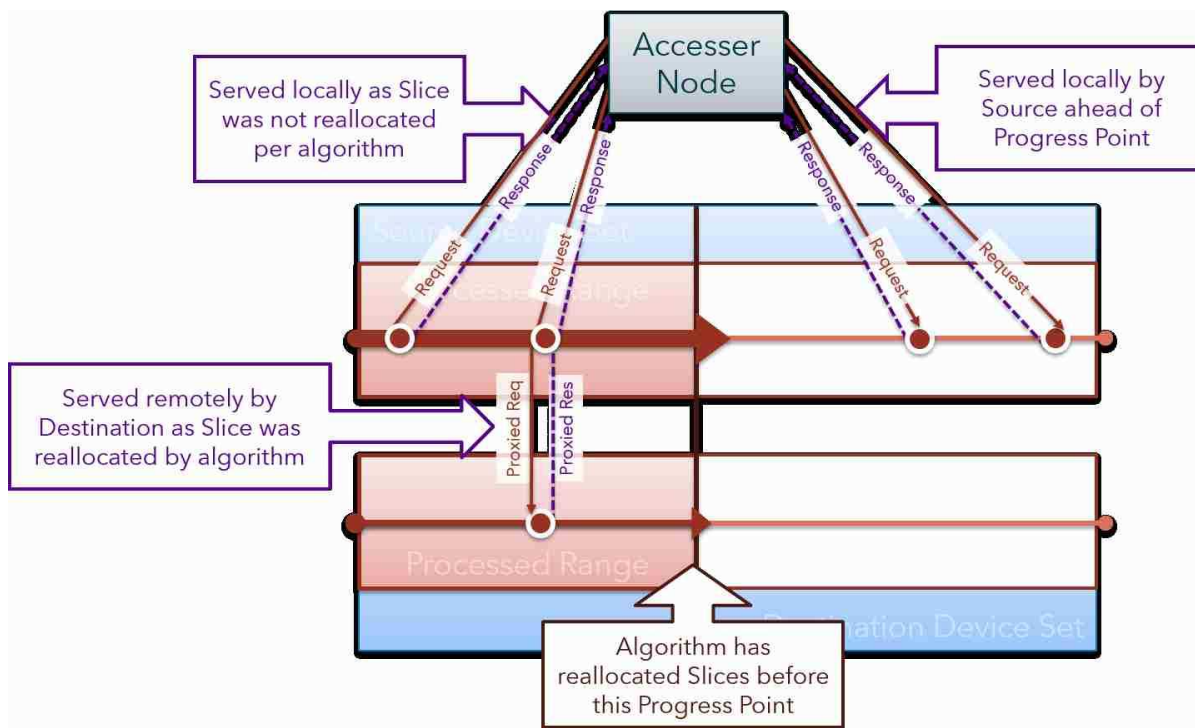
Figure 4. Requests proxy during reallocation

Chapter 5. Use cases

Use cases to consider.

Replace older device sets that contain outdated hardware, smaller device capacity, and so on.

Resize existing device sets to take advantage of newly added capacity.

Storage Pools should be expanded when projections of ingest for current pools suggest they will near their capacity. Reallocation takes time. Plan to start an expansion such that the reallocation can move data off the original Device Sets before they run out of capacity.

This feature replaces the current **Merge Storage Pool** function.

Attention: Two previously merged Storage Pools cannot be expanded.

Attention: A Storage Pool cannot be expanded then merged.

A case for creating new Storage Pools

While Storage Pool Expansion eliminates the requirement to form new Storage Pools to expand the system, it does not entirely eliminate all justifications for creating new Storage Pools.

In situations with differing storage requirements (for example, different physical locations, different levels of physical security, different IDA configurations, different users, and so on), you should create different Storage Pools to meet the different storage needs.

Chapter 6. Feature impact

Performance implications

Storage Pool Expansion impacts the theoretical maximum throughput, OPS, and latency of the system.

Impact to system after expansion

- The theoretical maximum throughput could be increased in proportion to the added network interface capacity of the Slicestor® Nodes in the new Device Set.
- The theoretical maximum OPS could be increased in proportion to the number of added drives of the Slicestor® Nodes in the new Device Set.
- The latency change should be negligible, unless the original latency was significant due to system load. In that case, latency should reduce to a meaningful degree.

Performance Improvement Dependencies

It can vary in practice, depending upon:
- The quantity of Accesser® Nodes to support the quantity of deployed Slicestor® Nodes.
- The type of Slicestor® Nodes that are deployed in the original and new Device Sets.
- Network bandwidth between sites for a dispersed geographic configuration.

Impact to system during reallocation

Latency increases due to reallocation processing and proxying of storage requests.

Throttling limits the impact of reallocating Slices but increases the time that is needed to complete reallocation.

As reallocation proceeds, more Slices are proxied to the new Device Set, which makes more OPS capacity available but the proxied requests use more bandwidth on the source Slicestor® Nodes. It is not an issue unless the source Slicestor® Nodes are operating at near their network usage limits.

Chapter 7. Design decisions and limitations

Expand across physical sites

Because data is reallocated between Slicestor® Appliance devices during system expansion, it is preferred that the new Slicestor® Appliance is physically located in the same sites as the existing sets.

If it is not possible, slices that need to be reallocated might need to cross WAN links between sites. It can result in a slower reallocation rate and a longer reallocation phase. Additionally, the higher latency that typically exists when traversing these links can result in a greater request latency for requests that the source store must proxy.

Note: If this situation arises, contact Customer Support to discuss the proposed system expansion and help to ensure that the new set of devices is provisioned in such a way that the inter-site traffic is minimized.

Health of Slicestor® Nodes

The system should be in as healthy as possible before performing expansion: replace any bad drives, perform any evacuations, and so on.

Quarantined drives stall reallocation and must be remediated, either failed or resumed.

Note: If the need arises to replace a node during reallocation, contact IBM Cloud Object Storage System™ Customer Support.

Compatibility with old clients

All Accesser® Nodes must upgrade to ClevOS version 3.6 before using Storage Pool Expansion, otherwise they cannot access Vaults in newly expanded Storage Pools.

Merge storage pools

Multiple original Storage Pools cannot be collapsed into a single Storage Pool.

Chapter 8. Storage pool expansion using the Manager Web Interface

Storage pool expansion requires an original Storage Pool and a user granted system administrator or super-user rights.

- Each Storage Pool may have only one expansion in progress.
- A Device Set can only expand one Storage Pool.

Prerequisites

These are prerequisites for expanding a Storage Pool.

To expand a Storage Pool, a System Administrator or Super User must assign each Slicestor® Node to a Site in the Manager Web Interfacee if the system is configured in a geographically dispersed manner.

Note: Consult with support about the effect on future reallocations before renaming or physically moving Slicestor® Nodes.

Expand a storage pool

How to expand a storage pool.

Before you begin

Before expanding a storage pool, a system operator must perform the following preconditions:

1. Upgrade all Accesser® Nodes and Slicestor® Nodes to ClevOS version 3.6.0 or later.
2. Approve a number of Slicestor® Nodes equal to a multiple of the least common multiple of all the vault IDA widths for the vaults using the Storage Pool to be expanded. If you are using Concentrated Dispersal you can expand using either the Concentrated Dispersal set size or an integer multiple of the full IDA Width.

Procedure

1. Under the **Configure** tab, click on the Storage Pool to be expanded.
2. Click the **Expand Storage Pool** button.
3. From the **Width** drop-down menu under the **General** bar, select the IDA Width for the expanded Storage Pool.
4. Select the Slicestor® Nodes to be used to expand this pool. These devices will be used to form a new **Device Set** within this Storage Pool.
5. If the nodes to be added are decided, select the Slicestor® Nodes for the new Device Set by one of three options:
   a. Model group using the radio buttons for a particular model
   b. Sites using the checkboxes for one or more Sites
   c. Individual nodes via the checkboxes in the **Devices** section
6. If the nodes to be added are undecided, click **Suggest Devices** to let the Manager application select Slicestor® Nodes on your behalf. Click **Save** to save your changes.
7. Monitor the progress of the reallocation and ensure it continues to completion.
8. Change the Storage Pool Reallocation Rate.

Change the storage pool reallocation rate

Storage Pool expansion can be throttled.

Procedure

1. Click on the Storage Pool from the **Monitor** tab.
2. Click on the **Change** button under the **Data Reallocation in Progress** window. The **Edit Data Reallocation** dialog box displays.
3. For all device sets:
    a. Check the **Enable** check box in the **Bulk Change** section.
    b. Type a transfer limit in MB per second in the **MB/s** field.
4. For one device set:
    a. Check the **Per-Device Rate Limiting** check box for that Device Set.
    b. Type a transfer limit in MB per second in the **MB/s** field.
5. Click **Submit** to accept these changes.

Pausing storage pool reallocation

Storage Pool expansion can be paused.

Procedure

1. Click on the Storage Pool from the **Monitor** tab.
2. Click **Change** on the **Data Reallocation in Progress** window. The **Edit Data Reallocation** dialog box displays.
3. To pause a reallocation:
    a. For all device sets, check the **Pause** check box in the **Bulk Change** section and click **Apply to All Sets**.
    b. For one device set, check the **Pause** check box for that Device Set.
4. Click **Submit** to accept these changes.

   Note: When a Device Set's reallocation is paused, it will appear with a **-Paused-** indicator next to the Device Set's name.

Replacing storage pool sets

How to replace storage pool sets.

Before you begin

Before you replace Storage Pool sets, you must perform the following preconditions:
1. Upgrade all Accesser® Nodes and Slicestor® Nodes to ClevOS 3.10.0 or newer.
2. Approve an amount of Slicestor® Nodes equal to a multiple of the least common multiple of all the vault IDA widths for the vaults that use the Storage Pool that you want to replace. If you are using Concentrated Dispersal, you can replace sets by using either the Concentrated Dispersal set size or an integer multiple of the full IDA Width.

To replace one or more Storage Pool sets:

Procedure

1. In the **Configure** tab, click the **Storage Pool** you want to replace.
2. In the Slicestor Devices section, click **Change Sets and Devices**.

3. Click **Configure Set Replacement**.
4. In the **Sets to Replace** section, select the **Storage Pool** sets you want to replace.
5. Select the Slicestor® Nodes you want to use to replace the Storage Pool sets. The devices are used to form a new Device Set within this Storage Pool.
6. Select the Slicestor® Nodes you want to add for the new Device Set in one of the following ways:
   - Select the check boxes for choosing Slicestor models with similar drive counts
   - Check the check boxes for one or more Sites to add nodes by Site.
   - Check the check boxes in the Devices section to add individual nodes.
   - Click **Suggest Devices** to allow the Manager application to select Slicestor® Nodes on your behalf.
7. Click **Continue**. A confirmation page appears.
8. Click **Save**.
9. Monitor the progress of the reallocation and ensure it continues to completion.

Resizing a storage pool

How to resize a storage pool.

About this task

You can resize an existing partially populated storage pool to take advantage of extra capacity.

Note: To resize a storage pool, the storage pool must contain multiple sets.

Procedure
1. Install more drives in each Slicestor® Node in the storage pool.
2. When most of the devices have the additional capacity, click the **Configure** tab.
3. In the navigation tree, click the storage pool that you want to resize. The Configure Storage Pool page displays a notification that the system identified more capacity.
4. When you finish installing extra drives, click **Done Adding Capacity** in the notification. The notification displays the estimated capacity before resize and the project capacity after resize.
5. Confirm the capacity values and click **Approve & Start Data Reallocation**. The resize process begins.
6. Monitor the resize process on the Monitor Storage Pool page.

Removing storage pool sets

To remove one or more Storage Pool sets, follow these instructions.

Before you begin

Before you can remove Storage Pool sets, you must upgrade all Accesser® Nodes and Slicestor® Nodes to release 3.10.0 or newer.

Procedure
1. In the **Configure** tab, click the **Storage Pool** that you want to replace.
2. In the Slicestor Devices section, click **Change Sets and Devices**.
3. Click **Configure Set Removal**.
4. Select the **Storage Pool** sets you want to remove and click **Continue**. A confirmation page appears.
5. To remove the Storage Pool sets, click **Save**.
6. Monitor the progress of the reallocation and ensure it continues to completion.

Chapter 9. Storage pool expansion using the Manager API

There are REST API methods related to Storage Pool expansion.

These APIs are defined in the **Manager Help**.

There are two relevant API methods:

Expanding storage pool

Expanding an existing storage pool.

Request

Security
Security for expanding a pool.

Table 2. Roles capable of executing the Expand Storage Pool (expandStoragePool) API method

Any	Super User	System Admin	Security Admin	Operator
	superUser	systemAdmin		

HTTP method

```
POST /manager/api/{apiResponseType}/1.0/expandStoragePool.adm HTTP/1.1
    Host:{manager.dsnet}
    id={spId}&width={idaWidth}
            &deviceIds={node1}
            &deviceIds={node2}
            &deviceIds={node3}
            &deviceIds={node4}
            &deviceIds={node5}
```

Curl method

```
curl
    -X POST
    -u admin:password
    -k "https://{manager.dsnet}/manager/api/{apiResponseType}/1.0/expandStoragePool.adm"
    -d "id={spId}&width={idaWidth}
                &deviceIds={node1}
                &deviceIds={node2}
                &deviceIds={node3}
                &deviceIds={node4}
                &deviceIds={node5}"
```

Parameters
These are the parameters used to expand a storage pool.

Table 3. Request Parameters for Expand Storage Pool (expandStoragePool) API method

Parameter	Type	Usage	Default	Description
id	long	Required		ID of the Storage Pool to expand
width	integer	Required		The width of the new Storage Pool

Parameter	Type	Usage	Default	Description
`deviceIds`	`List`	`Required`		Device IDs to expand the Storage Pool by

Response

Returns success or failure status.

JSON response example

Default JSON success or failure response shown in *(Default JSON Success Response)* and *(Default JSON Failure Response)*.

Parameters

No response parameters other than the default response parameters are shown in "Parameters" on page 27.

Editing data reallocation

Request

Security

Table 4. Roles capable of executing the Edit Data Reallocation (`editDataReallocation`) API method

Any	Super User	System Admin	Security Admin	Operator
	`superUser`	`systemAdmin`		

HTTP method

```
POST /manager/api/{apiResponseType}/1.0/editDataReallocation.adm HTTP/1.1
    Host:{manager.dsnet}
    id={spId}&rateLimitEnabled={true|false}
            &rateLimitMB={limitInMB}
            &reallocationPaused={true|false}
            &rateLimitEnabledMap[{spId2}]={true|false}
            &rateLimitEnabledMap[{spId3}]={true|false}
            &rateLimitMBMap[{spId2}]={limitInMB}
            &rateLimitMBMap[{spId3}]={limitInMB}
            &reallocationPausedMap[{spId2}]=false
            &reallocationPausedMap[{spId3}]=false
            &useGlobalPropertyMap[{spId1}]={true|false}
```

Curl method

```
curl
    -X POST
    -u admin:test
    -k "https://{manager.dsnet}/manager/api/{apiResponseType}/1.0/editDataReallocation.adm"
    -d "id={spId}&rateLimitEnabled={true|false}
            &rateLimitMB={limitInMB}
            &reallocationPaused={true|false}
            &rateLimitEnabledMap[{spId2}]={true|false}
            &rateLimitEnabledMap[{spId3}]={true|false}
            &rateLimitMBMap[{spId2}]={limitInMB}
            &rateLimitMBMap[{spId3}]={limitInMB}
            &reallocationPausedMap[{spId2}]=false
            &reallocationPausedMap[{spId3}]=false
            &useGlobalPropertyMap[{spId1}]={true|false}"
```

Parameters

Table 5. Request Parameters for Edit Data Reallocation `editDataReallocation` API method

Parameter	Type	Usage	Default	Description
Parameter	Optional		Type	Description
id	Long	Required		Storage Pool Id
rateLimitEnabled	Boolean	Required		Used when setting if rate limit is enabled on all storage sets marked as true in the useGlobalPropertyMap
rateLimitMB	Integer	Optional		Used when setting what rate limit (in MB) should be used for all storage sets marked as true in the useGlobalPropertyMap
reallocationPaused	Boolean	Required		Should reallocation be paused on all storage sets marked as true in the useGlobalPropertyMap
rateLimitEnabledMap	Map	Required		A map of Storage Pool ids to a boolean for if the rate limit should be enabled
rateLimitMBMap	Map	Optional		A map of Storage Pool IDs to a number that the rate limit should be set as in MB
reallocationPausedMap	Map	Required		A map of Storage Pool IDs to a boolean for if the reallocation should be paused
useGlobalPropertyMap	Map	Required		A map of Storage Pool ids mapped to a boolean value whether or not they should get overwritten with the top level rateLimitEnabled, rateLimitMB and isReallocationPaused properties

Response

Returns success or failure status.

JSON Response Example

Default JSON success or failure response shown in *(Default JSON Success Response)* and *(Default JSON Failure Response)*.

Parameters

None other than the default response parameters are shown in "Parameters."

Replace storage pool sets

Replaces one or more storage pool sets with a new set of devices.

Request

Security

Table 6. Roles capable of executing the Replace Storage Pool Sets (`replaceStoragePoolSets`) API method

Any	Super User	System Admin	Security Admin	Operator
	superUser	systemAdmin		

HTTP method

```
POST /manager/api/{apiResponseType}/1.0/replaceStoragePoolSets.adm HTTP/1.1
    Host:{manager.dsnet}
    id={spId}&width={idaWidth}
            &deviceIds={node1}
            &deviceIds={node2}
            &deviceIds={node3}
            &deviceIds={node4}
            &deviceIds={node5}
            &replacedSetIds={set1}
            &replacedSetIds={set2}
```

Curl method

```
curl
    -X POST
    -u admin:password
    -k "https://{manager.dsnet}/manager/api/{apiResponseType}/1.0/replaceStoragePoolSets.adm"
    -d "id={spId}&width={idaWidth}
                &deviceIds={node1}
                &deviceIds={node2}
                &deviceIds={node3}
                &deviceIds={node4}
                &deviceIds={node5}
                &replacedSetIds={set1}
                &replacedSetIds={set2}"
```

Parameters

Table 7. Request Parameters for Replace Storage Pool Sets (`replaceStoragePoolSets`) API method

Parameter	Type	Usage	Default	Description
`id`	Long	Required		ID of the Storage Pool to replace sets.
`replacedSetIds`	List	Required		IDs of the Storage Pool sets to replace.
`width`	Integer	Required		The width of the new replacement Storage Pool set.
`partial`	Boolean	Optional	false	Set to true if planning to create a replacement storage pool set with some devices temporarily missing. When true, the number of device IDs that are passed in (by using the `deviceIds` parameter) does not need to equal the width of the new storage pool set.
`storageEngine`	String			The storage engine to use for this storage pool. One of "packed" or "file".
`deviceIds`	List	Required		Device IDs to use in the replacement Storage Pool set.

Response

Returns success or failure status.

JSON response example

Default JSON success or failure response shown in *(Default JSON Success Response)* and *(Default JSON Failure Response)*.

Parameters

None other than the default response parameters shown in . Common Response Parameters.

Resize storage pool sets

Resize the Storage Pool sets to take advantage of extra capacity.

Request

Security

Table 8. Roles capable of executing the Resize Storage Pool Sets (`resizeStoragePoolSets`) API method

Any	Super User	System Admin	Security Admin	Operator
	superUser	systemAdmin		

HTTP method

```
POST /manager/api/{apiResponseType}/1.0/resizeStoragePoolSet.adm HTTP/1.1 Host:{manager.dsnet}
id={spId}
```

Curl method

```
curl -X POST -u admin:password -k "https://{manager.dsnet}/manager/api/{apiResponseType}/1.0/
resizeStoragePoolSet.adm" -d "id={spId}"
```

Table 9. Request Parameters for Resize Storage Pool Set (`resizeStoragePoolSet`) API method

Parameter	Type	Usage	Default	Description
id	Long	Required		ID of the Storage Pool to resize

Parameters

Table 10. Request Parameters for Resize Storage Pool Set (`resizeStoragePoolSet`) API method

Parameter	Type	Usage	Default	Description
id	Long	Required		ID of the Storage Pool to resize.

Response

Returns success or failure status.

JSON response example

Default JSON success or failure response shown in *(Default JSON Success Response)* and *(Default JSON Failure Response)*.

Parameters

None other than the default response parameters shown in *(Common Response Parameters)*.

Chapter 10. Operations

Monitoring reallocation progress

Procedure

1. Click the **Storage Pool** that is undergoing a reallocation, from the **Monitor** tab of the Manager Web Interface. A page that shows the overall progress of the reallocation appears.
2. Click **View Reallocation Progress** to show per Slicestor® Node progress.

Monitoring reallocation performance

The Manager Web Interface provides several metrics of reallocation performance at an aggregate per-Device Set pool and at a per-device level. Graphs of the rate of data transfer and the rate of Slices per second are provided.

Aggregate storage pool reallocation rates

Outgoing reallocation rates on source device sets

View the aggregate rate of all reallocation across an entire Storage Pool.

Procedure

1. Click the **Monitor** tab in the Manager Web Interface.
2. Click **Storage Pool** in the left navigation menu.

   Note: The reallocation rate in terms of bytes and Slices per second. It can be displayed for the entire Storage Pool or per Device Set.
3. Select a Device Set from the **Graph Selection** to monitor the rate for each Slicestor® Node independently.

Incoming reallocation rates on destination device sets

The total inbound rate of reallocation can be seen for Slicestor® Nodes in the destination Device Set.

Note: This rate represents an aggregate rate of inbound traffic that is received from all source Slicestor® Nodes that share responsibility for the same Slices as this device.

Troubleshooting

How to diagnose and recover from problems that can occur during Storage Pool expansion.

Manager events and incidents

The first indication that something is wrong generally appears through the Manager Web Interface.

These events all apply at the Storage Pool level.

Reallocation-related events

Reallocation started

> This event indicates that an expansion is started. The Storage Pool cannot be expanded again until the Reallocation completed event occurs.

Reallocation completed

This event indicates that reallocation is completed for every source Slicestor® Node in the Storage Pool and that the reallocation process is complete.

Reallocation-related incidents

Reallocation has not made progress for 6 hours.

This event indicates that the Slicestor® Node reported progress does not change in more than 6 hours. It can be caused by the following items:

- Quarantined drives in either a source or destination Slicestor® Node.

  Fail or resume the drive to continue. This event will clear automatically after it is done.

- A single slow Slicestor® Node. When this event occurs, resolve and correct the problem as soon as possible.

  It can be due to a lack of network connectivity between the source and destination Slicestor® Node.

Note: If no drives are quarantined and progress does not resume, contact IBM Customer Support.

Chapter 11. Additional information

1. How long do reallocations take?

    As all reallocations are throttled, the rate of the reallocation can be computed:

    ```
    T = (Y / (X+Y)) * U * C * R
    ```

Table 11. Variables and their Values for Reallocation

Variable	Description
X	Aggregate Capacity of Existing Pool (all sets)
Y	Capacity of New Set
U	Utilization Percentage of Storage Pool before Expansion
C	Largest Slicestor Node's Capacity
R	Reallocation Throttle Rate (MBps)
T	Reallocation Time

It implies that you:

a. Multiply the capacity per drive by the number of drives per Slicestor® Node.

b. Multiply this result (TB per node) by the amount of Slicestor® Nodes.

c. Divide this result (TB per pool) by 2 to get the amount of storage to be reallocated.

d. Multiply this result (TB reallocated) by the average percentage of utilization of each Slicestor® Node.

e. Multiply this result (TB used) by capacity size of the largest Slicestor® Node.

f. Multiply this result (TB to reallocate) by 1 million.

g. Divide this result (MB to reallocate) by the Throttle Level.

h. Multiply this result (Amount of time to reallocate in seconds) by the product of seconds per minute then minutes per hour then hours per day.

i. It gives you the total days to reallocate.

Example 2. Reallocation Time

Add a Device Set with 24 Slicestor 2400 appliances with 40 ├ 4 TB drives to a Device Set composed of the same number, type and configuration of Slicestor appliances throttled to 500 MBps.

```
T = (Y / (X+Y)) * U * C / R
T =           ((4 TB * 40 * 24) / ((4 TB * 40 * 24) + (4 TB * 40 * 24))) * 80% * (4 TB *
40) / 20        MBps

T = ((3,840 TB / 7,680 TB) *  80% *  160 TB) / (20 MBps)
T = (0.5 * 80% *  160 TB) / (20 MBps)
T = (64 TB) / (20 MBps)
T = (3,200,000) / (60 *  60 * 24)
T = (3,200,000) / (86,400)
T = 37.04
```

2. Is there a maximum number of expansions supported?

    Storage Pools have no limit on how many times they can be expanded. Limit Vault Mirrors to five Device Sets for performance considerations.

3. Do only certain types of data move?

   All types of data are treated identically for purposes of reallocation and proxying.

4. Are certain storage types supported?

   The source or destination Slicestor® Nodes can use Packed Slice Storage or File Slice Storage. When a Storage Pool is expanded, the Slicestor® Nodes in the new Device Set use the same storage type as the other Slicestor® Nodes.

5. Are only certain APIs supported?

   The SOH, CSO, and OpenStack APIs are supported.

Notices

This information was developed for products and services offered in the US. This material might be available from IBM® in other languages. However, you may be required to own a copy of the product or product version in that language in order to access it.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not grant you any license to these patents. You can send license inquiries, in writing, to:

IBM Director of Licensing
IBM Corporation
North Castle Drive
Armonk, NY 10504-1785
U.S.A.

For license inquiries regarding double-byte character set (DBCS) information, contact the IBM Intellectual Property Department in your country or send inquiries, in writing, to:

Intellectual Property Licensing
Legal and Intellectual Property Law
IBM Japan, Ltd.
19-21, Nihonbashi-Hakozakicho, Chuo-ku
Tokyo 103-8510, Japan

INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some jurisdictions do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM websites are provided for convenience only and do not in any manner serve as an endorsement of those websites. The materials at those websites are not part of the materials for this IBM product and use of those websites is at your own risk.

IBM may use or distribute any of the information you provide in any way it believes appropriate without incurring any obligation to you.

Licensees of this program who wish to have information about it for the purpose of enabling: (i) the exchange of information between independently created programs and other programs (including this one) and (ii) the mutual use of the information which has been exchanged, should contact:

IBM Director of Licensing
IBM Corporation
North Castle Drive, MD-NC119
Armonk, NY 10504-1785
US

Such information may be available, subject to appropriate terms and conditions, including in some cases, payment of a fee.

The licensed program described in this document and all licensed material available for it are provided by IBM under terms of the IBM Customer Agreement, IBM International Program License Agreement or any equivalent agreement between us.

The performance data discussed herein is presented as derived under specific operating conditions. Actual results may vary.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

Statements regarding IBM's future direction or intent are subject to change or withdrawal without notice, and represent goals and objectives only.

All IBM prices shown are IBM's suggested retail prices, are current and are subject to change without notice. Dealer prices may vary.

This information is for planning purposes only. The information herein is subject to change before the products described become available.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs. The sample programs are provided "AS IS", without warranty of any kind. IBM shall not be liable for any damages arising out of your use of the sample programs.

If you are viewing this information softcopy, the photographs and color illustrations may not appear.

Trademarks

IBM, the IBM logo, and ibm.com® are trademarks or registered trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the web at Copyright and trademark information at www.ibm.com/legal/copytrade.shtml.

Accesser®, Cleversafe®, ClevOS™, Dispersed Storage®, dsNet®, IBM Cloud Object Storage Accesser®, IBM Cloud Object Storage Dedicated™, IBM Cloud Object Storage Insight™, IBM Cloud Object Storage Manager™, IBM Cloud Object Storage Slicestor®, IBM Cloud Object Storage Standard™, IBM Cloud Object Storage System™, IBM Cloud Object Storage Vault™, SecureSlice™, and Slicestor® are trademarks or registered trademarks of Cleversafe, an IBM Company and/or International Business Machines Corp.

Other product and service names might be trademarks of IBM or other companies.

Homologation statement

This product may not be certified in your country for connection by any means whatsoever to interfaces of public telecommunications networks. Further certification may be required by law prior to making any such connection. Contact an IBM representative or reseller for any questions.

IBM ®

Printed in USA