WHITE PAPER

PURITY CLOUDSNAP

SETUP AND BEST PRACTICES GUIDE

PURESTORAGE®

TABLE OF CONTENTS

INTRODUCTION

Pure portable snapshots provide simple, built-in, local and cloud protection for Pure FlashArrays. Purity Snapshots enable free movement of space-efficient copies between FlashArrays, to FlashBlade, to 3rd party NFS servers, and to the cloud. Pure's portable snapshot technology encapsulates metadata along with data into the snapshot, making the snapshot portable, so it can be offloaded from a Pure FlashArray to the cloud in a format that is recoverable to any FlashArray. In Purity//FA 5.2, CloudSnap extends the mobility of snapshots: Pure snapshots can now be sent directly to the cloud from Pure FlashArray appliances.

Pure's portable snapshot technology was first introduced via the Snap-to-NFS feature in Purity//FA 5.1, which enables FlashArray to offload snapshots to Pure FlashBlade™ and to other NFS servers. Snapshots to FlashBlade enable rapid restore on-premises while consolidating siloed workloads. (Learn more about Snap-to-NFS.) This technology has now been extended to enable data protection in the cloud. The first replication target that has been added is the AWS S3 object-based cloud storage service.

CloudSnap™ enables FlashArray to offload snapshots directly to an S3 bucket, without the need for additional backup software or a cloud gateway. Think of it as a built-in "self-backup to cloud" feature for FlashArray.

Once the S3 bucket is configured as a replication target in FlashArray, it appears like any other destination, and users can simply create a protection group, select the data they want to replicate, add the S3 bucket to the protection group, and create a schedule to determine the replication frequency and the retention period. Once a protection group has been created and configured, FlashArray starts replicating snapshots to the S3 bucket. Restoring data is also simple: snapshots in the S3 bucket can be browsed via the FlashArray GUI and recovered with a few simple clicks from the cloud back to the source FlashArray, or to a different FlashArray.

CloudSnap can be managed natively on Pure FlashArrays via the GUI or the CLI. It is also integrated with Pure1®, so users can monitor snapshots in the cloud via the Pure1 Snapshot Catalog. In addition, there's a robust and open REST API which can be used by customers and third party data management software to move incremental snapshots from FlashArray to the cloud.

The Pure Storage website contains more information on CloudSnap.

CLOUDSNAP BENEFITS

CloudSnap is a self-backup technology built into FlashArray. It does not require the purchase of additional backup software or hardware, nor is there a need to learn and use an additional management interface. CloudSnap is natively managed via Pure FlashArray's GUI, CLI, and REST interfaces and is integrated with the Pure1 Snapshot Catalog.

Since FlashArray connects to AWS via https, data is encrypted in transit and stored in an encrypted format in the S3 bucket using server side encryption.

Since CloudSnap was built from scratch for FlashArray, it is deeply integrated with the Purity Operating Environment, resulting in highly efficient operation. A few examples of the efficiency of CloudSnap:

- CloudSnap preserves data compression on the wire, and in the S3 bucket, saving network bandwidth and increasing storage space efficiency.
- CloudSnap preserves data reduction across snapshots of a volume. After offloading the initial baseline snapshot of a volume, it only sends delta changes for subsequent snaps of the same volume. The snapshot differencing engine runs within the Purity Operating Environment in FlashArray and uses a local copy of the previous snapshot to compute the delta changes. Therefore, there is no back and forth network traffic between FlashArray and the cloud to compute deltas between snapshots, further reducing network congestion and data access costs in the cloud.
- CloudSnap knows which data blocks already exist on FlashArray, so during restores it only pulls back missing data blocks to rebuild the complete snapshot on FlashArray. In addition, CloudSnap uses dedupe-preserving restores, so when data is restored from the offload target to FlashArray, it is deduped to save space on FlashArray.

The highly efficient operation of CloudSnap provides the following benefits:

- Less space is consumed in the S3 bucket
- Network utilization is minimized
- Backup windows are much smaller
- Data retrieval costs from the S3 bucket are lower

CORE COMPONENTS

FlashArray, Purity, and Run

CloudSnap is available starting with Purity version 5.2. Since CloudSnap relies on an offload engine that operates in Purity Run, FlashArray models that support Purity Run are required.

The Pure Storage support website contains more information on Purity Run.

Network

FlashArray ports configured for CloudSnap must have network connectivity to AWS. An extra IP address is required to offload data, even if the ports already have IP addresses assigned.

AWS

AWS Account – Users should have an AWS account in which they can create an S3 bucket to store snapshots offloaded by CloudSnap.

S3 Bucket – At least one dedicated AWS S3 bucket must be created, and FlashArray must be given full access to this bucket. This bucket is used by CloudSnap to store snapshots offloaded from FlashArray. Customers must enable encryption on the S3 bucket, so that the offloaded snapshots will be encrypted when stored in the S3 bucket.

CLOUDSNAP ADMINISTRATION

This section covers basic CloudSnap administration. Please refer to the FlashArray User's Guide for a complete list of all CloudSnap capabilities and commands.

Creating an S3 Bucket in the Customer's AWS Account

Log in to the AWS management console, and go to S3. From the AWS S3 console dashboard, choose **Create Bucket**, as shown below:

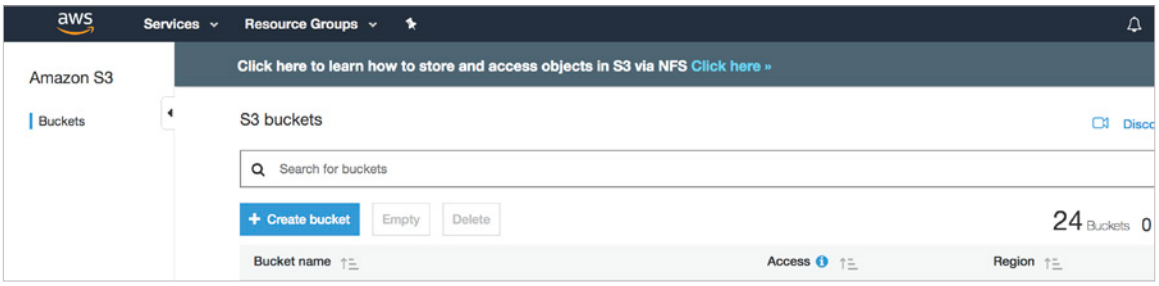

Figure 1. AWS management console

On the first **Create Bucket** screen, enter a name for the S3 bucket and an AWS region in which to create the bucket. In most cases, it is best to select the region which is geographically closest. The bucket name must be globally unique. For more information on S3 bucket naming requirements, please read the AWS documentation.

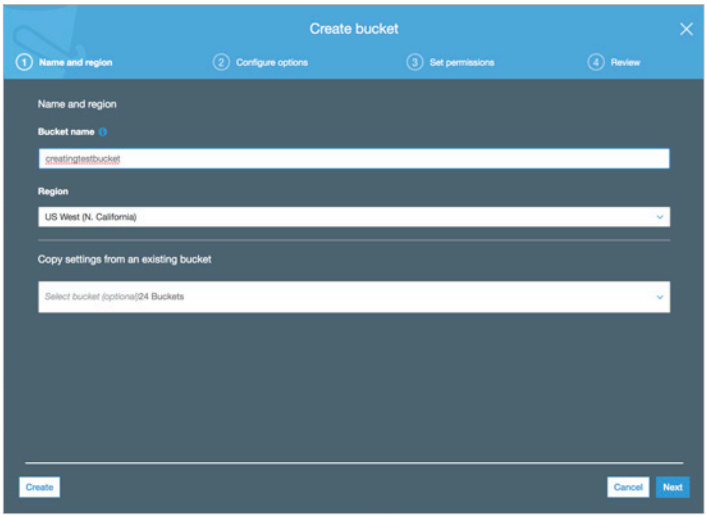

Figure 2. AWS Create bucket screen

Click on **Next** in the following few screens, leaving all the other options at default, unless there's a specific reason to change the default value of an option. The final screen will show a summary of the selected options. Click on **Create Bucket** to create the bucket.

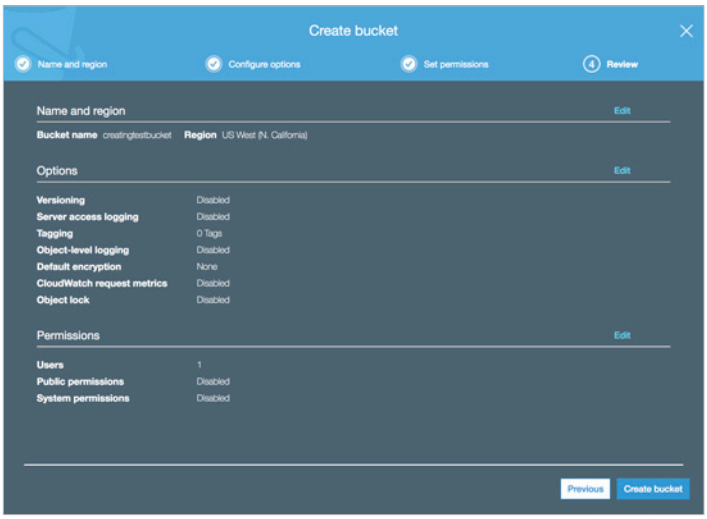

Figure 3. AWS Create Bucket summary

On the S3 dashboard, verify that the new bucket has been added to the list of S3 buckets.

Enabling Encryption on the Newly Created S3 Bucket

Encryption must be enabled on the S3 bucket in order for CloudSnap to work.

To enable encryption, from the list of buckets on the S3 dashboard, click on the name of the newly created bucket, and select the **Properties** tab. The following screen will appear.

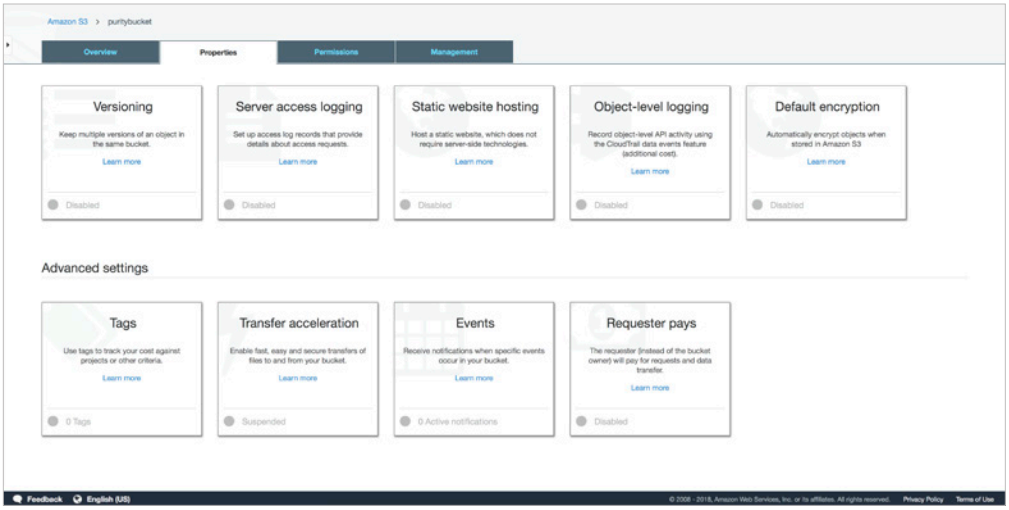

Figure 4. AWS Properties tab

Click on the **Default encryption** box, and the following encryption options will be displayed.

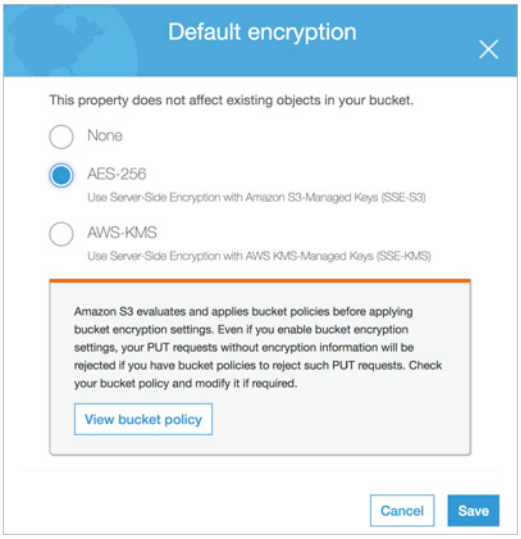

Figure 5. Default encryption screen

Currently, there's no additional charge for the default encryption using S3 managed keys. Please refer to the AWS website for their Key Management Services.

Select your preferred encryption method and click on **Save**. Verify that the Default encryption box displays your selected encryption method.

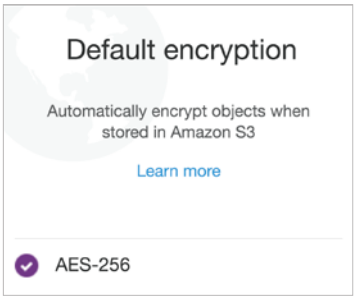

Figure 6. Default encryption box with selected encryption

Other Options on the S3 Bucket

A best practice is to leave all other bucket options at default (with the exception of encryption, which should be enabled as described above) unless there's a specific reason to change the default value of an option. The following options should not be enabled under any circumstances:

- Please do not add lifecycle rules. CloudSnap will not work properly if lifecycle rules have been added.
- **Important:** Do not make the bucket public, in order to keep your data secure.

Creating an IAM Policy to Allow Full Access to the S3 Bucket

This step goes through the process of creating an IAM policy to grant a user full access to the S3 bucket. Since there is no pre-configured **AWS Managed** policy that allows full access to one specific bucket in S3, the following 2 options are available to create the IAM policy:

OPTION 1, THE SIMPLE METHOD – USING AN EXISTING AWS MANAGED POLICY

The easier option is to use AWS's pre-configured policy called **AmazonS3FullAccess**. This AWS policy grants the IAM user (Pure FlashArray) full access to all S3 buckets in the customer's AWS account.

Users who wish to follow this method should proceed to the step "Create an IAM User, and Attach a Policy to the User".

OPTION 2, THE MORE RESTRICTIVE METHOD – CREATING A CUSTOMER MANAGED POLICY

This option is for users who want to create a **Customer managed** policy which would allow the IAM user (Pure FlashArray) full access to only the specific S3 bucket that will be used by CloudSnap to store offloaded data from FlashArray.

Please follow the steps below to create a **Customer managed** policy that allows the IAM user full access to one specific S3 bucket:

From the Identity and Access Management console, go to **Policies** and click on **Create policy**.

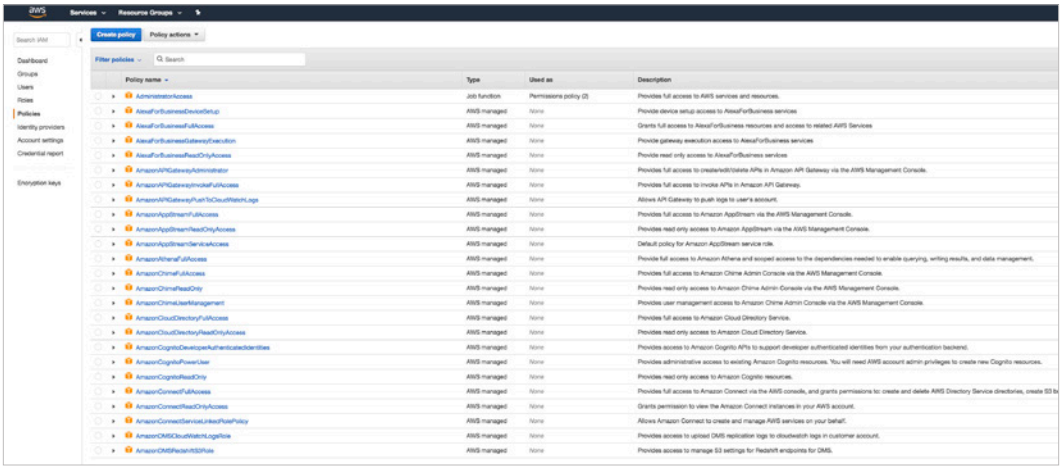

Figure 7. AWS Policies screen

The following screen will appear. Select the **JSON** tab, and either copy and paste the sample JSON code below, or write your own JSON script to allow an IAM user full access to the newly created bucket in S3.

The bucket name in the following sample policy is **snaptos3**. Replace this name with the name of the bucket created in the previous step.

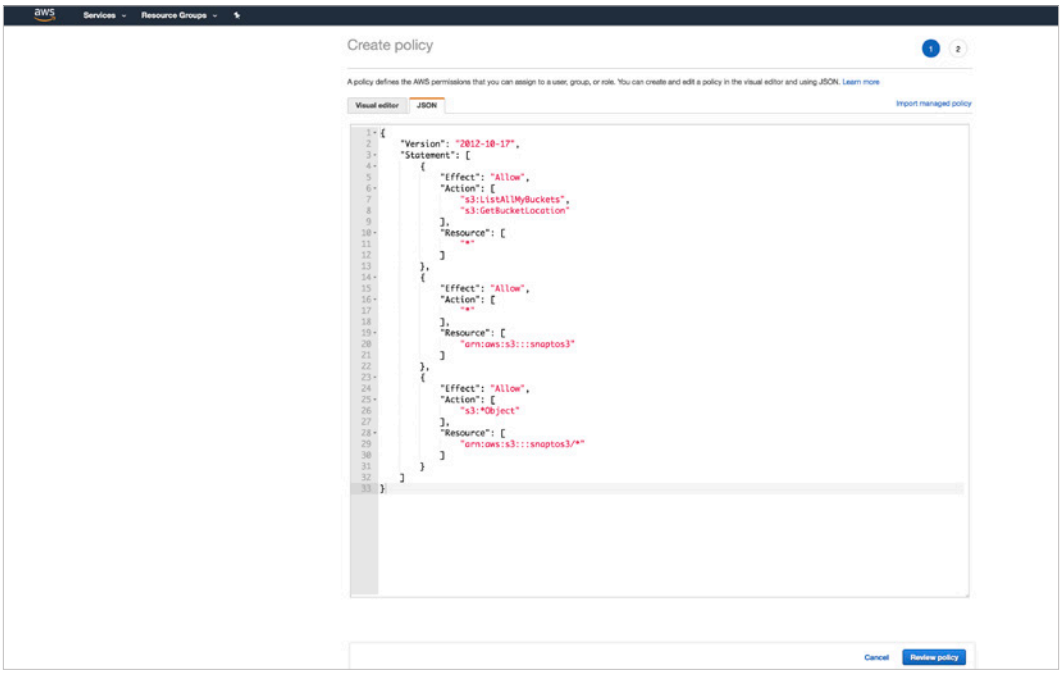

Figure 8. JSON script for full user access

Following is the sample JSON code in text format, for copy & paste:

```
{
"Version": "2012-10-17",
"Statement":
      [
           {
                                "Effect": "Allow",
                                "Action": ["s3:ListAllMyBuckets",
"s3:GetBucketLocation"],
                                "Resource": ["*"]
               },
               {
                                "Effect": "Allow",
                                "Action": ["*"],
                                "Resource": ["arn:aws:s3:::snaptos3"]
               },
               {
                                "Effect": "Allow",
                                "Action": ["s3:*Object"],
                                "Resource": ["arn:aws:s3:::snaptos3/*"]
               }
]
}
```

After pasting the sample JSON code, click on **Review policy**, and the following screen will appear.

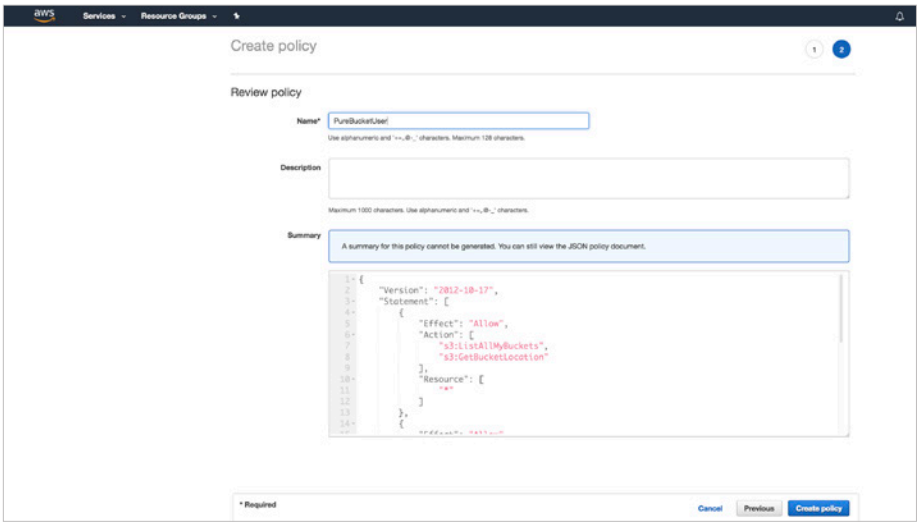

Figure 9. Review policy screen

Enter a name for the user-created policy, and click on **Create Policy**. After the policy is created, verify that it shows up in the policy list as a **Customer managed** policy.

Creating an IAM User and Attaching the IAM Policy to the User

Log in to the AWS management console and navigate to IAM. Click on **Add User** to create a new IAM user.

Figure 10. Add user screen

Select a name for the new user and select **Programmatic Access** for the **Access Type**. Then click on **Next: Permissions**.

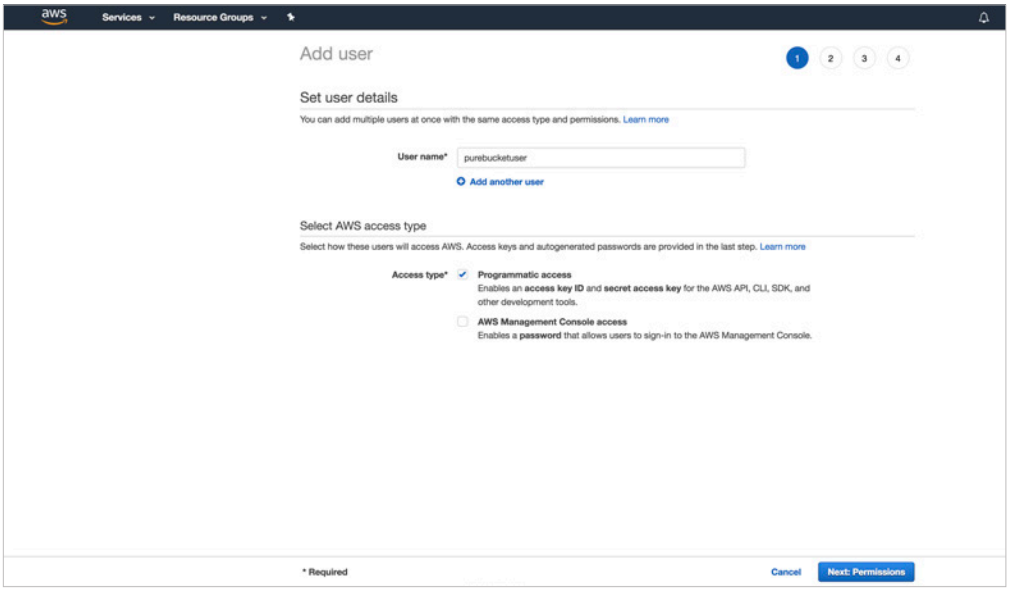

Figure 11. Add user details

The following screen will appear. This screen is where permissions are granted to the newly created IAM user. A permissions policy can either be attached directly to a user, or it can be attached to a group, and the user can be added to the group.

In this example, we will attach a policy directly to the user. Click on the **Attach existing policies directly** box in the top menu. A list of existing policies will appear. Search for and select the desired policy from the list.

If you created a **Customer Managed** policy in Step 4, search for and select that policy. If you want to use the pre-configured **AWS Managed** policy, select the **AmazonS3FullAccess** policy, as shown in the figure on the next page.

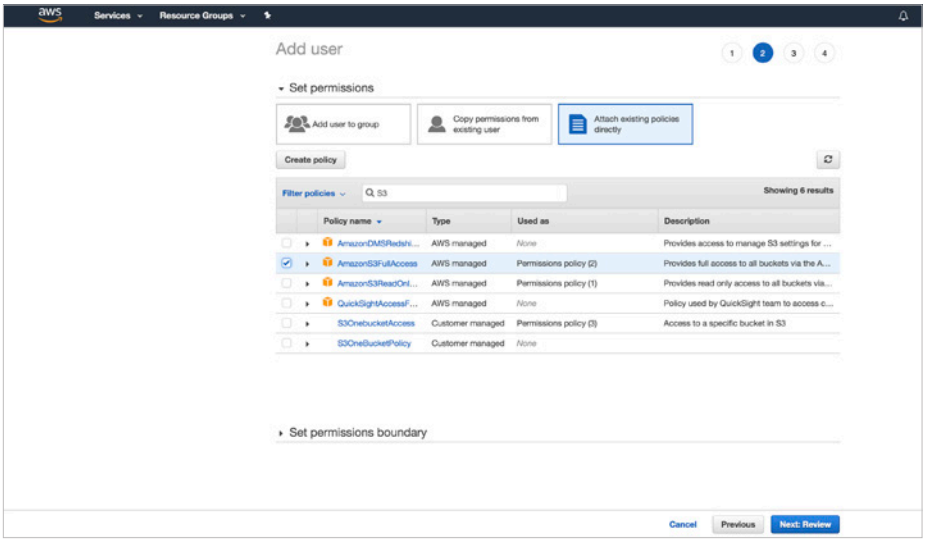

Figure 12. User permissions screen

Click on **Next: Review**. The following screen will appear. Verify that the information is correct and click on **Create user**.

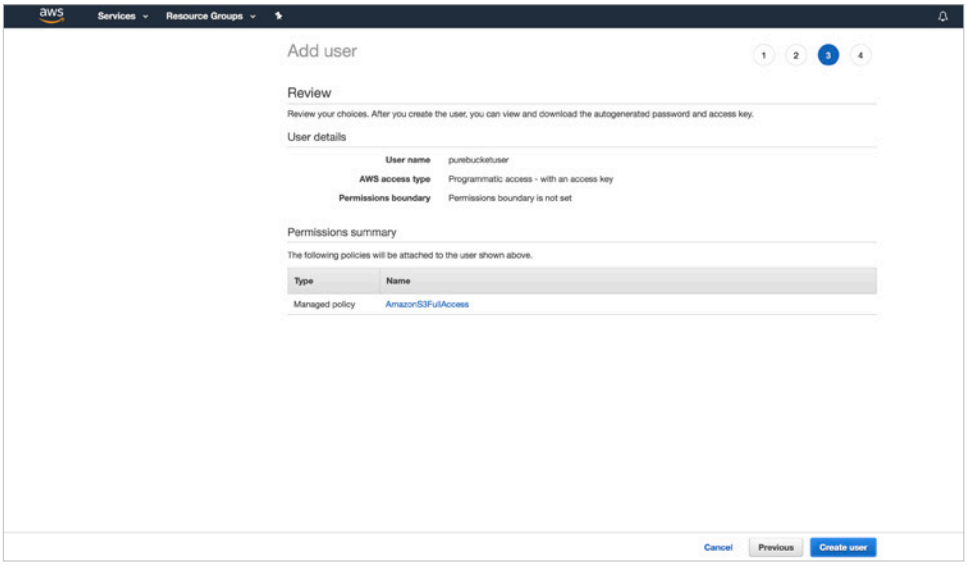

Figure 13. User permissions review screen

The following screen will appear. Be sure to either click on the **Download .csv** file or copy the **Access key ID** and the **Secret access key** at this point, because once this screen is closed, there is no way to get the security credentials for the IAM user. Save the security credentials and use them when connecting FlashArray to the S3 bucket.

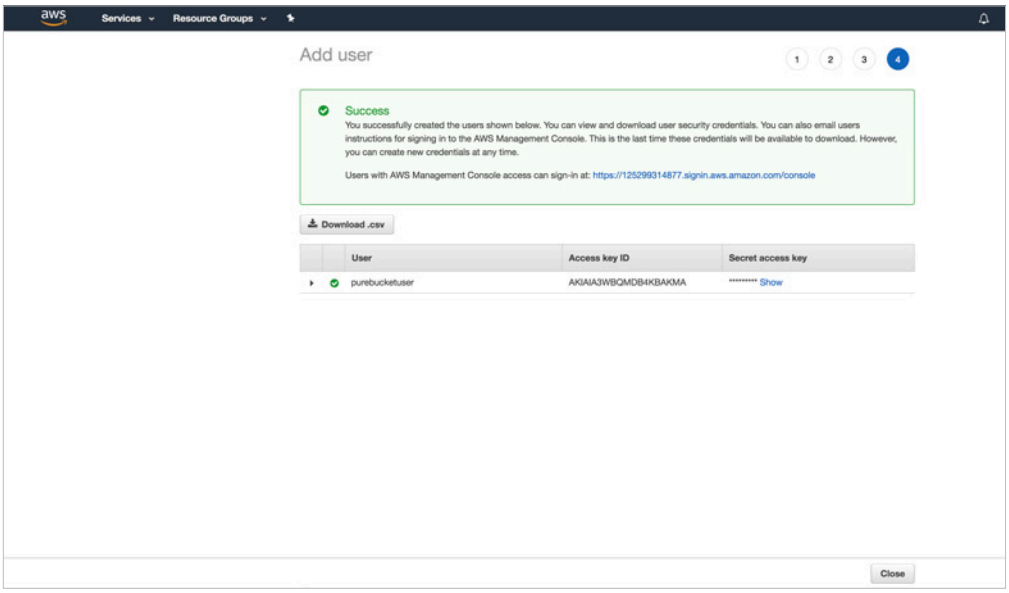

Figure 14. Add User Success screen

Connecting FlashArray to the AWS S3 Bucket

The following screenshots show how to connect FlashArray to the S3 bucket via the FlashArray GUI.

After logging into the FlashArray GUI, go to **Storage > Array**, and under **Offload Targets**, click on the **+** sign.

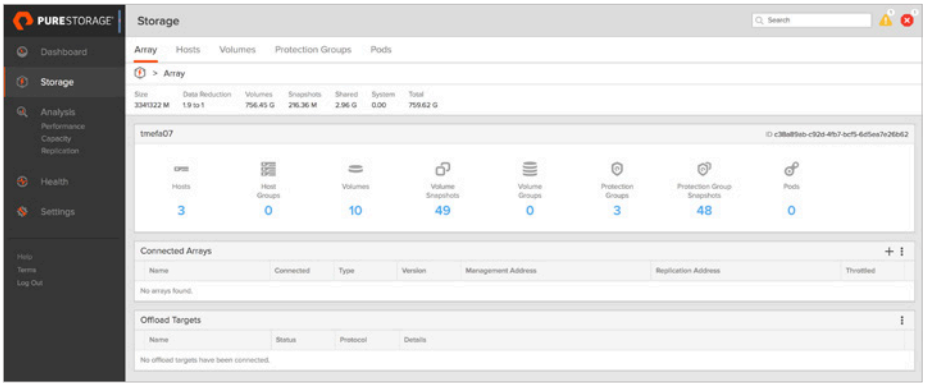

Figure 15. FlashArray Storage > Array screen

Select **Connect to S3 Offload Target** from the options menu, as shown on the next page.

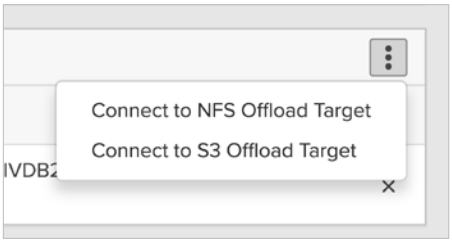

Figure 16. Target options

The following window will appear. Select a name for the S3 target on FlashArray. Enter the Access Key ID and the Secret Access Key for the AWS IAM user, and enter the name of the S3 bucket that was created earlier. This is the bucket that will be used to store snapshots offloaded from FlashArray by CloudSnap. If connecting to the bucket for the first time, select the **Initialize** option.

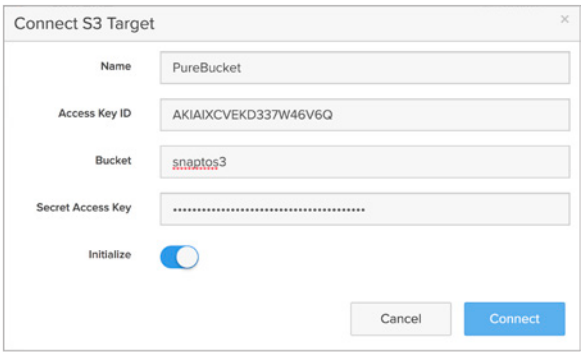

Figure 17. Connect S3 target screen

Once FlashArray successfully connects to the S3 bucket, the new S3 target will be listed under **Offload Targets**, as shown below.

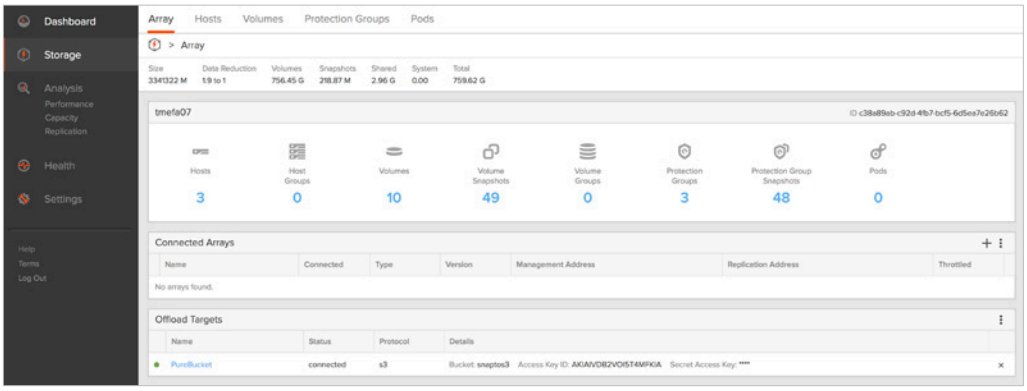

Figure 18. Offload targets screen

When the S3 bucket is connected to FlashArray, the next step is to create and configure a protection group on FlashArray.

Creating & Configuring Protection Groups

A protection group is the unit of replication on FlashArray. This means that all volumes added to a protection group are replicated to the target selected in the replication group, according to the replication schedule set in the protection group.

Following are the steps to create and configure a protection group via the GUI. Please refer to the FlashArray User's Guide for how to create and configure protection groups via the CLI.

First, we show how to create a protection group via the GUI.

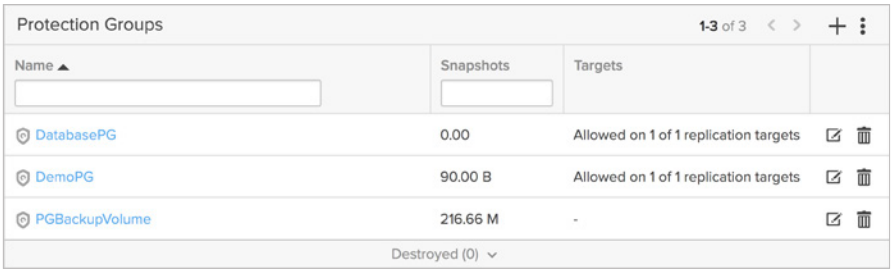

Figure 19. Protection Groups screen

Click on the **+** sign on the right, and the following window will appear. Enter a name for the protection group and click **Create**.

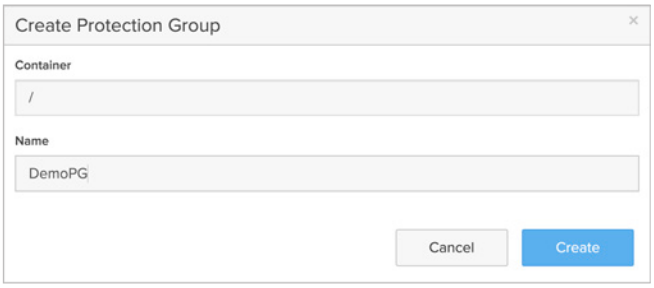

Figure 20. Create Protection Group screen

After a protection group has been created, it will appear in the list of existing protection groups. Next, follow the steps below to configure the protection group.

Protection Group Configuration

The following three steps are required to configure a protection group:

A. Add volumes to the protection group (this is the data that needs to be offloaded to the S3 target)
B. Add the S3 target to the protection group (to specify where to offload the data)
C. Create a replication schedule (to specify how frequently the data should be offloaded to the S3 target and how long it should be retained on the S3 target before expiring)

A. ADDING VOLUMES TO THE PROTECTION GROUP

After a protection group has been created, add volumes to the protection group. The screenshots below show how to add volumes to a protection group.

From the protection groups listed under **Storage > Protection Groups**, select the desired protection group. The following screen will appear. Select **Add Volumes** from the options menu under **Members**.

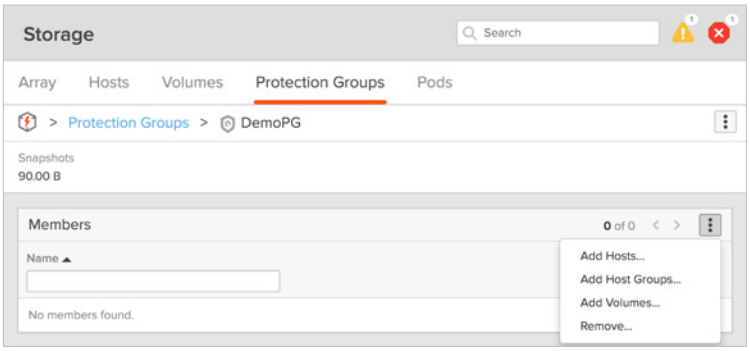

Figure 21. Adding Members to the Protection Group

From the next screen that appears, select the volume/volumes that you want to replicate to the S3 target, and click **Add**.

The newly added volumes should appear in the list of Members. Next, add the S3 target to the protection group.

B. ADDING AN OFFLOAD TARGET TO THE PROTECTION GROUP

The following screenshots show how to connect an offload target to a protection group using the GUI.

Go to **Storage > Protection Groups** and select the desired protection group. The following screen will appear. Select **Add** from the options menu under **Targets**.

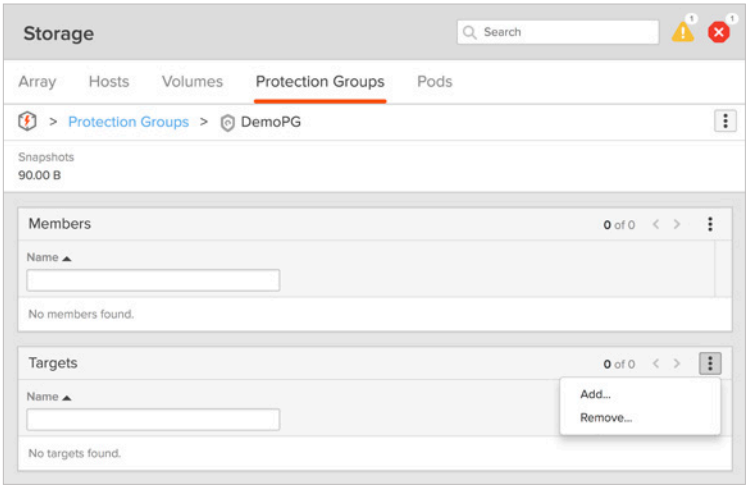

Figure 22. Adding Targets to a Protection Group

The following screen will appear. Select the S3 target and click **Add**.

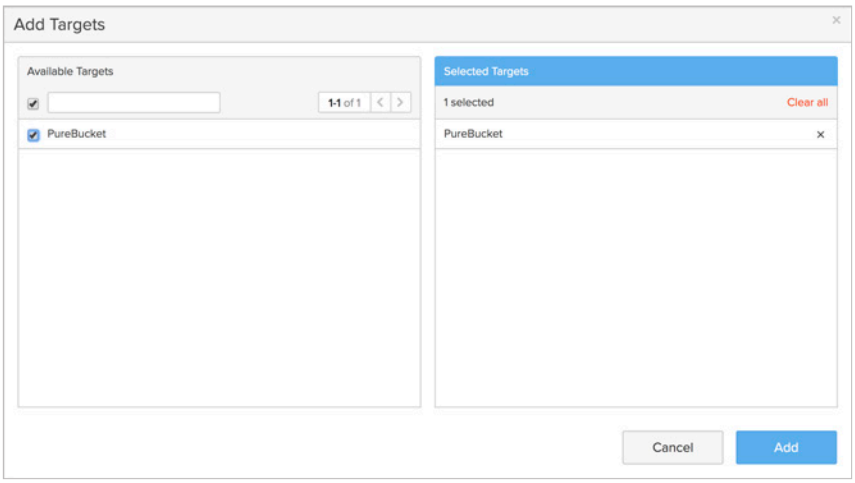

Figure 23. Targets added

When the S3 target has been added, it should appear under the **Targets** tab for the protection group as shown below.

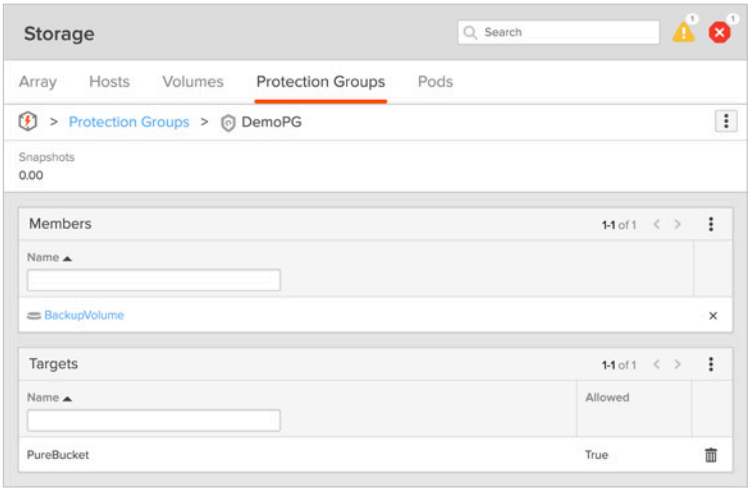

Figure 24. Targets displayed

The last step in configuring a protection group is to create a replication schedule.

C. CREATING A REPLICATION SCHEDULE FOR THE PROTECTION GROUP

The following steps show how to create a schedule for the protection group to take snapshots of the selected volumes and offload them to the S3 target.

Go to **Storage > Protection Groups**, select the desired protection group, and the following screen will appear. Click on the small square box to the right of **Replication Schedule**, as shown on the next page.

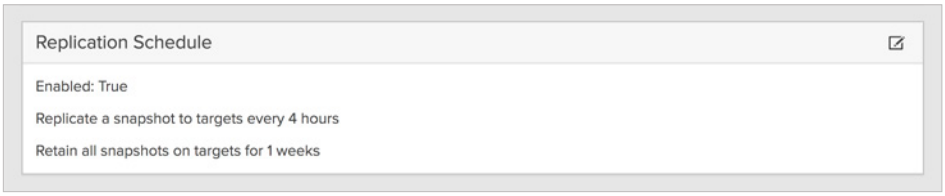

Figure 25. Repolication Schedule set

The following screen will appear. Enable the replication schedule using the radio button on the top left and select **hours** or **days** on the **Replicate a snapshot to targets every** line. Enter the number of hours or days. Note that Snap to S3 does not allow the replication frequency to be more than once every 4 hours.

Next, set the retention period by entering the number of hours or days on the **Retain all snapshots on target for** line.

Optionally, you can also choose to enter an extended retention schedule by entering non-zero values in the **then retain X snapshots per day for Y more days** line. In the example below, snapshots are taken and offloaded every 4 hours; they are retained on the S3 target for 24 hours; finally, one snapshot per day is retained on the S3 target for an additional 30 days.

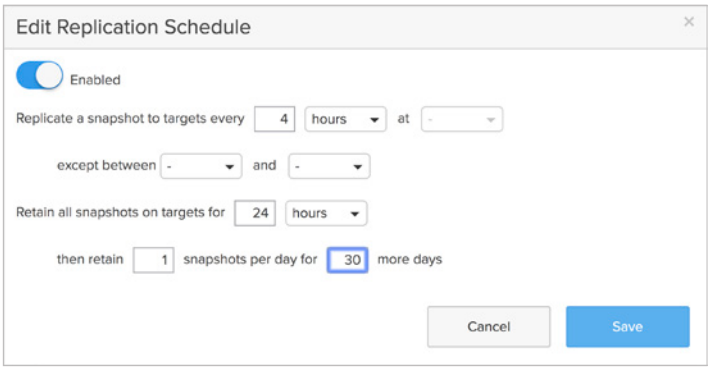

Figure 26. Replication scheduling

Once the schedule is created, the replication process will start immediately. Snapshots will be taken at the scheduled times and offloaded to the S3 target.

CLOUDSNAP REPLICATION FREQUENCY BEST PRACTICE

Though the ideal replication frequency depends on several factors, including the size of the dataset, the network bandwidth between FlashArray and AWS, and the data change rate, etc., in most cases, the best practice for CloudSnap is to replicate data once or twice per day at the most.

Displaying the Offloaded Snapshots in the S3 Bucket

To view protection groups & snapshots on the S3 target, go to **Storage > Array** and click on the S3 target listed under **Offload Targets**. A list of Protections Groups will be displayed in the top half of the screen, and a list of the snapshots on the S3 target will be displayed in the bottom half of the screen, as shown below.

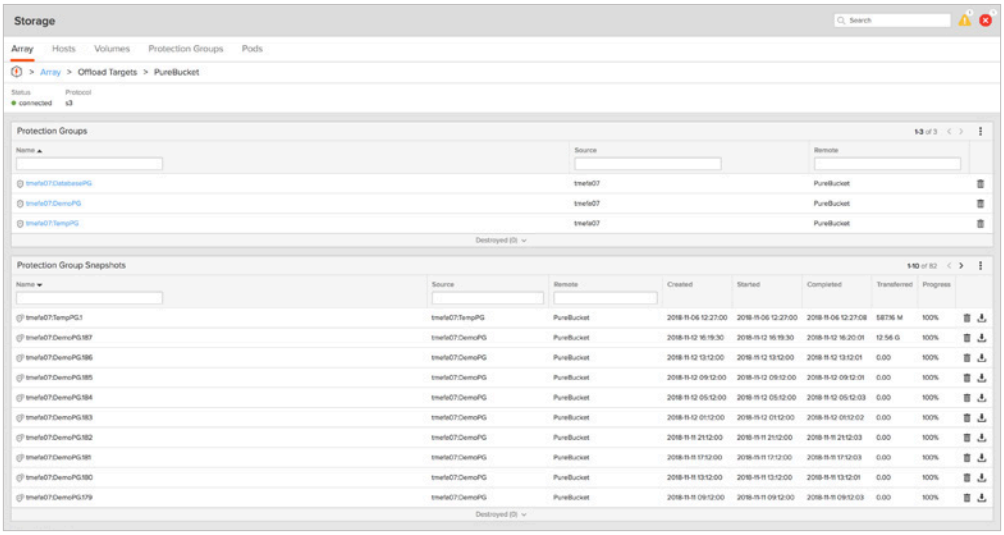

Figure 27. Displaying offloaded snapshots

Restoring a Snapshot from the S3 Target to FlashArray

The following screenshots show how to restore a snapshot from the S3 target to FlashArray.

Under **Storage > Array > Offload Targets**, click on the S3 target to view a list of snapshots on the S3 targets. Select a protection group snapshot to restore by clicking on the download button on its right, as shown below.

Protection Group Snapshots 1-10 of 82

Name	Source	Remote	Created	Started	Completed	Transferred	Progress
tmefa07:TempPG.1	tmefa07:TempPG	PureBucket	2018-11-06 12:27:00	2018-11-06 12:27:00	2018-11-06 12:27:08	587.16 M	100%
tmefa07:DemoPG.187	tmefa07:DemoPG	PureBucket	2018-11-12 16:19:30	2018-11-12 16:19:30	2018-11-12 16:20:01	12.56 G	100%
tmefa07:DemoPG.186	tmefa07:DemoPG	PureBucket	2018-11-12 13:12:00	2018-11-12 13:12:00	2018-11-12 13:12:01	0.00	100%
tmefa07:DemoPG.185	tmefa07:DemoPG	PureBucket	2018-11-12 09:12:00	2018-11-12 09:12:00	2018-11-12 09:12:01	0.00	100%
tmefa07:DemoPG.184	tmefa07:DemoPG	PureBucket	2018-11-12 05:12:00	2018-11-12 05:12:00	2018-11-12 05:12:03	0.00	100%
tmefa07:DemoPG.183	tmefa07:DemoPG	PureBucket	2018-11-12 01:12:00	2018-11-12 01:12:00	2018-11-12 01:12:02	0.00	100%
tmefa07:DemoPG.182	tmefa07:DemoPG	PureBucket	2018-11-11 21:12:00	2018-11-11 21:12:00	2018-11-11 21:12:03	0.00	100%
tmefa07:DemoPG.181	tmefa07:DemoPG	PureBucket	2018-11-11 17:12:00	2018-11-11 17:12:00	2018-11-11 17:12:03	0.00	100%
tmefa07:DemoPG.180	tmefa07:DemoPG	PureBucket	2018-11-11 13:12:00	2018-11-11 13:12:00	2018-11-11 13:12:01	0.00	100%
tmefa07:DemoPG.179	tmefa07:DemoPG	PureBucket	2018-11-11 09:12:00	2018-11-11 09:12:00	2018-11-11 09:12:03	0.00	100%

Destroyed (0)

Figure 28. Snapshots on S3 targets

The following screen will appear, listing all the volume snapshots in the protection group. Select the volume snapshots that you want to restore. When selecting snapshots to restore, you can optionally add a suffix to the names of the restored snapshots to make it easier to identify the restored snapshots. Then click **Get**.

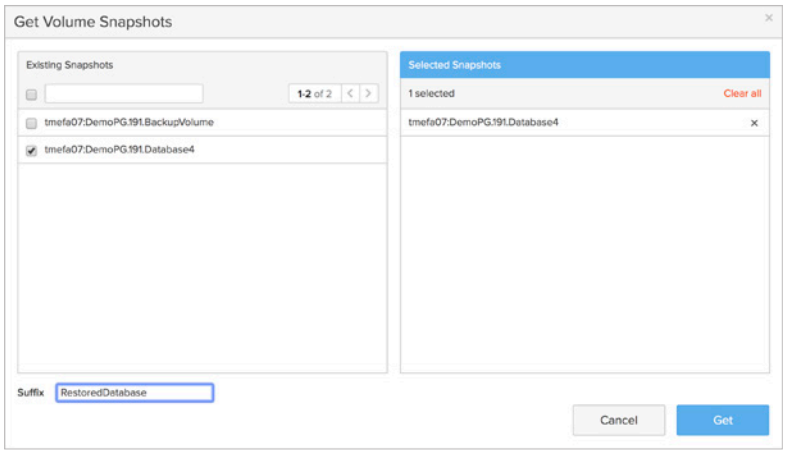

Figure 30. Volume snapshots in the protection group

Once the snapshots are restored to FlashArray, they appear in the **Volume Snapshots** tab under the **Storage > Volumes** menu, as shown below.

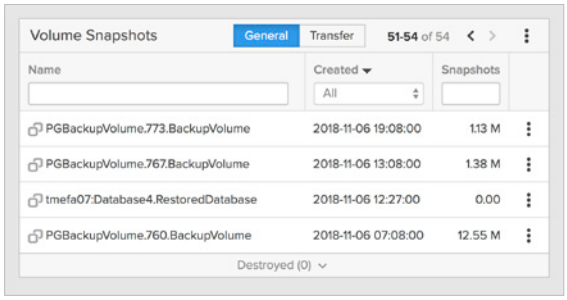

Figure 31. Volume Snapshots in the Volumes menu

Creating a Volume from the Restored Snapshot

After a snapshot has been restored to FlashArray, a new volume can be created from the snapshot, or an existing volume can be overwritten by it. The screens on the next page show how to create a volume from a restored snapshot using the FlashArray GUI. Under the **Storage > Volumes** menu, click on the options menu to the right of the snapshot, and select **Copy**.

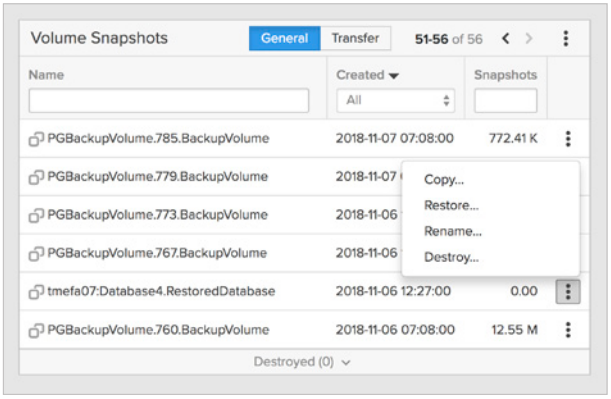

Figure 32. Copying a snapshot

The following screen will appear. Enter a name for the volume to be created and click on **Copy**.

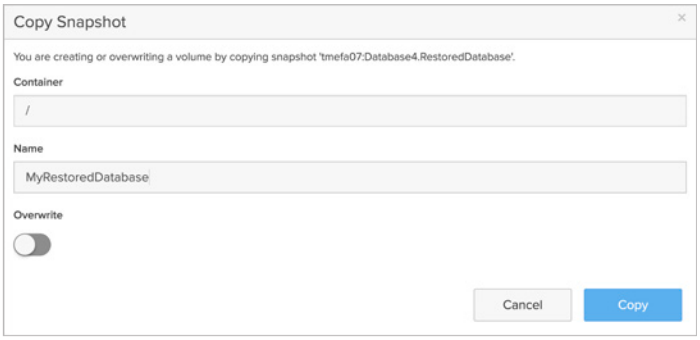

Figure 33. Naming the copied snapshot

When the volume has been created, it will appear in the list of volumes under **Storage > Volumes**.

Connecting the Newly Created Volume to a Host

To access the newly created volume from a host, connect the volume to a host. Click on the volume, and under **Connected Hosts**, from the options menu, select **Connect**, as shown below.

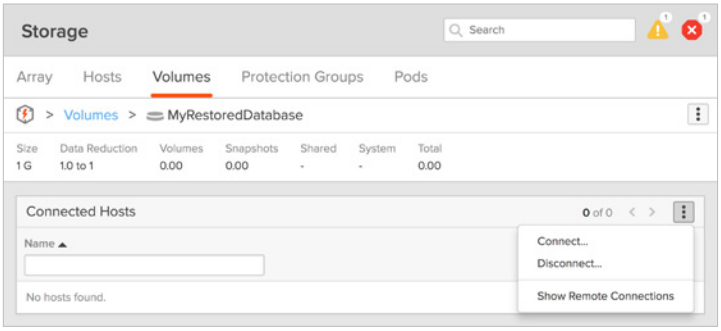

Figure 34. Connected hosts

The following screen will appear. Select a host and click **Connect**.

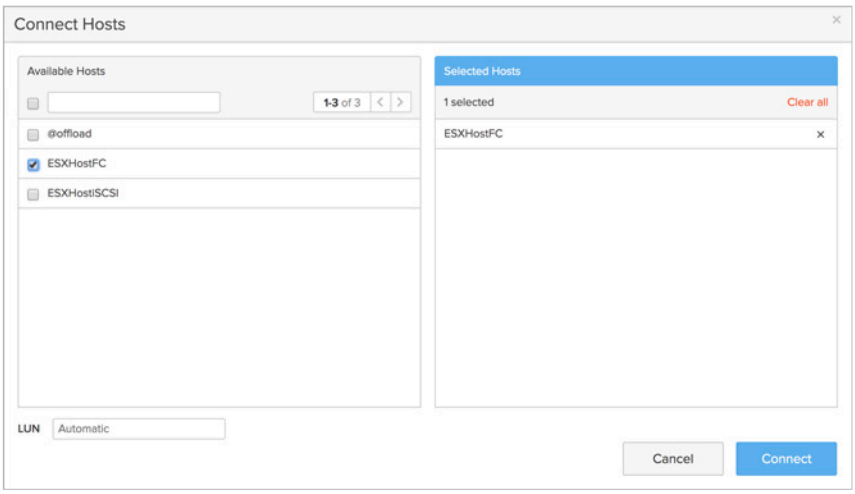

Figure 35. Selecting a host to connect

When the volume is connected to a host, it will be visible from the host. Connect to the volume from the host to access the restored data.

ps_wp22p_purity-cloudsnap-setup-best-practices_01