Dataprotection for OpenShift Virtualization

NetApp Solutions

NetApp
April 17, 2024

Table of Contents

Dataprotection for OpenShift Virtualization

Data protection for VMs in OpenShift Virtualization using OpenShift API for Data Protection (OADP)

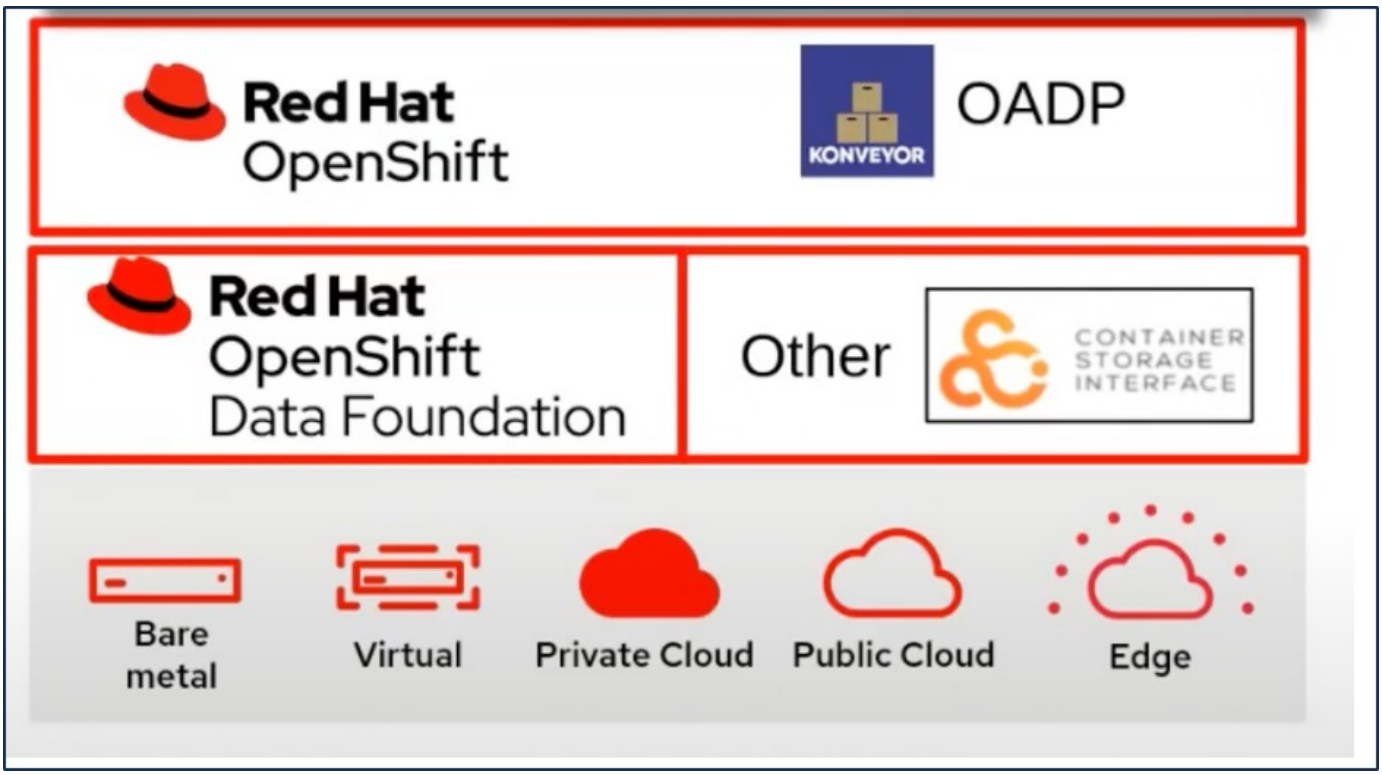
Banu Sundhar, NetApp

This reference document provides details for creating backups of VMs using the OpenShift API for Data Protection (OADP) with Velero and moving it to ONTAP S3. The backups of PVCs of the VMs are created using CSI Astra Trident Snapshots.

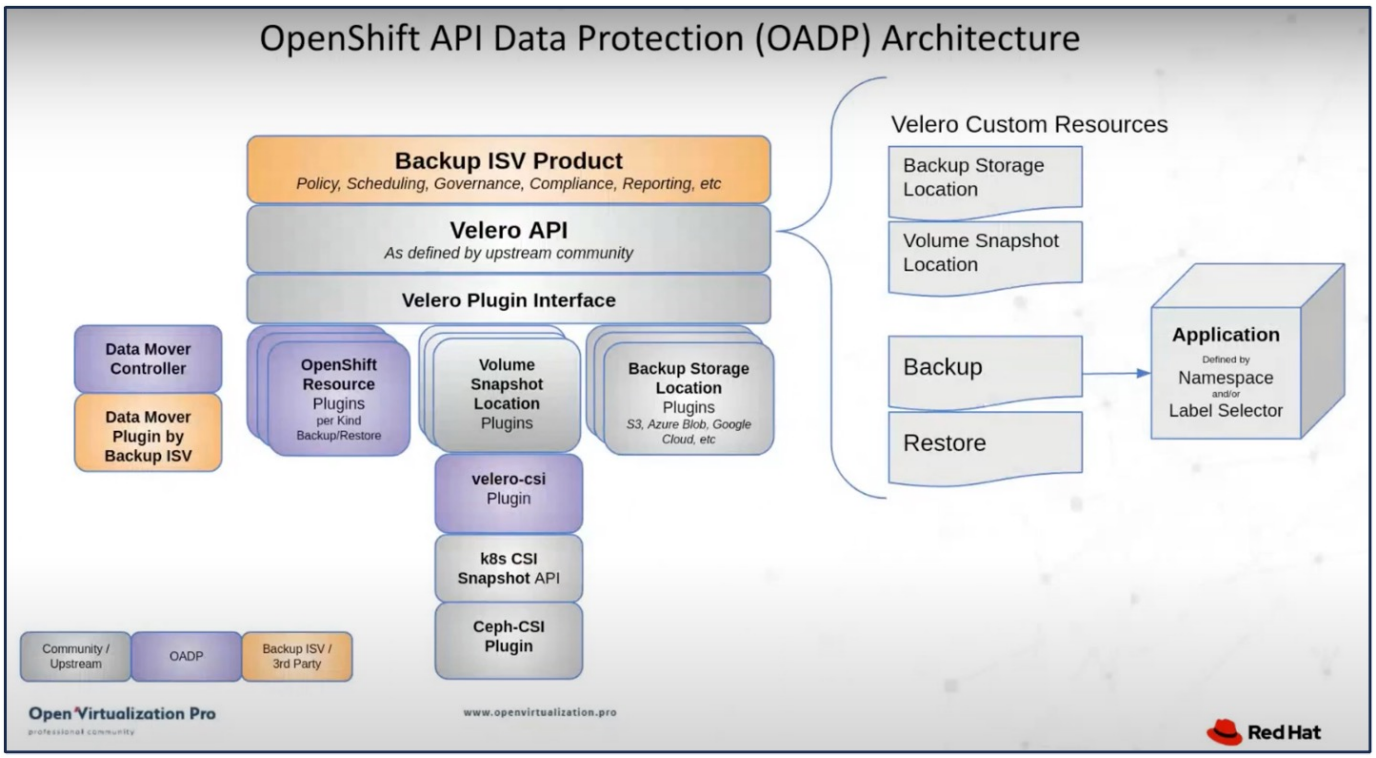
Virtual machines in the OpenShift Virtualization environment are containerized applications that run in the worker nodes of your OpenShift Container platform. It is important to protect the VM metadata as well as the persistent disks of the VMs, so that when they are lost or corrupted, you can recover them.

The persistent disks of the OpenShift Virtualization VMs can be backed by ONTAP storage integrated to the OpenShift Cluster using Astra Trident CSI. In this section we use OpenShift API for Data Protection (OADP) to perform backup of VMs including its data volumes to ONTAP Object Storage. We then restore from the backup when needed.

OADP enables backup, restore, and disaster recovery of applications on an OpenShift cluster. Data that can be protected with OADP include Kubernetes resource objects, persistent volumes, and internal images.



Red Hat OpenShift has leveraged the solutions developed by the OpenSource communities for data protection. Velero is an open-source tool to safely backup and restore, perform disaster recovery, and migrate Kubernetes cluster resources and persistent volumes. To use Velero easily, OpenShift has developed the OADP operator and the Velero plugin to integrate with the CSI storage drivers. The core of the OADP APIs that are exposed are based on the Velero APIs. After installing the OADP operator and configuring it, the backup/restore operations that can be performed are based on the operations exposed by the Velero API.



OADP 1.3 is available from the operator hub of OpenShift cluster 4.12 and later. It has a built-in Data Mover that can move CSI volume snapshots to a remote object store. This provides portability and durability by moving snapshots to an object storage location during backup. The snapshots are then available for restoration after disasters.

The following are the component versions for the examples in this section

- OpenShift Cluster 4.14
- OpenShift Virtualization installed via OperatorOpenShift Virtualization Operator provided by Red Hat
- OADP Operator 1.13 provided by Red Hat
- Velero CLI 1.13 for Linux
- Astra Trident 24.02
- ONTAP 9.12

Installation of OpenShift API for Data Protection (OADP) Operator

Prerequisites

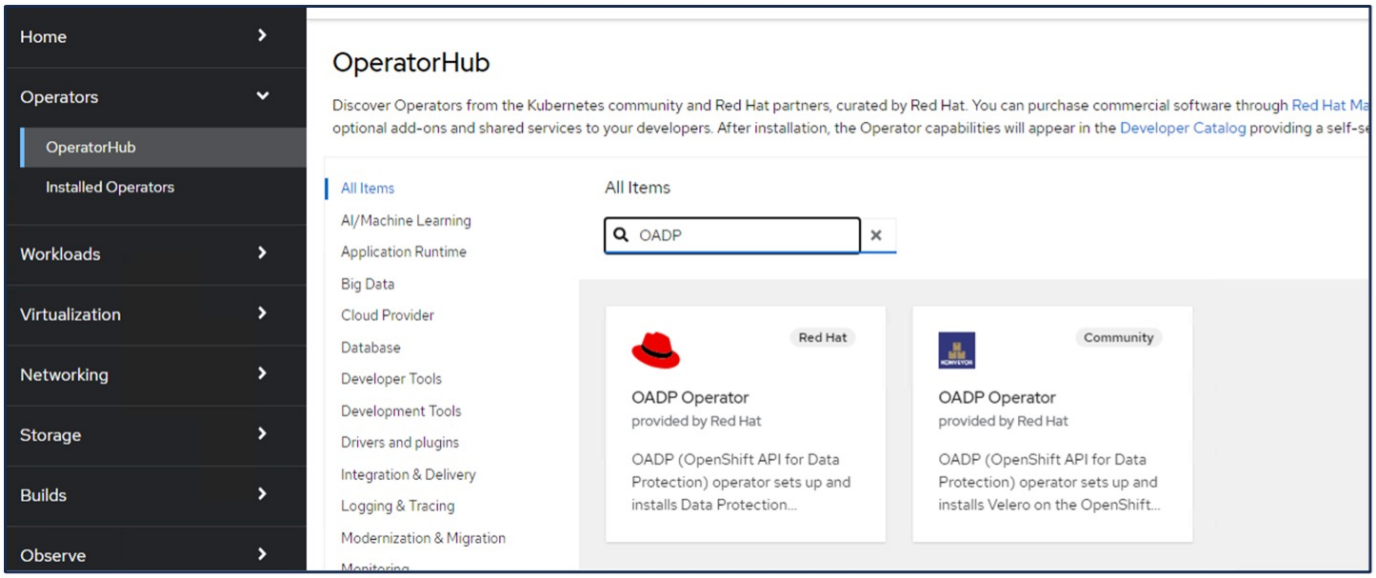
- A Red Hat OpenShift cluster (later than version 4.12) installed on bare-metal infrastructure with RHCOS worker nodes
- A NetApp ONTAP cluster integrated with the cluster using Astra Trident
- A Trident backend configured with an SVM on ONTAP cluster
- A StorageClass configured on the OpenShift cluster with Astra Trident as the provisioner
- Trident Snapshot class created on the cluster

- Cluster-admin access to Red Hat OpenShift cluster
- Admin access to NetApp ONTAP cluster
- OpenShift Virtualization operator installed and configured
- VMs deployed in a Namespace on OpenShift Virtualization
- An admin workstation with tridentctl and oc tools installed and added to $PATH

> If you want to take a backup of a VM when it is in the Running state, then you must install the QEMU guest agent on that virtual machine. If you install the VM using an existing template, then QEMU agent is installed automatically. QEMU allows the guest agent to quiesce in-flight data in the guest OS during the snapshot process, and avoid possible data corruption. If you do not have QEMU installed, you can stop the virtual machine before taking a backup.

Steps to install OADP Operator

1. Go to the Operator Hub of the cluster and select Red Hat OADP operator. In the Install page, use all the default selections and click install. On the next page, again use all the defaults and click Install. The OADP operator will be installed in the namespace called openshift-adp.



OADP Operator

1.3.0 provided by Red Hat

Install

Channel

stable-1.3

Version

1.3.0

Capability level

- Basic Install
- Seamless Upgrades
- Full Lifecycle
- Deep Insights
- Auto Pilot

Source

Red Hat

Provider

Red Hat

Infrastructure features

Disconnected

OpenShift API for Data Protection (OADP) operator sets up and installs Velero on the OpenShift platform, allowing users to backup and restore applications.

Backup and restore Kubernetes resources and internal images, at the granularity of a namespace, using a version of Velero appropriate for the installed version of OADP.

OADP backs up Kubernetes objects and internal images by saving them as an archive file on object storage. OADP backs up persistent volumes (PVs) by creating snapshots with the native cloud snapshot API or with the Container Storage Interface (CSI). For cloud providers that do not support snapshots, OADP backs up resources and PV data with Restic or Kopia.

- Installing OADP for application backup and restore
- Installing OADP on a ROSA cluster and using STS, please follow the Getting Started Steps 1-3 in order to obtain the role ARN needed for using the standardized STS configuration flow via OLM
- Frequently Asked Questions

Project: All Projects

Installed Operators

Installed Operators are represented by ClusterServiceVersions within this Namespace. For more information, see the Understanding Operators documentation Operator and ClusterServiceVersion using the Operator SDK.

Name | Search by name... /

Name	Namespace	Managed Namespaces	Status
OpenShift Virtualization 4.14.4 provided by Red Hat	NS openshift-cnv	NS openshift-cnv	Succeeded Up to date
OADP Operator 1.3.0 provided by Red Hat	NS openshift-adp	NS openshift-adp	Succeeded Up to date
Package Server 0.0.1-snapshot provided by	NS openshift-operator-lifecycle-manager	NS openshift-operator-lifecycle-manager	Succeeded

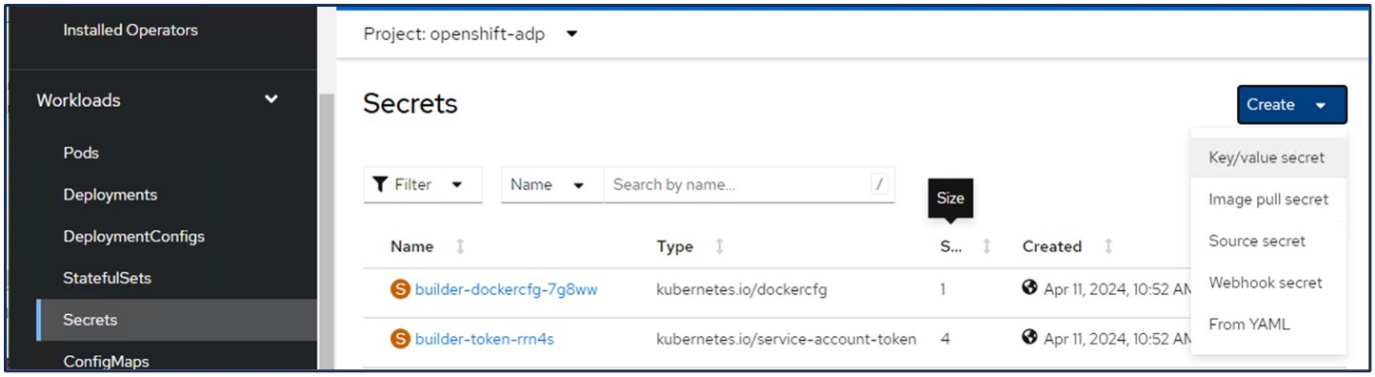
Prerequisites for Velero configuration with Ontap S3 details:

After the installation of the operator succeeds, configure the instance of Velero.
Velero can be configured to use S3 compatible Object Storage. Configure ONTAP S3 using the procedures shown in the Object Storage Management section of ONTAP documentation. You will need the following information from your ONTAP S3 configuration to integrate with Velero.

- A Logical Interface (LIF)IP that can be used to access S3
- User credentials to access S3 that includes the access key and the secret access key
- A bucket name in S3 for backups with access permissions for the user
- For secure access to the Object storage, TLS certificate should be installed on the Object Storage server.

Steps to configure Velero

- First, create a secret for the ONTAP S3 user credentials. This will be used to configure Velero later. You can create a secret from the CLI or from the web console.
  To create a secret from the web console, select Secrets, then click on Key/Value Secret.
  Provide the values for the credential name, key and the value as shown. Be sure to use the Access Key Id and Secret Access Key of your S3 user.



Project: openshift-adp

Create key/value secret

Key/value secrets let you inject sensitive data into your application as files or environment variables.

Secret name *

cloud-credentials

Unique name of the new secret.

Key *

cloud

Value

Browse...

Drag and drop file with your value here or browse to upload it.

```
[default]
aws_access_key_id=<Access Key Id of S3 user>
aws_secret_access_key=<Secret Access Key of S3 user>
```

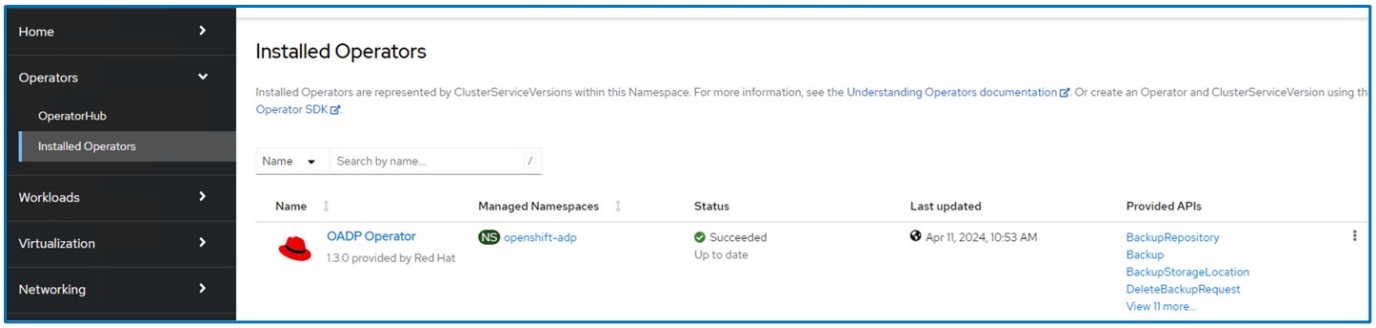
Add key/value

Create Cancel

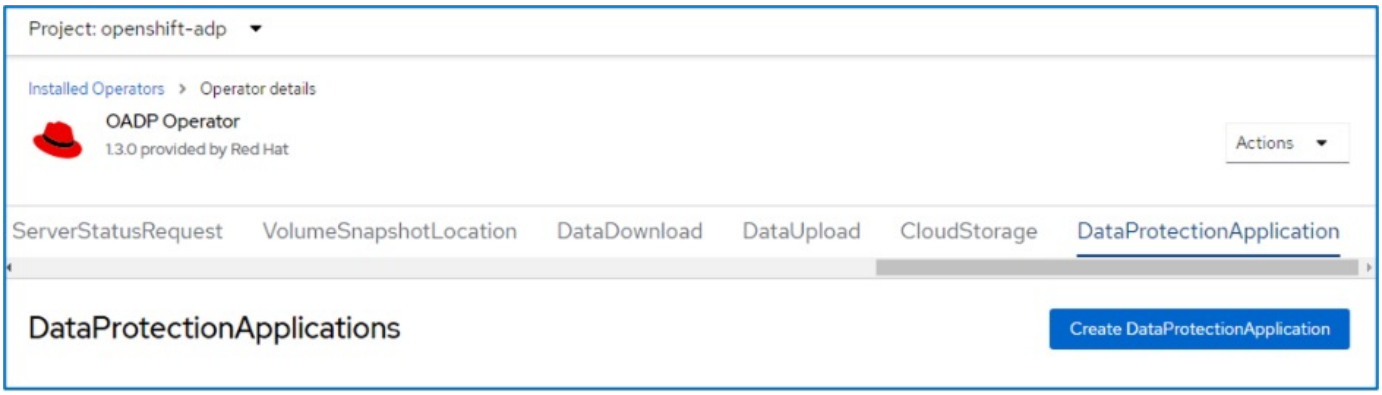
To create the default secret named cloud credentials from the CLI you can use the following command. If the backup and snapshot locations use the same credentials, you just need to create the default secret as shown above. For other scenarios, please see the OADP documentation.

```
# oc create secret generic cloud-credentials --namespace openshift-adp --
from-file cloud=cloud-credentials.txt

credentials.txt file contains the Access Key Id and the Secret Access Key of
the S3 user in the following format:

[default]
aws_access_key_id=<Access Key Id of S3 user>
aws_secret_access_key=<Secret Access Key of S3 user>
```

- Next, to Configure Velero, select Installed Operators from the menu item under Operators, click on OADP operator, and then select the DataProtectionApplication tab.



Click on Create DataProtectionApplication. In the form view, provide a name for the DataProtection Application or use the default name.



Now go to the YAML view and replace the default information or add the information as shown in the yaml file below.

```
spec:
  backupLocations:
    - velero:
        config:
          insecureSkipTLSVerify: 'true' //use this for https communication
with ONTAP S3
          profile: default
          region: us-east
          s3ForcePathStyle: 'True' //This allows use of IP in s3URL
          s3Url: 'https://10.xx.xx.xx' //Ensure TLS certificate for S3 is
configured
        credential:
          key: cloud
          name: cloud-credentials //previously created secret named cloud-
credentials
        default: true
        objectStorage:
          bucket: velero //Your bucket name previously created in S3 for
backups
          prefix: demobackup //The folder that will be created in the
bucket
        provider: aws
  configuration:
    nodeAgent:
      enable: true
      uploaderType: kopia
                    //default Data Mover uses Kopia to move snapshots to
Object  Storage
    velero:
      defaultPlugins:
        - csi //Add this plugin
        - openshift
        - aws
        - kubevirt //Add this plugin
  snapshotLocations:
    - velero:
        config:
          profile: default
          region: us-east
        provider: aws
```

The above YAML has the following sections in the spec that needs to be configured appropriately similar to the example:

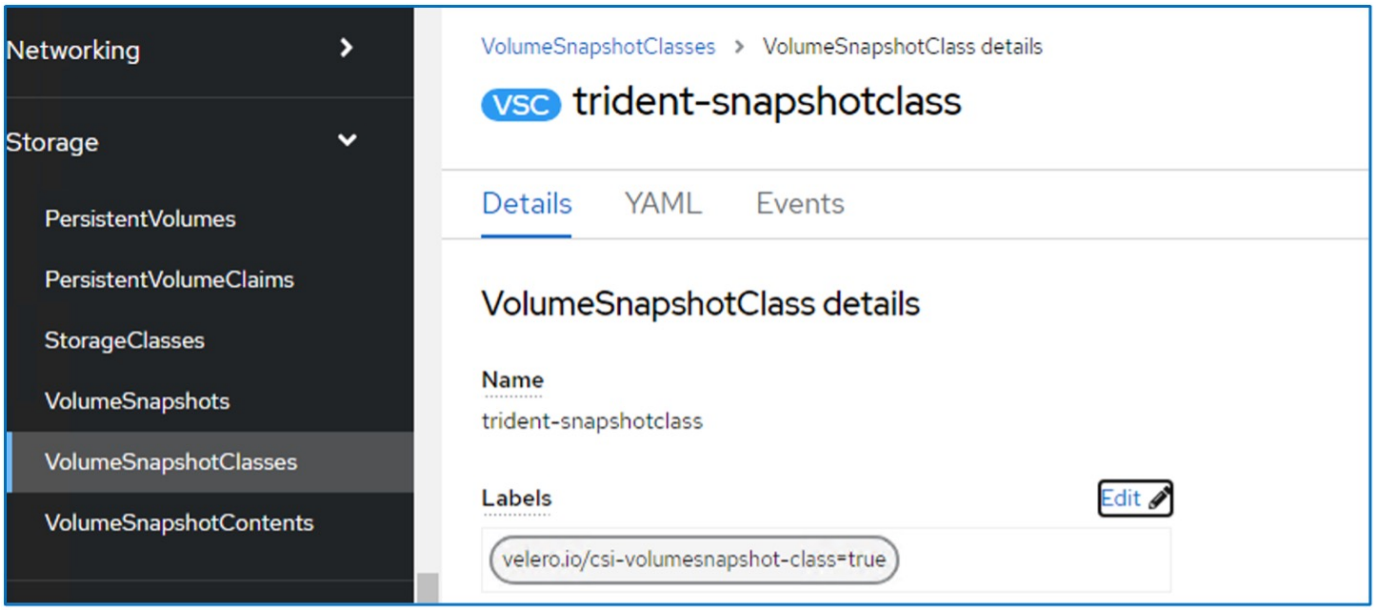
backupLocations
ONTAP S3 (with its credentials and other information as shown in the yaml) is configured as the default

BackupLocation for velero.

snapshotLocations
ONTAP S3 is configured as the default location for PVC snapshots for Velero.

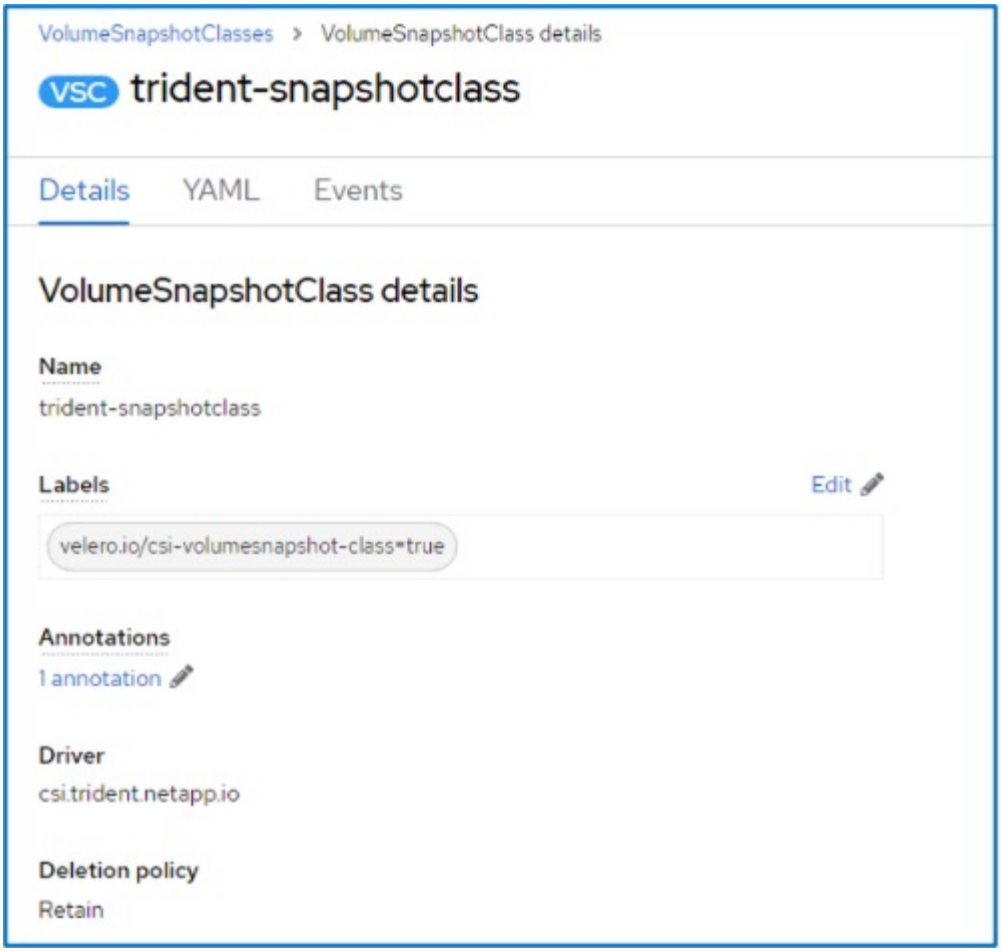
Enable CSI
Add csi to the defaultPlugins for Velero to back up persistent volumes with CSI snapshots.
The Velero CSI plugins, to backup CSI backed PVCs, will choose the VolumeSnapshotClass in the cluster that has **velero.io/csi-volumesnapshot-class** label set on it. For this

- You must have the trident VolumeSnapshotClass created.
- Edit the label of the trident-snapshotclass and set it to **velero.io/csi-volumesnapshot-class=true** as shown below.

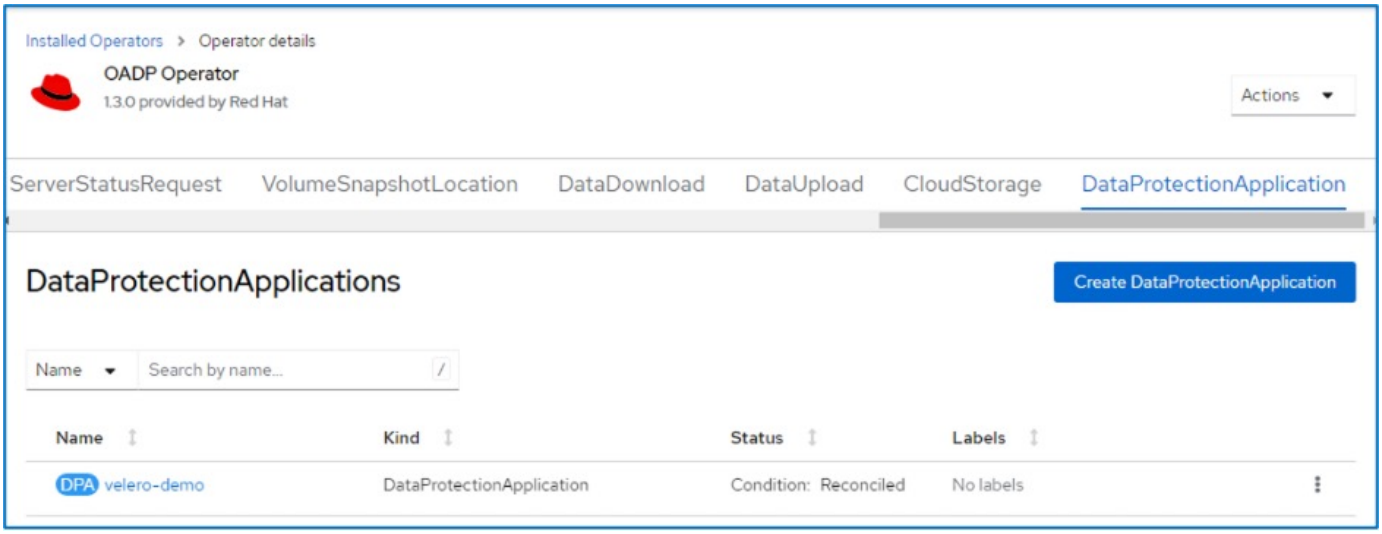


Ensure that the snapshots can persist even if the VolumeSnapshot objects are deleted. This can be done by setting the deletionPolicy to Retain. If not, deleting a namespace will completely lose all PVCs ever backed up in it.

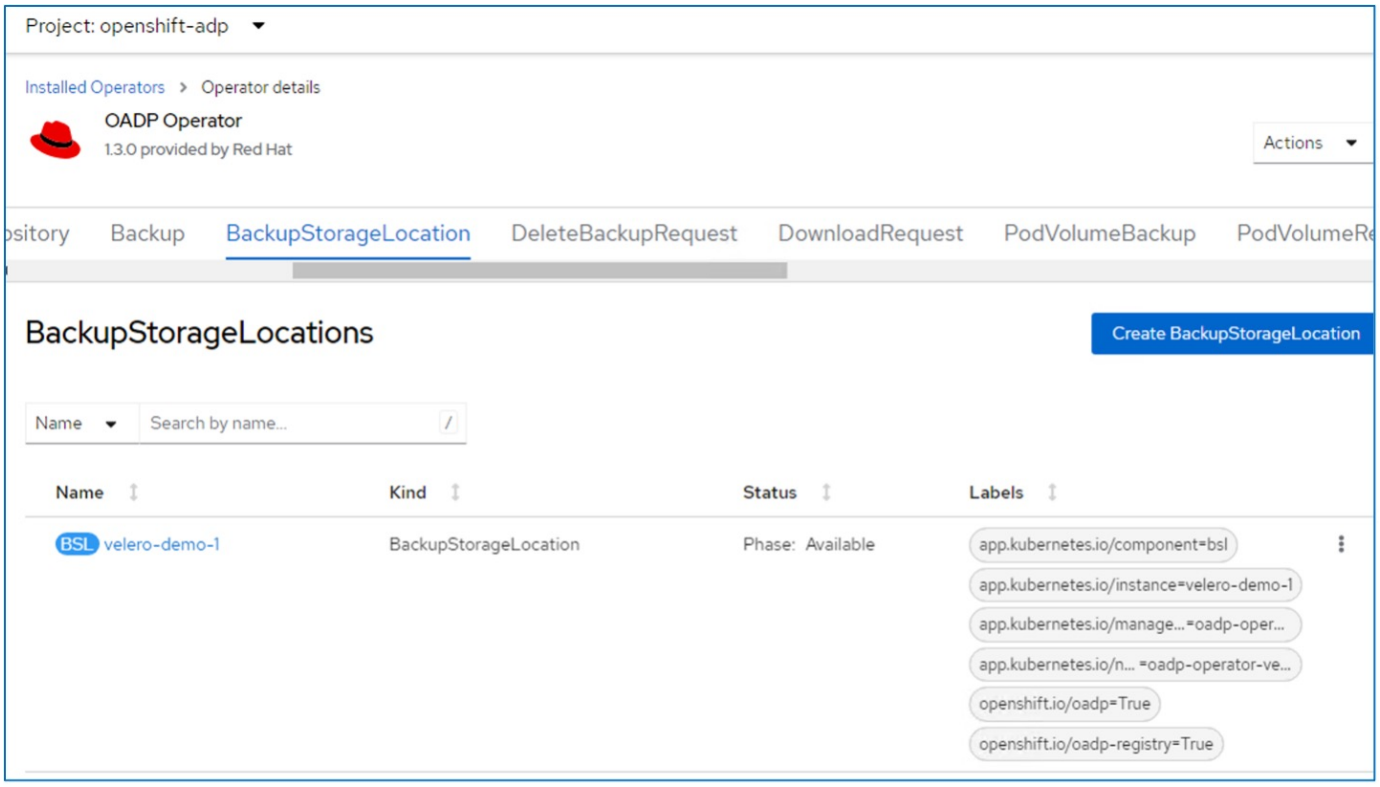
```
apiVersion: snapshot.storage.k8s.io/v1
kind: VolumeSnapshotClass
metadata:
  name: trident-snapshotclass
driver: csi.trident.netapp.io
deletionPolicy: Retain
```



Ensure that the DataProtectionApplication is created and is in condition:Reconciled.



The OADP operator will create a corresponding BackupStorageLocation.This will be used when creating a backup.



Creating on-demand backup for VMs in OpenShift virtualization

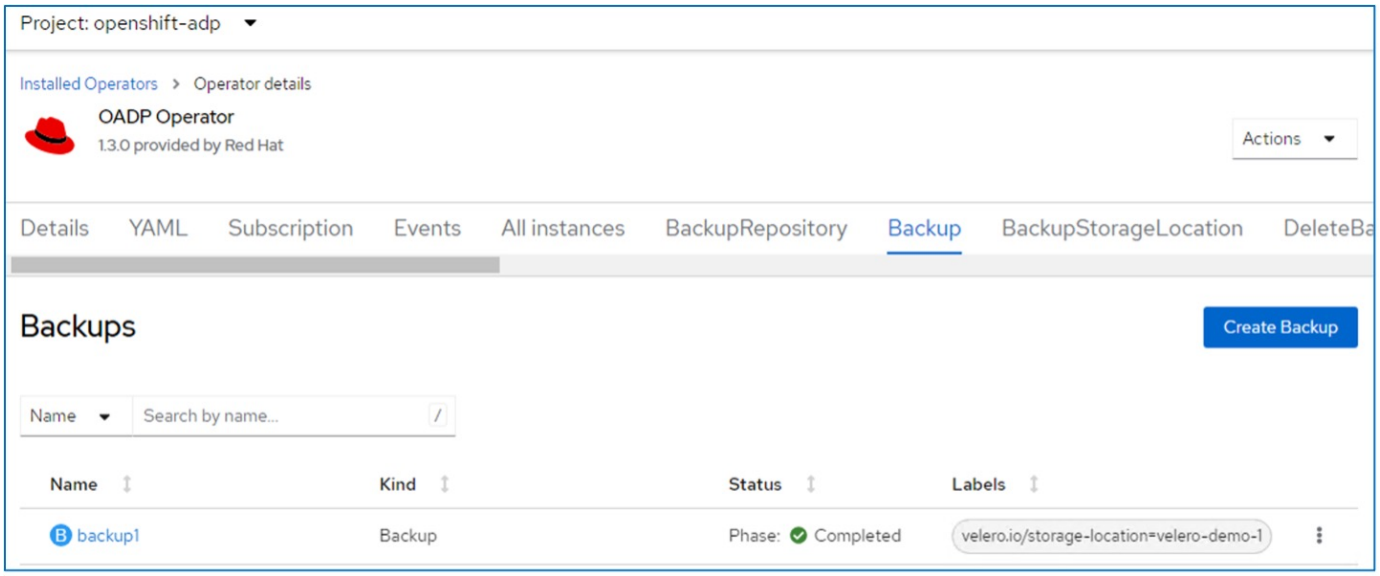
Steps to create a backup of a VM

To create an on-demand backup of the entire VM (VM metadata and VM disks), click on the **Backup** tab. This creates a Backup Custom Resource (CR). A sample yaml is provided to create the Backup CR. Using this yaml, the VM and its disks in the specified namespace will be backed up. Additional parameters can be set as shown in the documentation.

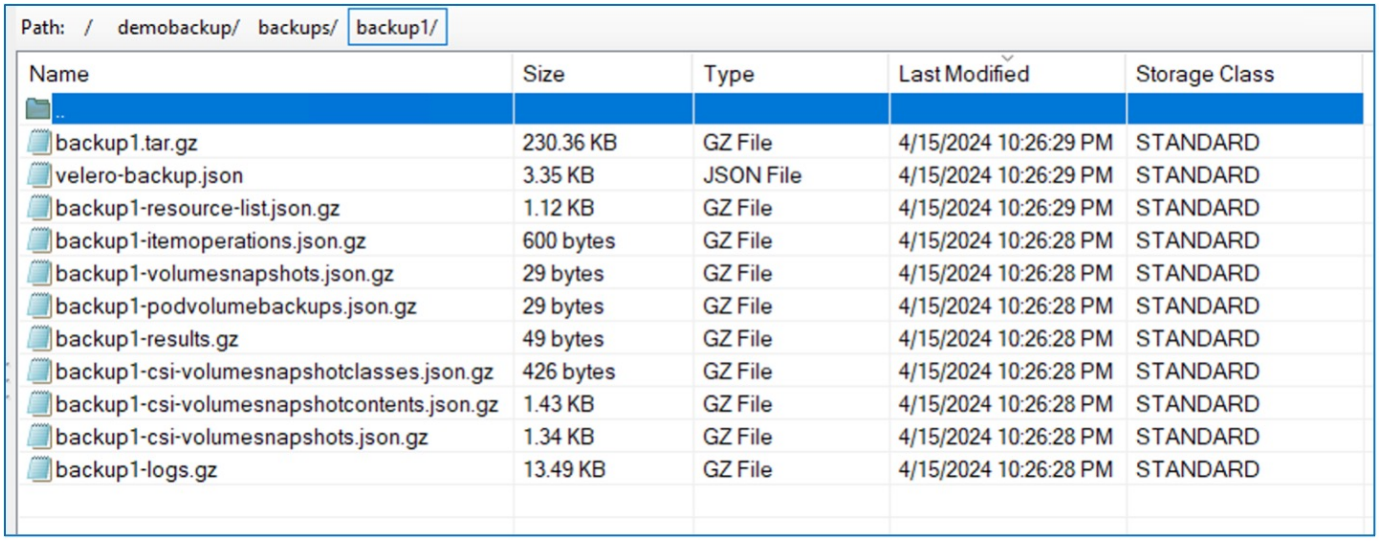
A snapshot of the persistent volumes backing the disks will be created by the CSI and will be moved to the object storage location provided in the yaml. The backup will remain in the system for 30 days as specified in the ttl.

```
apiVersion: velero.io/v1
kind: Backup
metadata:
  name: backup1
  namespace: openshift-adp
spec:
  includedNamespaces:
  - virtual-machines-demo
  snapshotVolumes: true
  storageLocation: velero-demo-1
  ttl: 720h0m0s
```

Once the backup completes, its Phase should show as completed.



You can inspect the backup in the Object storage with the help of an S3 browser application. The path of the backup shows in the configured bucket with the prefix name (velero/demobackup). You can see the contents of the backup includes the volume snapshots, logs, and other metadata of the virtual machine.



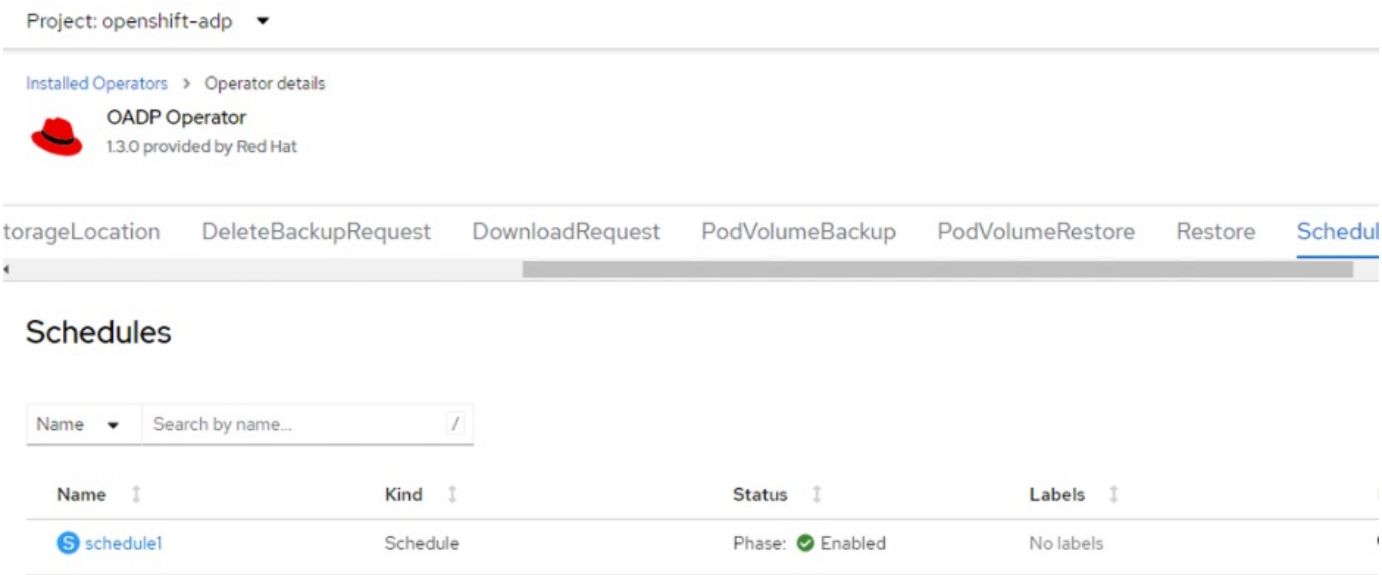
Creating scheduled backups for VMs in OpenShift virtualization

To create a backups on a schedule, you need to create a Schedule Custom Resource.
The schedule is simply a Cron expression allowing you to specify the time at which you want to create the backup. A sample yaml to create a Schedule CR.

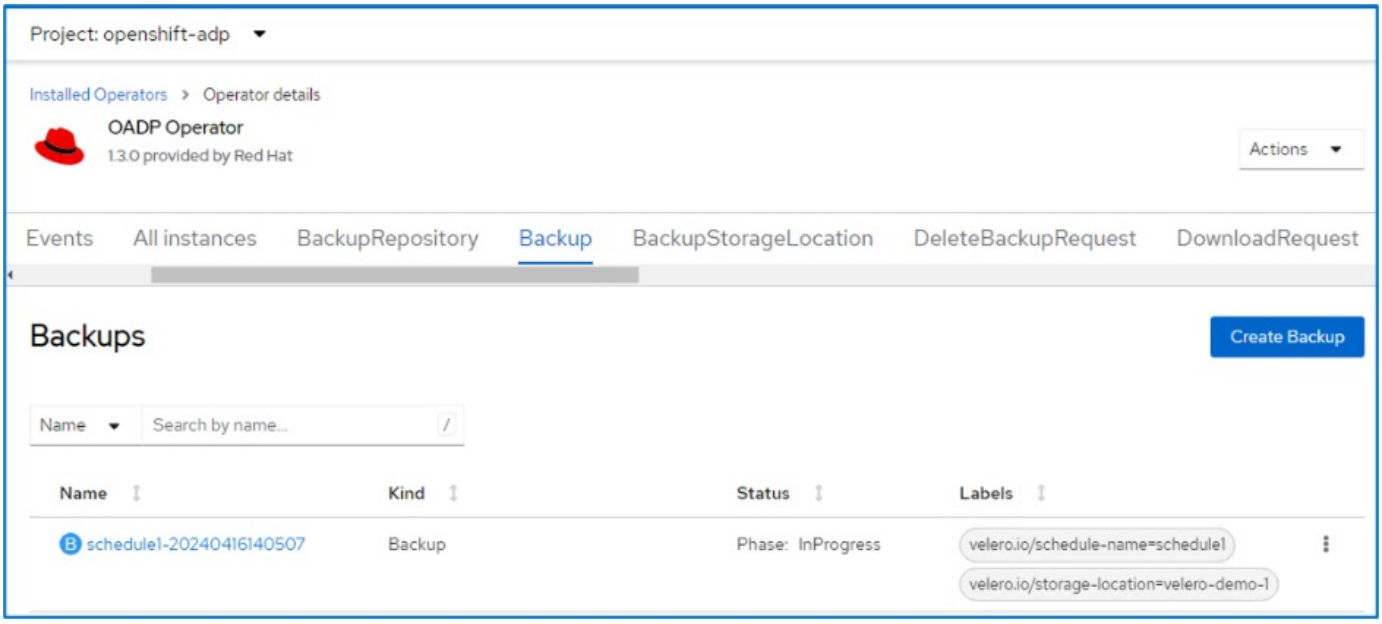
```
apiVersion: velero.io/v1
kind: Schedule
metadata:
  name: <schedule>
  namespace: openshift-adp
spec:
  schedule: 0 7 * * *
  template:
    hooks: {}
    includedNamespaces:
    - <namespace>
    storageLocation: velero-demo-1
    defaultVolumesToFsBackup: true
    ttl: 720h0m0s
```

The Cron expression 0 7 * * * means a backup will be created at 7:00 every day.
The namespaces to be included in the backup and the storage location for the backup are also specified. So instead of a Backup CR, Schedule CR is used to create a backup ar the specified time and frequency.

Once the schedule is created, it will be Enabled.



Backups will be created according to this schedule, and can be viewed from the Backup tab.



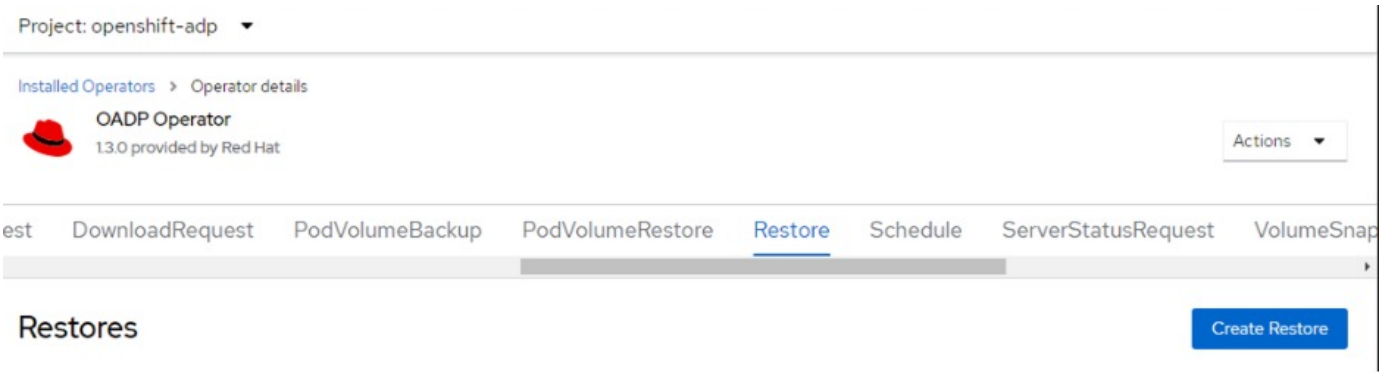
Restore a VM from a backup

Prerequisites

To restore from a backup, let us assume that the namespace where the virtual machine existed got accidentally deleted.

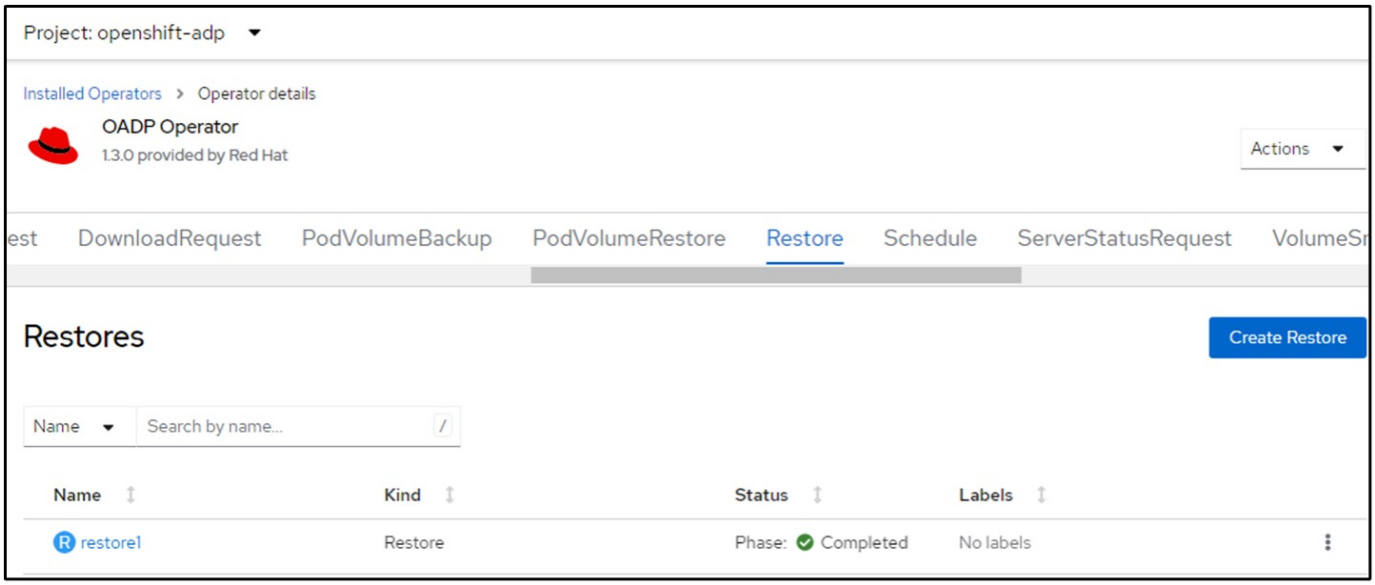
Steps to perform a Restore

To restore from the backup that we just created, we need to create a Restore Custom Resource (CR). We need to provide it a name, provide the name of the backup that we want to restore from and set the restorePVs to true.
Additional parameters can be set as shown in the documentation. Click on Create button.



```
apiVersion: velero.io/v1
kind: Restore
metadata:
  name: restore1
  namespace: openshift-adp
spec:
  backupName: backup1
  restorePVs: true
```

When the phase shows completed, you will see that the virtual machines have been restored to the state when the snapshot was taken. (If the backup was created when the VM was running, restoring the VM from the backup will start the restored VM and bring it to a running state)



Deleting backups and restores in using Velero

Deleting a backup

You can delete a Backup CR without deleting the Object Storage data by using the OC CLI tool.

```
oc delete backup <backup_CR_name> -n <velero_namespace>
```

If you want the delete the Backup CR and delete the associated object storage data, you can do so by using the Velero CLI tool.

Download the CLI as given in the instructions in the Velero documentation.

Execute the following delete command using the Velero CLI

```
velero backup delete <backup_CR_name> -n <velero_namespace>
```

You can also delete the Restore CR using the Velero CLI

```
velero restore delete restore --namespace openshift-adp
```

You can use oc command as well as the UI to delete the restore CR

```
oc delete backup <backup_CR_name> -n <velero_namespace>
```