

Data Protection of VMs using Third Party Tools

NetApp Solutions

NetApp July 31, 2024

This PDF was generated from https://docs.netapp.com/us-en/netapp-solutions/containers/rh-os-n_use_case_openshift_virtualization_dataprotection_overview.html on July 31, 2024. Always check docs.netapp.com for the latest.

Table of Contents

D	ata Protection of VMs using Third Party Tools	1
	Data protection for VMs in OpenShift Virtualization using OpenShift API for Data Protection (OADP)	1
	Installation of OpenShift API for Data Protection (OADP) Operator.	3
	Creating on-demand backup for VMs in OpenShift Virtualization	. 12
	Restore a VM from a backup	. 15
	Deleting backups and restores in using Velero	. 21

Data Protection of VMs using Third Party Tools

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

Author: Banu Sundhar, NetApp

This section of the reference document provides details for creating backups of VMs using the OpenShift API for Data Protection (OADP) with Velero on NetApp ONTAP S3 or NetApp StorageGRID S3. The backups of Persistent Volumes(PVs) of the VM disks are created using CSI Astra Trident Snapshots.

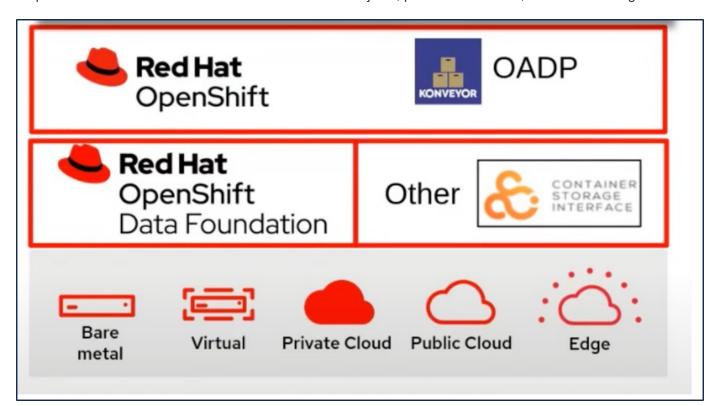
Virtual machines in the OpenShift Virtualization environment are containerized applications that run in the worker nodes of your OpenShift Container platform. It is important to protect the VM metadata as well as the persistent disks of the VMs, so that when they are lost or corrupted, you can recover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

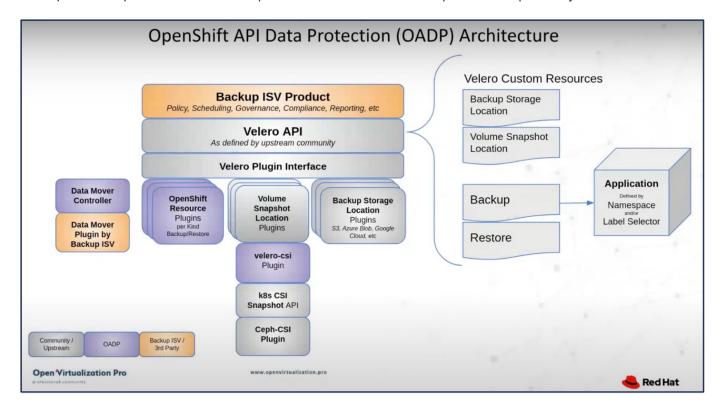
- ONTAP Object Storage
- · StorageGrid

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 versions of the various components used 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

Astra Trident CSI OpenShift API for Data Protection (OADP) Velero

Installation of OpenShift API for Data Protection (OADP) Operator

This section outlines the installation of OpenShift API for Data Protection (OADP) Operator.

Prerequisites

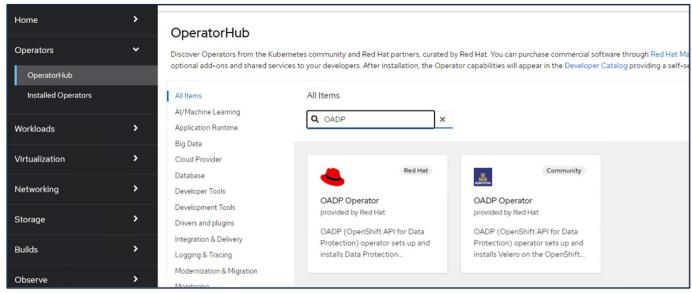
- A Red Hat OpenShift cluster (later than version 4.12) installed on bare-metal infrastructure with RHCOS worker nodes
- A NetApp ONTAP cluster integrated with the cluster using Astra Trident
- A Trident backend configured with an SVM on ONTAP cluster
- A StorageClass configured on the OpenShift cluster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

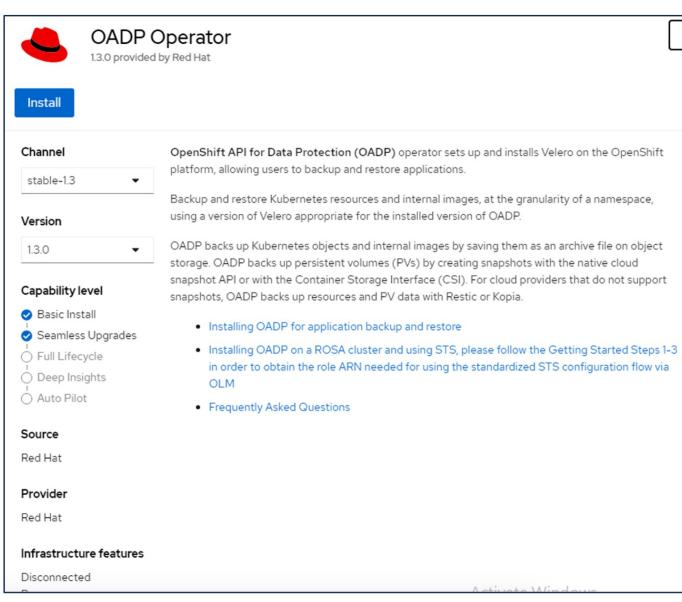


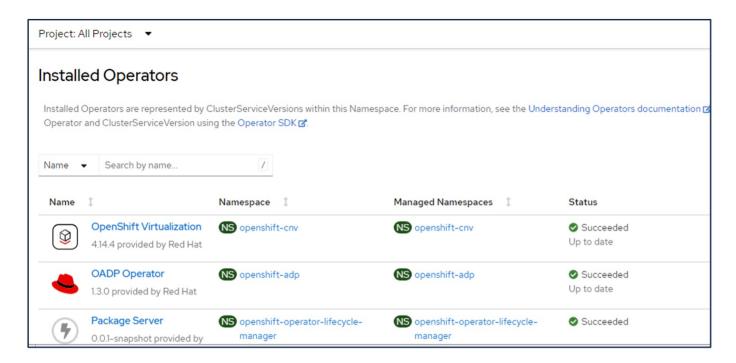
If you want to take a backup of a VM when it is in the Running state, then you must install the QEMU guest agent on that virtual machine. If you install the VM using an existing template, then QEMU agent is installed automatically. QEMU allows the guest agent to quiesce in-flight data in the guest OS 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 openshift-adp.







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) 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.

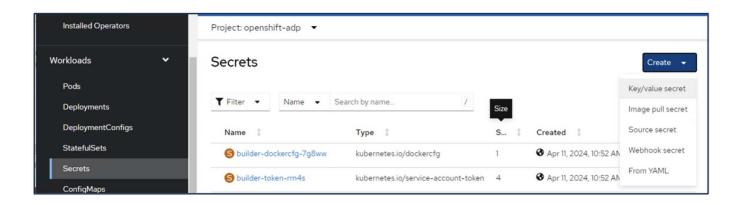
Prerequisites for Velero configuration with StorageGrid S3 details

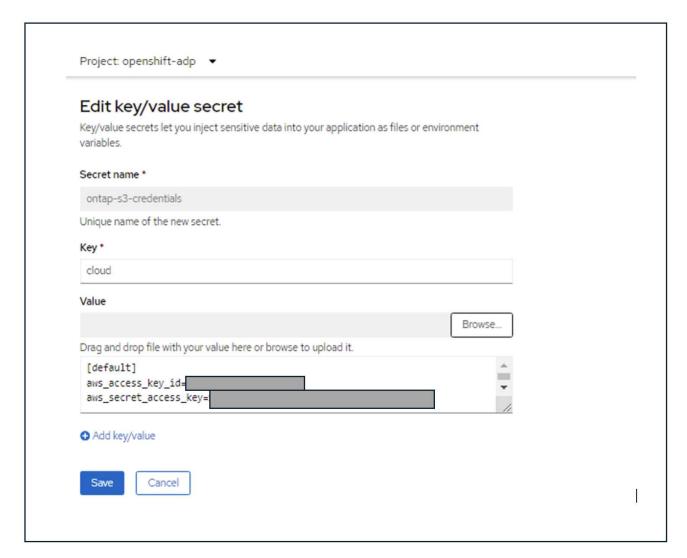
Velero can be configured to use S3 compatible Object Storage. You can configure StorageGrid S3 using the procedures shown in the StorageGrid documentation. You will need the following information from your StorageGrid S3 configuration to integrate with Velero.

- The endpoint 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 an ONTAP S3 user credential or StorageGrid Tenant 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. Name the secret appropriately. In the sample below, a secret with ONTAP S3 user
credentials named ontap-s3-credentials is created.





To create a secret named sg-s3-credentials from the CLI you can use the following command.

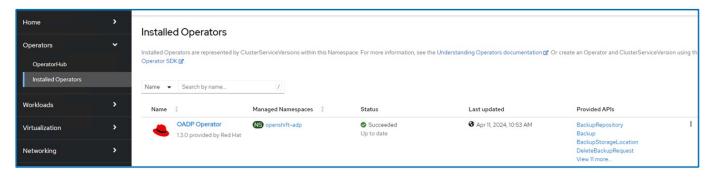
oc create secret generic sg-s3-credentials --namespace openshift-adp --from-file cloud=cloud-credentials.txt

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

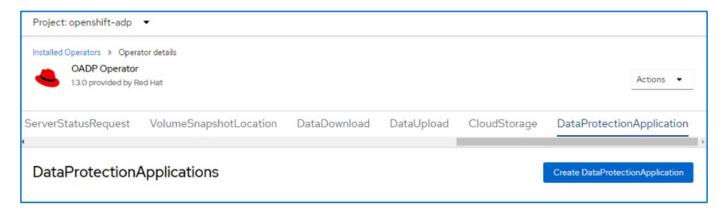
[default]

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 DataProtectionApplication or use the default name.



Now go to the YAML view and replace the spec information as shown in the yaml file examples below.

Sample yaml file for configuring Velero with ONTAP S3 as the backupLocation

```
spec:
 backupLocations:
    - velero:
        config:
          insecureSkipTLSVerify: 'false' ->use this for https
communication with ONTAP S3
          profile: default
          region: us-east-1
          s3ForcePathStyle: 'True' ->This allows use of IP in s3URL
          s3Url: 'https://10.xx.xx.xx' ->LIF to access S3. Ensure TLS
certificate for S3 is configured
        credential:
          key: cloud
          name: ontap-s3-credentials ->previously created secret
        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
```

Sample yaml file for configuring Velero with StorageGrid S3 as the backupLocation and snapshotLocation

```
spec:
 backupLocations:
    - velero:
        confiq:
          insecureSkipTLSVerify: 'true'
          profile: default
          region: us-east-1 ->region of your StorageGrid system
          s3ForcePathStyle: 'True'
          s3Url: 'https://172.21.254.25:10443' ->the IP used to access S3
        credential:
          key: cloud
          name: sq-s3-credentials ->secret created earlier
        default: true
        objectStorage:
          bucket: velero
          prefix: demobackup
        provider: aws
 configuration:
    nodeAgent:
      enable: true
      uploaderType: kopia
    velero:
      defaultPlugins:
        - csi
        - openshift
        - aws
        - kubevirt
```

The spec section in the yaml file should be configured appropriately for the following parameters similar to the example above

backupLocations

ONTAP S3 or StorageGrid S3 (with its credentials and other information as shown in the yaml) is configured as the default BackupLocation for velero.

snapshotLocations

If you use Container Storage Interface (CSI) snapshots, you do not need to specify a snapshot location because you will create a VolumeSnapshotClass CR to register the CSI driver. In our example, you use Astra Trident CSI and you have previously created VolumeSnapShotClass CR using the Trident CSI driver.

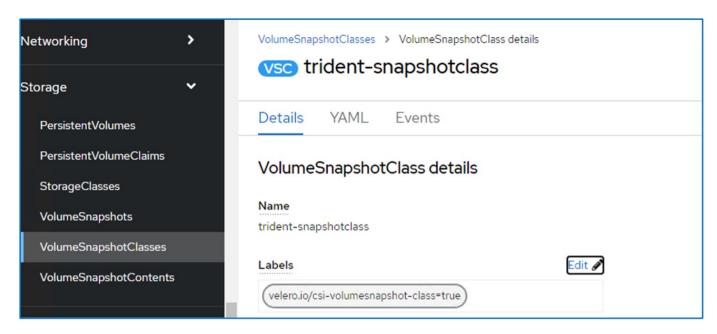
Enable CSI plugin

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

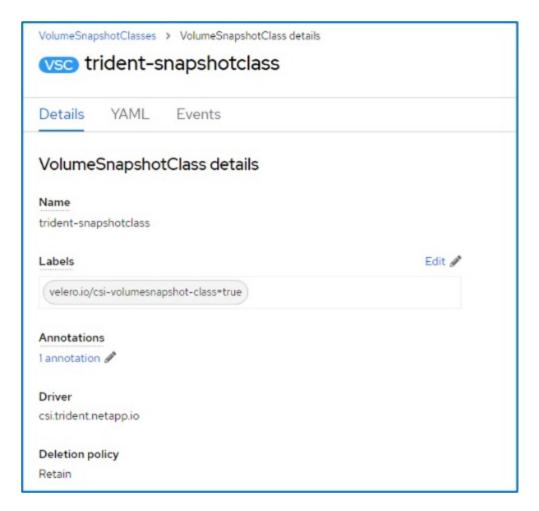
- You must have the trident VolumeSnapshotClass created.
- Edit the label of the trident-snapshotclass and set it to

velero.io/csi-volumesnapshot-class=true as shown below.

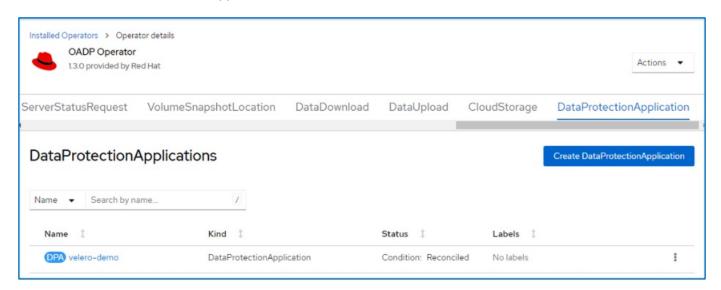


Ensure that the snapshots can persist even if the VolumeSnapshot objects are deleted. This can be done by setting the **deletionPolicy** to Retain. If not, deleting a namespace will completely lose all PVCs ever backed up in it.

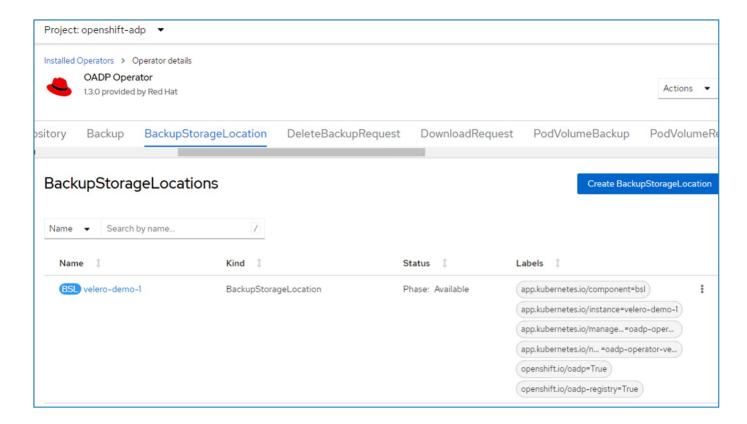
```
apiVersion: snapshot.storage.k8s.io/v1
kind: VolumeSnapshotClass
metadata:
   name: trident-snapshotclass
driver: csi.trident.netapp.io
deletionPolicy: Retain
```



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

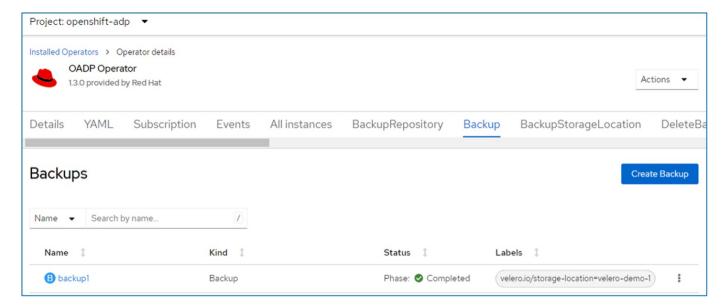
This section outlines how to create on-demand backup for VMs in OpenShift Virtualization.

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. A backup of the VM along with the snapshot of its disks are created and stored in the backup location specified in the yaml. The backup will remain in the system for 30 days as specified in the ttl.

Once the backup completes, its Phase will 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.



In StorageGrid, you can also use the S3 console that is available from the Tenant Manager to view the backup objects.

lame	Size	Type	Last Modified	Storage Class
backup1.tar.gz	230.36 KB	GZ File	4/15/2024 10:26:29 PM	STANDARD
velero-backup.json	3.35 KB	JSON File	4/15/2024 10:26:29 PM	STANDARD
backup1-resource-list.json.gz	1.12 KB	GZ File	4/15/2024 10:26:29 PM	STANDARD
backup1-itemoperations.json.gz	600 bytes	GZ File	4/15/2024 10:26:28 PM	STANDARD
backup1-volumesnapshots.json.gz	29 bytes	GZ File	4/15/2024 10:26:28 PM	STANDARD
backup1-podvolumebackups.json.gz	29 bytes	GZ File	4/15/2024 10:26:28 PM	STANDARD
backup1-results.gz	49 bytes	GZ File	4/15/2024 10:26:28 PM	STANDARD
backup1-csi-volumesnapshotclasses.json.gz	426 bytes	GZ File	4/15/2024 10:26:28 PM	STANDARD
backup1-csi-volumesnapshotcontents.json.gz	1.43 KB	GZ File	4/15/2024 10:26:28 PM	STANDARD
backup1-csi-volumesnapshots.json.gz	1.34 KB	GZ File	4/15/2024 10:26:28 PM	STANDARD
backup1-logs.gz	13.49 KB	GZ File	4/15/2024 10:26:28 PM	STANDARD

Creating scheduled backups for VMs in OpenShift Virtualization

To create backups on a schedule, you need to create a Schedule CR.

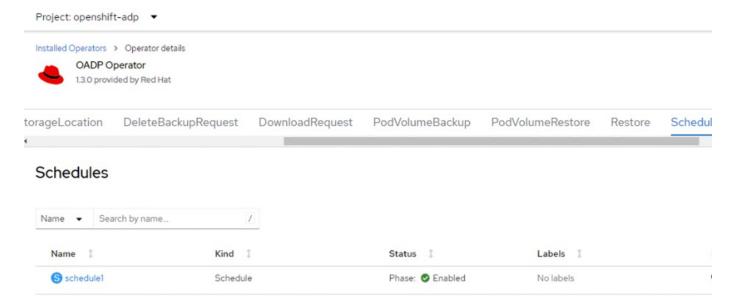
The schedule is simply a Cron expression allowing you to specify the time at which you want to create the backup. A sample yaml to create a Schedule CR.

```
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
```

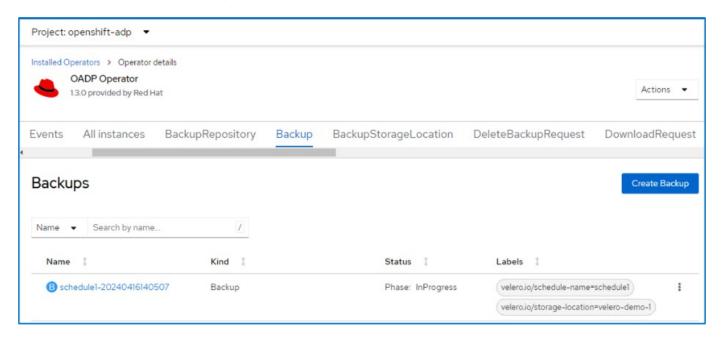
The Cron expression 0 7 * * * means a backup will be created at 7:00 every day.

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 at 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

This section describes how to restore virtual machine(s) from a backup.

Prerequisites

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

Restore to the same namespace

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:

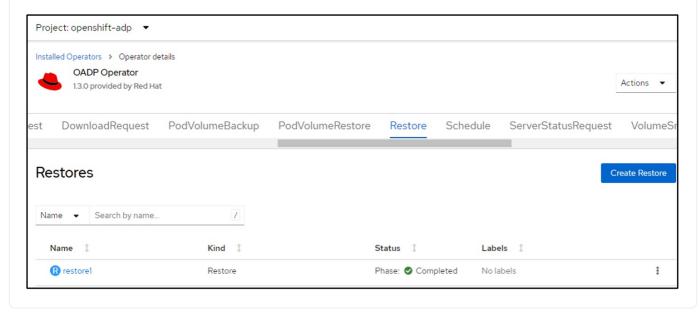
name: restore1

namespace: openshift-adp

spec:

backupName: backup1
restorePVs: true

When the phase shows completed, you can 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). The VM is restored to the same namespace.



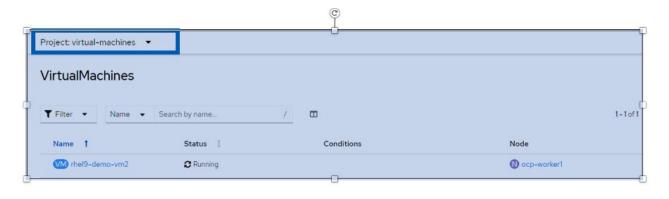
Restore to a different namespace

To restore the VM to a different namespace, you can provide a namespaceMapping in the yaml definition of the Restore CR.

The following sample yaml file creates a Restore CR to restore a VM and its disks in the virtual-machinesdemo namespace when the backup was taken to the virtual-machines namespace.

```
apiVersion: velero.io/v1
kind: Restore
metadata:
   name: restore-to-different-ns
   namespace: openshift-adp
spec:
   backupName: backup
   restorePVs: true
   includedNamespaces:
   - virtual-machines-demo
   namespaceMapping:
     virtual-machines-demo: virtual-machines
```

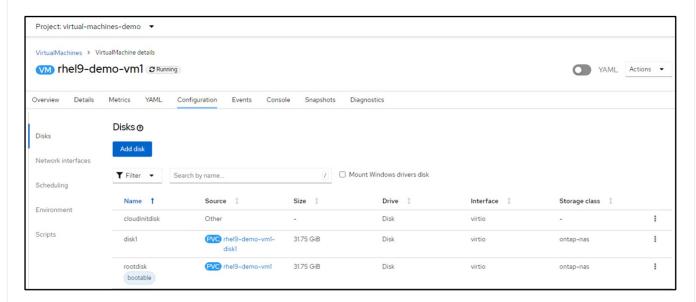
When the phase shows completed, you can see that the virtual machines have been restored to the state when the snapshot was taken. (If the backup was created when the VM was running, restoring the VM from the backup will start the restored VM and bring it to a running state). The VM is restored to a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

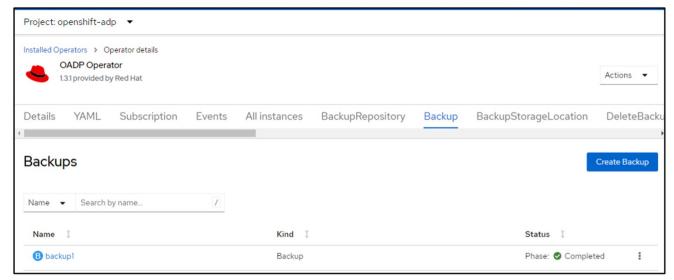


Restore to a different storage class

Velero provides a generic ability to modify the resources during restore by specifying json patches. The json patches are applied to the resources before they are restored. The json patches are specified in a configmap and the configmap is referenced in the restore command. This feature enables you to restore using different storage class.

In the example below, the virtual machine, during creation uses ontap-nas as the storage class for its disks. A backup of the virtual machine named backup1 is created.





Simulate a loss of the VM by deleting the VM.

To restore the VM using a different storage class, for example, ontap-nas-eco storage class, you need to do the following two steps:

Step 1

Create a config map (console) in the openshift-adp namespace as follows:

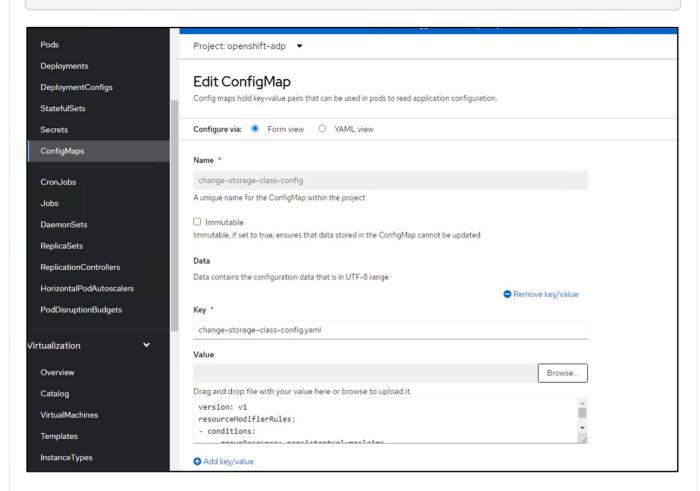
Fill in the details as shown in the screenshot:

Select namespace : openshift-adp

Name: change-storage-class-config (can be any name)

Key: change-storage-class-config.yaml:
Value:

version: v1
 resourceModifierRules:
 - conditions:
 groupResource: persistentvolumeclaims
 resourceNameRegex: "^rhel*"
 namespaces:
 - virtual-machines-demo
 patches:
 - operation: replace
 path: "/spec/storageClassName"
 value: "ontap-nas-eco"



The resulting config map object should look like this (CLI):

```
# kubectl describe cm/change-storage-class-config -n openshift-
adp
Name:
              change-storage-class-config
Namespace:
              openshift-adp
Labels:
              velero.io/change-storage-class=RestoreItemAction
              velero.io/plugin-config=
Annotations:
              <none>
Data
change-storage-class-config.yaml:
version: v1
resourceModifierRules:
- conditions:
     groupResource: persistentvolumeclaims
     resourceNameRegex: "^rhel*"
     namespaces:
     - virtual-machines-demo
  patches:
  - operation: replace
    path: "/spec/storageClassName"
    value: "ontap-nas-eco"
BinaryData
Events: <none>
```

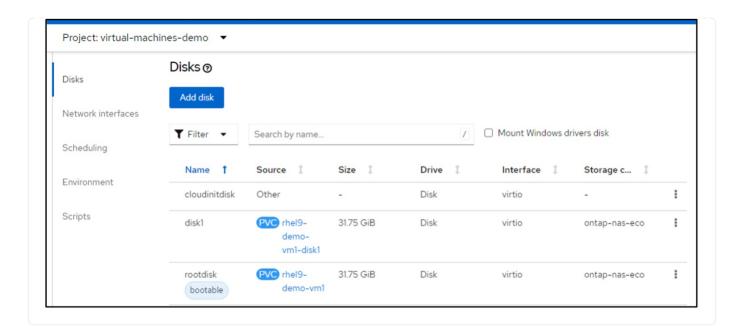
This config map will apply the resource modifier rule when the restore is created. A patch will be applied to replace the storage class name to ontap-nas-eco for all persistent volume claims starting with rhel.

Step 2

To restore the VM use the following command from the Velero CLI:

```
#velero restore create restorel --from-backup backup1 --resource
-modifier-configmap change-storage-class-config -n openshift-adp
```

The VM is restored in the same namespace with the disks created using the storage class ontap-nas-eco.



Deleting backups and restores in using Velero

This section outlines how to delete backups and restores for VMs in OpenShift Virtualization 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>
```

Deleting a Restore

You can 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>

Copyright information

Copyright © 2024 NetApp, Inc. All Rights Reserved. Printed in the U.S. No part of this document covered by copyright may be reproduced in any form or by any means—graphic, electronic, or mechanical, including photocopying, recording, taping, or storage in an electronic retrieval system—without prior written permission of the copyright owner.

Software derived from copyrighted NetApp material is subject to the following license and disclaimer:

THIS SOFTWARE IS PROVIDED BY NETAPP "AS IS" AND WITHOUT ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE, WHICH ARE HEREBY DISCLAIMED. IN NO EVENT SHALL NETAPP BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

NetApp reserves the right to change any products described herein at any time, and without notice. NetApp assumes no responsibility or liability arising from the use of products described herein, except as expressly agreed to in writing by NetApp. The use or purchase of this product does not convey a license under any patent rights, trademark rights, or any other intellectual property rights of NetApp.

The product described in this manual may be protected by one or more U.S. patents, foreign patents, or pending applications.

LIMITED RIGHTS LEGEND: Use, duplication, or disclosure by the government is subject to restrictions as set forth in subparagraph (b)(3) of the Rights in Technical Data -Noncommercial Items at DFARS 252.227-7013 (FEB 2014) and FAR 52.227-19 (DEC 2007).

Data contained herein pertains to a commercial product and/or commercial service (as defined in FAR 2.101) and is proprietary to NetApp, Inc. All NetApp technical data and computer software provided under this Agreement is commercial in nature and developed solely at private expense. The U.S. Government has a non-exclusive, non-transferrable, nonsublicensable, worldwide, limited irrevocable license to use the Data only in connection with and in support of the U.S. Government contract under which the Data was delivered. Except as provided herein, the Data may not be used, disclosed, reproduced, modified, performed, or displayed without the prior written approval of NetApp, Inc. United States Government license rights for the Department of Defense are limited to those rights identified in DFARS clause 252.227-7015(b) (FEB 2014).

Trademark information

NETAPP, the NETAPP logo, and the marks listed at http://www.netapp.com/TM are trademarks of NetApp, Inc. Other company and product names may be trademarks of their respective owners.