



HPE Solution Architecture for backup and recovery of Red Hat OpenShift Container Platform on HPE Synergy and HPE 3PAR Storage

Contents

Executive summary 3

Solution overview 3

 OpenShift master node 4

 OpenShift etcd node 5

 OpenShift node 5

 OpenShift infrastructure node 5

 Persistent storage 5

Solution hardware 6

 HPE StoreOnce 6

 HPE Recovery Manager Central software 7

Backup and recovery considerations 7

 Backup OpenShift Container Platform components 8

 Restore OCP components 9

Appendix A: Backup and restore OpenShift node components 10

 Master node backup 11

 Worker node backup 11

 Infra node backup 12

 etcd backup 12

 Restoring OCP components from a backup 13

 Restoring persistent volumes 14

Appendix B: HPE 3PAR StoreServ Volume Express Protection 17

 RMC Express Protect 17

Appendix C: HPE 3PAR StoreServ Storage replication protection 24

 Active/Passive replication (Disruptive) 24

 Peer Persistence replication (Non-Disruptive) 24

Change Tracker 25

Resources and additional links 26



Executive summary

In today's digital world, organizations are under increasing pressure to deliver applications faster while reducing costs. As these applications grow more complex, they put stress on IT infrastructure, teams, and processes. To remain competitive, organizations must quickly adapt and developers need to be more effective, efficient, and agile. Container technology provides the right application platform to help organizations become more responsive and iterate across multiple IT environments as well as develop, deploy, and manage applications faster. But implementing a containerized environment across existing infrastructure is a complex undertaking that can require weeks or months to mobilize, particularly for enterprises. To help accelerate container application delivery, Hewlett Packard Enterprise and Red Hat® are collaborating to optimize Red Hat OpenShift Container Platform (OCP) on HPE platforms, including HPE Synergy, the industry's first composable infrastructure, and HPE 3PAR StoreServ Storage.

Red Hat OCP on HPE Synergy provides an end-to-end, fully-integrated container solution which, once assembled, can be configured within hours. This eliminates the complexities associated with implementing a container platform across an enterprise data center and provides the automation of hardware and software configuration to quickly provision and deploy a containerized environment at scale. Red Hat OCP provides organizations with a reliable platform for deploying and scaling container-based applications. HPE Synergy Composable Infrastructure provides the flexibility customers need to run that container platform and dynamically provision and scale applications, whether they run in virtual machines or containers, hosted on-premises, in the cloud, or as a Hybrid cloud.

Containers have dramatically increased in popularity as organizations recognize the benefits with respect to both time and resource efficiency. This explosive growth of container applications overwhelms traditional data protection approaches. Applying traditional data protection strategies to containerized applications will simply not work. This paper highlights the importance of protecting each component of the OCP cluster in order to restore the data in case of corruption or system failures. It also addresses persistent volume backup using HPE Recovery Manager Central (RMC) software with the HPE 3PAR StoreServ snapshot feature and HPE StoreOnce.

Target audience: This document is intended for systems engineers, systems administrators, architects, and installers responsible for installing and maintaining Red Hat Enterprise Linux (RHEL) and Red Hat OCP, on a large scale, running on HPE Synergy Composable Infrastructure and HPE 3PAR StoreServ storage. The reader of this document should be familiar with RHEL, HPE Synergy Composable Infrastructure, HPE 3PAR StoreServ Storage, HPE StoreOnce, and general backup and recovery concepts and common practices.

Solution overview

An OpenShift cluster is made up of several nodes and each node type has different roles. To protect the environment, it is very important to understand how these components fit together and the services provided by each component. A successful backup and recovery solution is highly unlikely without this understanding in place. This section provides details of each OpenShift node and what components require protection within the environment.



For the purpose of making the infrastructure highly available, there are three (3) master nodes, three (3) infrastructure nodes, and three (3) etcd nodes. Worker nodes may be deployed in variable quantities. It is necessary to create a backup of the important components within the Red Hat OpenShift Container Platform (OCP) cluster in order to recreate the nodes in the event of a failure. Figure 1 describes the major components involved in the deployment of Red Hat OCP.

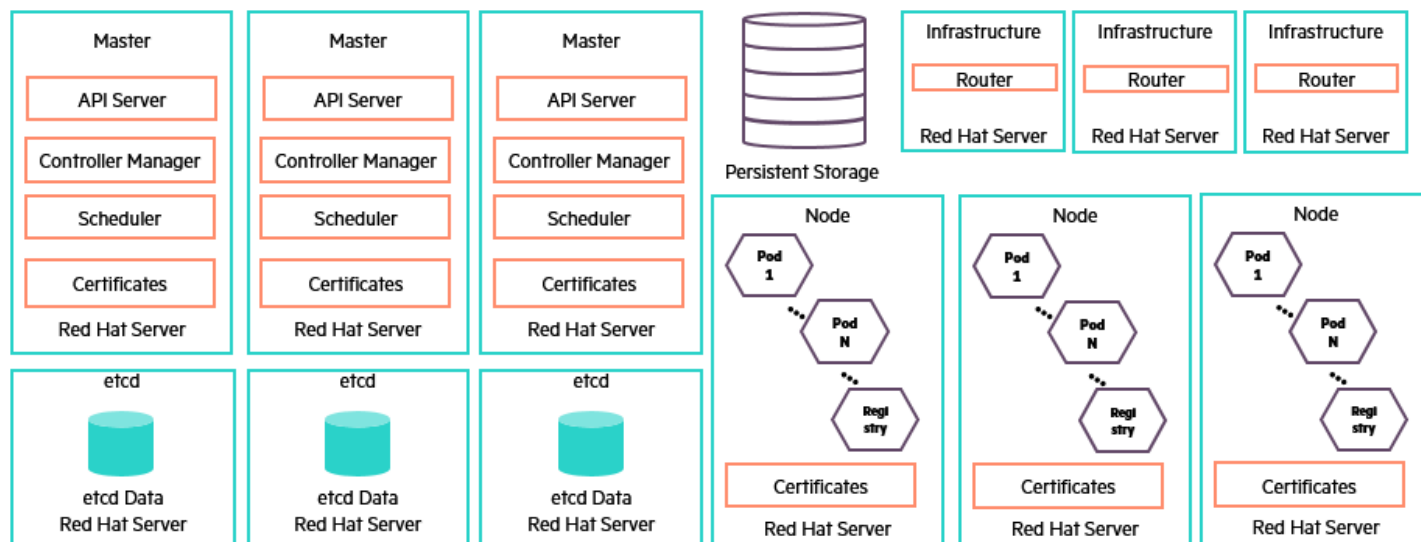


Figure 1. Solution diagram

OpenShift master node

The OpenShift master nodes are the brains of the cluster and are comprised of a set of core components including:

- API server
- Controller manager server
- Scheduler
- Certificates

The master nodes maintain the cluster's configuration, manage nodes in the OpenShift cluster, and schedule pods to run on nodes. If the OpenShift master nodes are unable to function, this will not impact the end users as the container application traffic will remain functional. However, administrators and users will not be able to make any new adjustments to the OpenShift cluster.

API server

The API server provides the management entry point of the OpenShift cluster. It mediates the interactions between the OpenShift master node components via RESTful API calls. It is responsible for storing API objects into the persistent etcd store. API server high availability is built on the persistent etcd store and deploys multiple instances of API server roles on the OpenShift cluster.

Controller manager

The controller manager monitors the state of the cluster through the API server watch feature. When a state change notification is received, the controller manager makes the necessary changes attempting to move the current state towards the desired state to keep the OpenShift cluster functioning correctly. Multiple controller manager roles are configured on OpenShift master nodes to provide high availability.

Scheduler

The scheduler ensures that container applications are scheduled to run on worker nodes within the OpenShift cluster. The scheduler reads data from the pod and attempts to find a node that is a good fit based on configured policies. To ensure high availability, more than one OpenShift master node must be configured for the scheduler roles.



Certificates

Certificates are used by the API server when securing inbound requests, authenticating users, making outbound requests, and for mutual TLS¹ between the API server and all other API objects in OpenShift Cluster. Certificates are copied to all the master nodes during the deployment. If more than one master host is deployed on an OpenShift cluster, the certificates are considered highly available.

OpenShift etcd node

etcd stores the persistent master state while other components watch etcd for changes to bring themselves into the desired state. It implements the key-value stores where all of the objects in OpenShift cluster master node components are stored. The etcd store implements a distributed consensus algorithm to ensure that even if one of the storage nodes fail, there is sufficient replication to maintain data availability. Optionally etcd role can be configured within in the master node itself.

OpenShift node

An OpenShift node or a worker node provides the runtime environment for containers. Each node in an OpenShift cluster has the required services to be managed by the master. The master uses information from nodes to validate nodes with health checks. A node is ignored until it passes the health checks, and the master continues checking nodes until they are valid. Other than running pods, worker nodes contain certificates, services, and authorization files. Large numbers of OpenShift nodes may be deployed in a cluster and if one node fails, it can be easily replaced without losing valuable data. However, certificates need to be deployed on the new node. Typically, certificates and authorization files are redeployed using Ansible playbooks and the Ansible Engine/Tower will hold the files. As a result, the Ansible Engine/Tower must be protected to ensure the file is highly available.

OpenShift makes use of its local Docker registry for storing Docker images. In a highly available deployment such as the one HPE has created at <https://github.com/HewlettPackard/hpe-solutions-openshift/tree/master/synergy/scalable>, the worker node nodes are responsible for hosting these registry pods and the accompanying local container images. Registry pods are assigned with a persistent volume (PV) from external storage. To protect the data, it is recommended to take a backup, snapshot or clone of the volume or replicate it to a disaster recovery site.

OpenShift infrastructure node

Pods inside an OpenShift cluster are only available via their IP addresses on the cluster network. An edge load balancer can be used to accept traffic from outside networks and route the traffic to pods inside the OpenShift cluster. An OpenShift administrator can deploy routers in an OpenShift cluster through infrastructure nodes. These enable routes created by developers to be used by external clients. OpenShift routers provide external hostname mapping and load balancing to services over protocols that pass information directly to the router. The hostname must be present in the protocol in order for the router to determine where to send it. For high availability, an external load balancer such as F5 BIG-IP can be used along with multiple infrastructure nodes.

Persistent storage

Containers were originally designed to run stateless applications, so there was no need for persistent storage. Once enterprises began adopting containers and they wanted to run stateful applications, persistent storage became necessary to meet the demands of the application data. HPE 3PAR StoreServ Storage provides persistent storage capabilities to an OpenShift cluster using plugins. Persistent storage and the data it houses need to be protected for business continuity, disaster recovery, and archival purpose. HPE 3PAR StoreServ integration with HPE Recovery Manager Central (RMC) and HPE StoreOnce provides snapshot protection and express protect² backup directly to HPE StoreOnce.

¹ Transport Layer Security

² RMC Express Protect Backups are faster than traditional ISV-based backups. Data is moved directly from primary storage to protection storage without any intermediate server bottlenecks.



Solution hardware

For this solution, OCP is built on an HPE Converged Architecture 750 which offers an improved time to deployment and tested firmware recipe. That baseline can be retrieved at the HPE Information Library. The user also has the flexibility of customizing the HPE components throughout this stack as per their unique IT and workload requirements or building with individual components. Figure 2 shows the physical configuration of the two racks used in this solution.

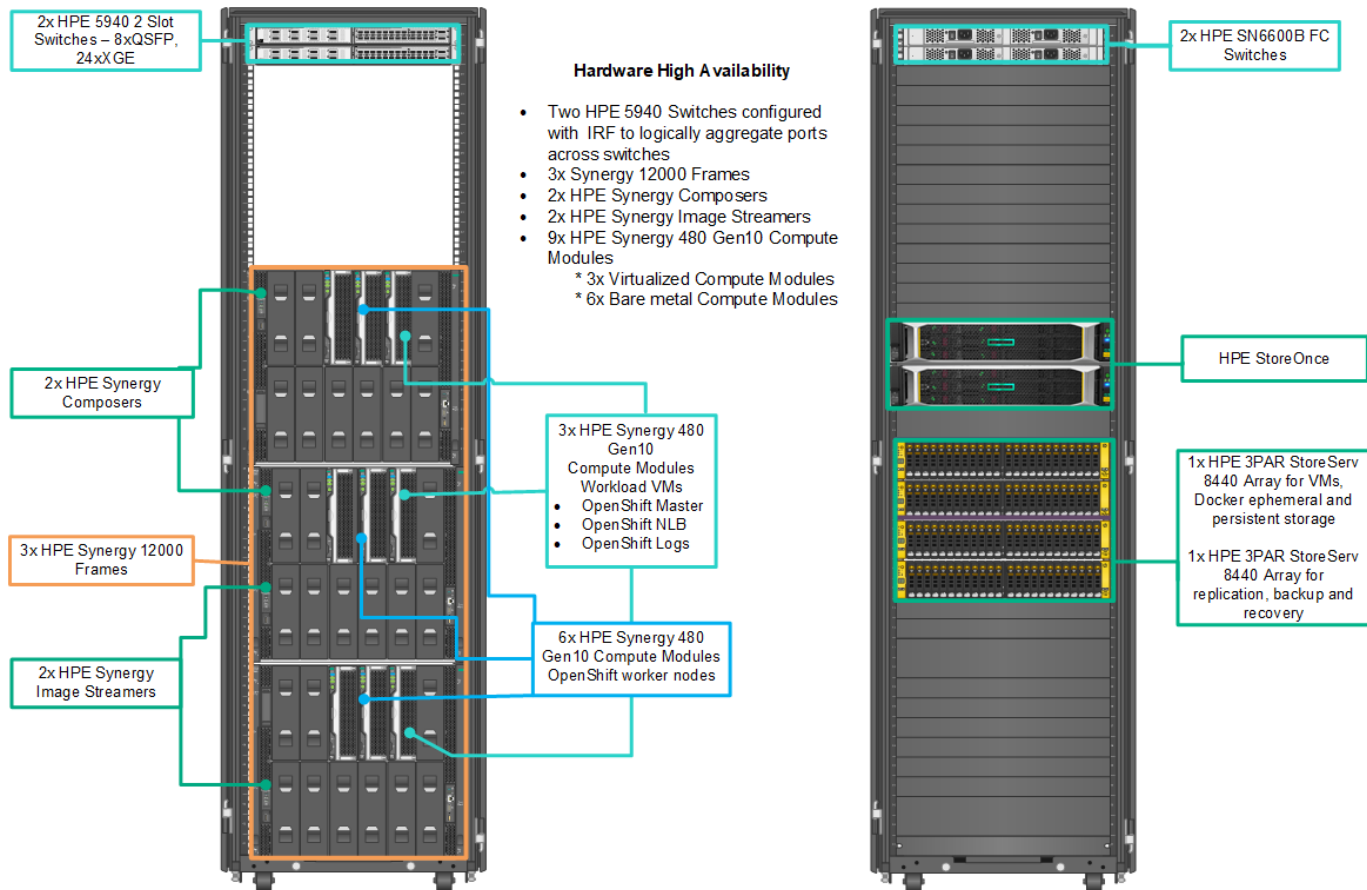


Figure 2. Physical layout of the compute within the solution, front view

For detailed configuration information including a list of components, software versions, and how the solution is built, refer to the Reference Configuration document at <https://h20195.www2.hpe.com/V2/GetDocument.aspx?docname=a00056102enw>.

HPE StoreOnce

HPE StoreOnce addresses the needs of customers ranging from entry level to large-scale enterprises. HPE StoreOnce systems deliver scale-out capacity and performance to keep pace with shrinking backup windows, reliable disaster recovery, simplified protection of remote offices, and rapid file restore to meet today's service level agreements (SLAs). HPE StoreOnce models vary by capacity and connectivity protocol. It is possible to start with a single HPE StoreOnce base unit or Virtual Server Appliance (VSA) and then expand with additional units and expansion shelves.

The HPE StoreOnce VSA extends the deployment options for HPE StoreOnce with the agility and flexibility of a virtual appliance, thus removing the need to install dedicated data protection hardware. All the features of the purpose-built StoreOnce Systems are available in a software defined backup target with up to 500 TB of usable capacity. This provides a flexible and cost-effective backup target for virtualized server environments as part of a pure software defined data protection solution or in conjunction with StoreOnce purpose-built appliances for mixed environments.



HPE StoreOnce Catalyst is a data protection optimized interface unique to HPE StoreOnce systems. It provides higher performance and more flexible control than traditional emulated tape virtual tape libraries (VTL) targets or network attached storage (NAS) shares.

Cloud Bank Storage is an extension to HPE StoreOnce Catalyst that combines the low cost of object storage with the storage efficiency of StoreOnce deduplication. HPE StoreOnce connects to external object storage to provide capacity for the Catalyst Cloud Bank store. Using external storage, in addition to the local system storage, can triple the effective capacity of the StoreOnce system. Through cloud optimized data transfer and storage, Cloud Bank Storage minimizes cloud transfer and storage costs.

For more information on HPE StoreOnce systems, refer <https://h20195.www2.hpe.com/v2/gethtml.aspx?docname=c04328820>.

HPE Recovery Manager Central software

HPE RMC software facilitates policy-driven, converged data protection, and copy data management for business-critical applications. RMC integrates HPE 3PAR StoreServ Storage and HPE Nimble All-Flash Arrays with HPE StoreOnce systems, leveraging snapshot performance with storage-integrated backups to deliver flash speed application protection and copy data management with less cost and complexity than legacy solutions. RMC is also built for cloud, allowing users to leverage public cloud for cost-effective, long-term retention of user's backups.

For more information on HPE RMC, refer https://support.hpe.com/hpsc/doc/public/display?docId=emr_na-c04552659.

Backup and recovery considerations

A container application data protection architecture should be defined by business requirements. These requirements include factors such as the speed of recovery, the maximum permissible data loss, and data retention needs. The data protection plan must also take into consideration various regulatory requirements for data retention and restoration. Finally, different data recovery scenarios must be considered, ranging from the typical and foreseeable recovery resulting from user or application errors to disaster recovery scenarios that include the complete loss of a datacenter.

Small changes in data protection and recovery policies can have a significant impact on the overall architecture of container infrastructure. It is critical to define and document standards before starting the design work to avoid complicating a data protection architecture. Unnecessary features or levels of protection lead to unnecessary costs and management overhead, and an initially overlooked requirement can lead a project in the wrong direction or require last-minute design changes.

Two of the important parameters that define a backup and recovery plan are recovery point objective (RPO) and recovery time objective (RTO):

- RPO refers to the maximum amount of data you can afford to lose. It is measured backward in time from a failure to the last available backup.
- RTO refers to the maximum downtime it takes to restore and recover an application after a disaster or corruption. It is stated in terms of time.

RPO and RTO requirements may vary based on the business needs.

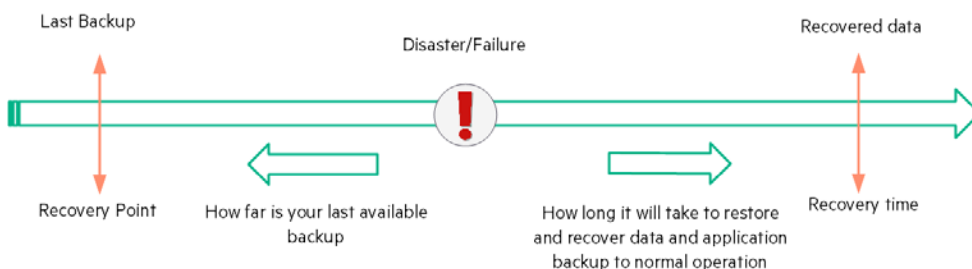


Figure 3. RPO/RTO description

Backup is defined as the procedure to create extra copies of data in case the original data is lost or corrupted. This helps to restore the data to a newly-created environment or to relevant components in case of a disaster or corruption.

For the purpose of this solution, HPE recommends the use of HPE StoreOnce as a backup target for backing up OCP components including persistent volumes. For the backup of the configuration files, create an HPE StoreOnce NAS share and export it to the OCP nodes. For high availability, the NAS share can be replicated to a remote HPE StoreOnce.

For persistent volume protection, use RMC to initiate a crash consistent snapshot at the volume level, and using the RMC express protect feature, move the snapshot to the HPE StoreOnce Catalyst store. As it is moved to a backup appliance, data can be stored for archival purposes as well. In this scenario, there is no external data mover involved. Either HPE StoreOnce or RMC acts as the data mover. This reduces the cost and complexity of the solution. HPE 3PAR StoreServ Storage also supports replication of the volume to a remote array which can be used to reduce the RPO/RTO requirements. RPO/RTO can be further reduced with peer persistent (active/active) replication³.

Backup OpenShift Container Platform components

This section disseminates the components that need to be protected or backed up within a Red Hat OCP cluster. Hewlett Packard Enterprise plans to update this section with new information over time.

A backup should be taken before any upgrade or modifications. It is recommended to perform periodic backup to make sure you have the most recent configuration available in the event of a failure. An OCP backup involves taking backup of current state to external storage at the cluster level. This means creating individual backups of the following components:

- Master node components
- Worker node components
- Infrastructure node components
- etcd data and configurations
- Persistent Storage

In Red Hat OpenShift, all of the components are treated as objects and are stored in files. This means that creating a configuration backup is equivalent to taking file level backups.

Master node

The master node is responsible for maintaining the desired state of a cluster. It is recommended to perform a master node backup before you make any modifications to the OCP infrastructure. In high availability environments, make sure to perform the backup on all master nodes.

Master node components are stored under the `/etc/origin/` and `/etc/sysconfig/` directories. Refer to [Appendix A](#) for more details.

Worker node

The nature of worker nodes is that any specific configuration pertaining to running pods are replicated over the nodes in case of a failover. They typically do not contain data that is necessary to run an environment. In an HPE 3PAR StoreServ Storage OCP deployment, the worker node is responsible for hosting the registry pods. These pods are deployed with a volume from the HPE 3PAR StoreServ Storage volume.

Apart from running pods, worker nodes contain certificates, which are generated during installation, services, authorization files and more. These files are stored under the `/etc/origin/` and `/etc/sysconfig/` directories. Please refer to [Appendix A](#) for more details.

Infrastructure node

The infrastructure node is responsible for hosting the router pods. These pods are replicated over other nodes in case of a failover. They typically do not contain data that is necessary to run the environment.

Registry certificates must be backed up from the `/etc/docker/certs.d` directory. Refer to [Appendix A](#) for more details.

etcd

OpenShift uses etcd to store system configuration and state as well as metadata. etcd is a key value store for all the object definitions. It allows nodes in the cluster to read and write data.

When you back up etcd, you need to take a backup of the configuration and data.

³ <https://www.hpe.com/in/en/product-catalog/storage/storage-software/pip.hpe-3par-peer-persistence-software.5335710.html>



etcd configurations are stored in the `/etc/etcd` directory where etcd instances are running. Unlike other configurations, etcd configurations are unique across etcd instance. etcd data can be backed up using the `etcd snapshot save` command and by copying the `/var/lib/etcd/member/snap/db` file to a desired location. Refer to [Appendix A](#) for more details.

Persistent storage

Containers were designed to run stateless applications. In the beginning, there was no need for persistent storage, but then enterprises started adopting containers and they wanted to run their applications on stateful containers. Once persistent data is present, a need is created for persistent storage. Backing up and protecting this data becomes very important.

For persistent volume level backups, traditional agent-based backup software won't work natively with a container orchestrator such as Red Hat OpenShift. Backup schemes need to be consumed as a data protection service from the underlying container-aware storage such as HPE 3PAR StoreServ Storage. Using HPE RMC Express Protect Backup, we can take a crash-consistent snapshot-based persistent volume backup directly to HPE StoreOnce.

Using the HPE 3PAR Docker Storage plugin feature, you can create replicated volumes. The HPE 3PAR StoreServ Storage plugin extends Docker's `volume create` command interface via optional parameters in order to make it possible. The HPE 3PAR StoreServ Storage plugin assumes that an already working 3PAR Remote Copy setup is present. The plugin has to be configured with the details of this setup.

Refer to [Appendix A](#) and [Appendix B](#) for more details.

Restore OCP components

It is important to restore the OCP components in case of a system failure or corruption and to ensure the nodes are in a previous working state.

Master node

Restore means recreating the components from the point in time the backup is available. In the case where a master host is corrupted or failed due to system error, reinstall the master host, copy the important configuration files, and then restart the OCP services.

If you are restoring to a master, which is behind a highly available load balancer pool, restarting OCP service may cause downtime. Make sure you remove the master from the pool, restart the service, and then add it back to the load balancer pool.

If you are recreating a master after the system failure, apply the backup, reboot, and then add the master to the cluster. Refer to [Appendix A](#) for more details.

Worker node

In case a worker node host is corrupted or failed due to a system error, reinstall the worker node the way you did initially, copy the important configuration files, and then restart the OCP services.

If you are recreating a worker node after the system failure, apply the backup, reboot, and then add the worker to the cluster. Refer to [Appendix A](#) for more details.

Infrastructure node

If an infrastructure node is down due to a system error, reinstall the infrastructure node and reapply the certificates.

etcd

If the etcd configuration is corrupted or lost, restore the `/etc/etcd/etcd.conf` file from the backup and restart the service.

If the etcd data is corrupted and you want to restore from the snapshot, this can be performed on a single etcd node. After that, add the rest of the etcd nodes to the cluster. Please refer to [Appendix A](#) for more details.

Persistent storage

When utilizing the native storage capability to take crash consistent snapshots of a persistent volume, that volume can be restored from a point in time snapshot in case the PV is corrupted. This solution makes use of RMC and HPE 3PAR StoreServ Storage's snapshot capability to take crash consistent snapshots of persistent volumes. These volumes can be restored to a point in time either from the array snapshot or from the Express Protect Backup available on the HPE StoreOnce in the event that the PV is corrupted. Even if the persistent volume is deleted at the container level, the volume will still be available to restore from RMC. Please refer to [Appendix B](#) for more details.



Appendix A: Backup and restore OpenShift node components

To protect the OCP nodes, it is recommended to take a backup of any important configuration files using HPE StoreOnce Storage. Before a backup of the configuration files is initiated, a NAS share dedicated for backup targets, as explained in the following section, must be created:

- 1. Log in to the HPE StoreOnce Administration web GUI and navigate to **Data Services > NAS Shares** and then select **Create Share** as in Figure A1.

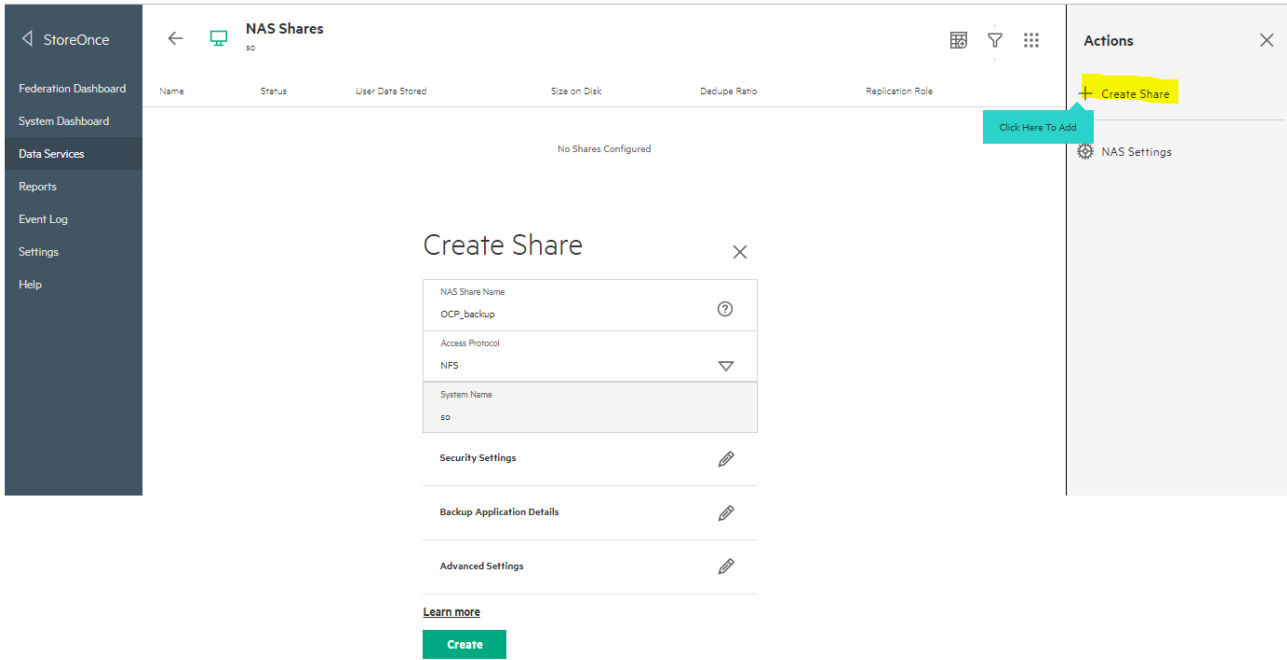


Figure A1. Creating a new NAS share in HPE StoreOnce

- 2. Create a name for the share and then select the access protocol by selecting **NFS**. Select the application details (if any), select the data type as file, and then click **Create** to finalize.
- 3. To mount the NAS share, provide access to the servers as in Figure A2.

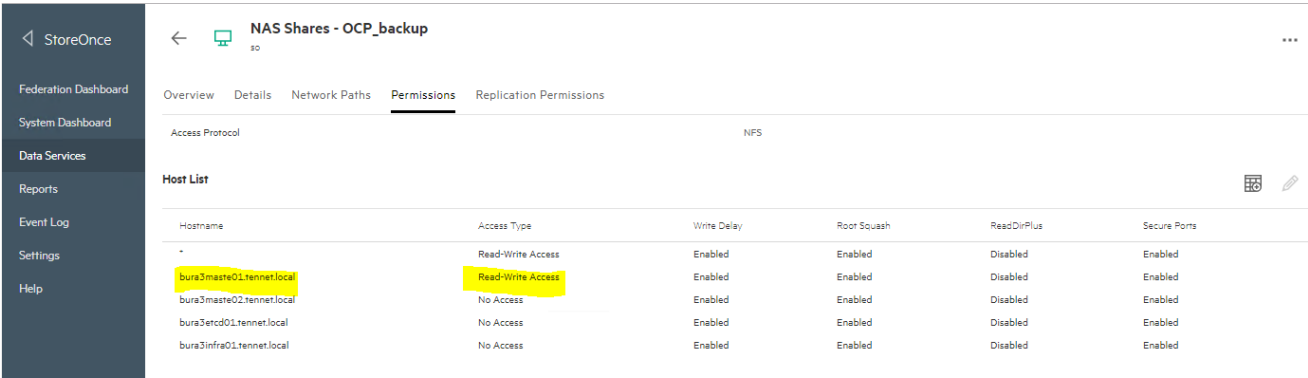


Figure A2. HPE StoreOnce NAS share access details

- 4. Mount the NAS share on the master server and trigger the backup with the following command:

```
# mount 10.0.20.xx:/nas/OCP_backup /StoreOnce_vol
```



Master node backup

To create a backup of the important configuration files on the master node, do the following:

1. Run the following commands using the mount point:

```
# MYBACKUPDIR=/backup/${hostname}/${date +%Y%m%d}
# sudo mkdir -p ${MYBACKUPDIR}/etc/sysconfig
# sudo cp -aR /etc/origin ${MYBACKUPDIR}/etc
# sudo cp -aR /etc/sysconfig/ ${MYBACKUPDIR}/etc/sysconfig/
# sudo cp -aR /etc/sysconfig/{iptables,docker-*} ${MYBACKUPDIR}/etc/sysconfig/
# sudo cp -aR /etc/dnsmasq* /etc/cni ${MYBACKUPDIR}/etc/
# rpm -qa | sort | sudo tee $MYBACKUPDIR/packages.txt
# cp -aR /etc/docker/certs.d/ ${MYBACKUPDIR}/docker-registry-certs-${hostname}
# tar -zcvf /backup/${hostname}-${date +%Y%m%d}.tar.gz $MYBACKUPDIR
```

2. Copy this tar file to the HPE StoreOnce NAS share mount point at the master server by running the following command:

```
# scp -aR /backup/${hostname}-${date +%Y%m%d}.tar.gz <user>@<master>:/StoreOnce_vol
```

Worker node backup

To create a backup of the important configuration files on the worker nodes, do the following:

1. Run the following commands:

```
# MYBACKUPDIR=/backup/${hostname}/${date +%Y%m%d}
# sudo mkdir -p ${MYBACKUPDIR}/etc/sysconfig
# sudo cp -aR /etc/origin ${MYBACKUPDIR}/etc
# sudo cp -aR /etc/sysconfig/atomic-openshift-node ${MYBACKUPDIR}/etc/sysconfig/
# sudo mkdir -p ${MYBACKUPDIR}/etc/sysconfig
# sudo cp -aR /etc/sysconfig/{iptables,docker-*} ${MYBACKUPDIR}/etc/sysconfig/
# sudo cp -aR /etc/dnsmasq* /etc/cni ${MYBACKUPDIR}/etc/
# rpm -qa | sort | sudo tee $MYBACKUPDIR/packages.txt
# cp -aR /etc/docker/certs.d/ ${MYBACKUPDIR}/docker-registry-certs-${hostname}
# tar -zcvf /backup/${hostname}-${date +%Y%m%d}.tar.gz $MYBACKUPDIR
```

2. Copy the tar file to the HPE StoreOnce NAS share mount point at the master server with the following command:

```
# sudo scp -aR /backup/${hostname}-${date +%Y%m%d}.tar.gz <user>@<master>:/StoreOnce_vol
```

3. Use the following Ansible play to take backups of master and worker nodes and move the data to the HPE StoreOnce NAS share. To retrieve the play, run the following commands:

```
# mkdir ~/git
# cd ~/git
# git clone https://github.hpe.com/Solutions/Openshift-Synergy-RA.git
```



```
# cd /Openshift-Synergy-RA/synergy/scalable/bura
# ansible-playbook -i hosts site.yaml
```

Infra node backup

To create a backup of certificate files and copy it to the appropriate mount point, do the following:

1. Run the following commands:

```
# MYBACKUPDIR=/backup/${hostname}/${date +%Y%m%d}
# mkdir -p ${MYBACKUPDIR}/etc/docker
# cp -R /etc/docker/certs.d ${MYBACKUPDIR}/etc/docker/certs-${date +%Y%m%d}
# tar -zcvf /backup/${hostname}-${date +%Y%m%d}.tar.gz $MYBACKUPDIR
```

2. Copy the tar file to the HPE StoreOnce NAS share mount point at the master server using the following command:

```
# scp -aR /backup/${hostname}-${date +%Y%m%d}.tar.gz <user>@<master>:/StoreOnce_vol
```

etcd backup

With this backup solution, etcd is running on a separate host and not as a static pod on the master. As a result, the etcd backup process is comprised of two different procedures, etcd configuration backup, including the required etcd configuration and certificates, and etcd data backup.

etcd configuration backup

The etcd configuration files to be preserved are all stored in the `/etc/etcd` directory of the instances where etcd is running. This includes the etcd configuration file (`/etc/etcd/etcd.conf`) and the required certificates for cluster communication. Make a backup of the configurations from all etcd members of the cluster.

```
# MYBACKUPDIR=/backup/${hostname}/${date +%Y%m%d}
# mkdir -p ${MYBACKUPDIR}/etcd-config-${date +%Y%m%d}
# cp -R /etc/etcd/ ${MYBACKUPDIR}/etcd-config-${date +%Y%m%d}
```

etcd data backup

Before taking a backup, ensure that the OpenShift Container Platform API service is running, connectivity with the etcd cluster (port 2379/tcp) is working, and proper certificates are available to connect to the etcd cluster, by running the following commands:

1. To cross-check the above services, run the following command, using etcd API version V3:

```
# systemctl show etcd --property=ActiveState,SubState
# etcdctl -C https://xx.0.62.xx:2379,https://xx.0.62.xx:2379,https://xx.0.62.2xx:2379 --ca-file=/etc/etcd/ca.crt --cert-file=/etc/etcd/peer.crt --key-file=/etc/etcd/peer.key cluster-health
# etcdctl -C https://xx.0.62.xx:2379,https://xx.0.62.xx:2379,https://xx.0.62.xx:2379 --ca-file=/etc/etcd/ca.crt --cert-file=/etc/etcd/peer.crt --key-file=/etc/etcd/peer.key member list
```

2. With the etcd v3 API, a snapshot from a live member can be taken with the `etcdctl snapshot` command and saved to an external storage:

```
# MYBACKUPDIR=/backup/${hostname}/${date +%Y%m%d}
# mkdir -p ${MYBACKUPDIR}/etcd-data-${date +%Y%m%d}
# ETCDCTL_API=3 etcdctl snapshot save ${MYBACKUPDIR}/etcd-data-${date +%Y%m%d}/snapshot.db --endpoints=https://xx.0.62.xx:2379 --cacert=/etc/etcd/ca.crt --cert=/etc/etcd/server.crt --key=/etc/etcd/server.key
```



3. Create the etcd data backup and copy the etcd db file:

```
# etcdctl backup --data-dir /var/lib/etcd --backup-dir ${MYBACKUPDIR}/etcd-data-$(date +%Y%m%d)
# tar -zcvf /backup/${hostname}-${date +%Y%m%d}.tar.gz $MYBACKUPDIR
```

4. Copy the tar file to the HPE StoreOnce NAS share mount point on the master server:

```
# scp -aR /backup/${hostname}-${date +%Y%m%d}.tar.gz <user>@<master>:/StoreOnce_vol
```

Restoring OCP components from a backup

1. For restoring the master node or only restoring certain files, mount the NAS backup share from the HPE StoreOnce to the node and copy the required files to the desired location. When complete, restart the services. An example appears below:

```
# mount 10.0.20.xx:/nas/OCP_backup /StoreOnce_vol
# bzip2 -d /StoreOnce_vol/${hostname}-${date +%Y%m%d}.tar.bz2
# tar -xvf /StoreOnce_vol/${hostname}-${date +%Y%m%d}.tar
# cp /StoreOnce_vol/${hostname}/${date +%Y%m%d}/origin/master/master-config.yaml
/etc/origin/master/master-config.yaml
# systemctl restart atomic-openshift-master-api
# systemctl restart atomic-openshift-master-controllers
```

Note

Restart the server for replacing the IP tables, if required.

2. For restoring the worker node or only restoring certain files, mount the NAS backup share from HPE StoreOnce to the node and copy the files to the desired location and restart the service:

```
# mount 10.0.20.xx:/nas/OCP_backup /StoreOnce_vol
# bzip2 -d /StoreOnce_vol/${hostname}-${date +%Y%m%d}.tar.bz2
# tar -xvf /StoreOnce_vol/${hostname}-${date +%Y%m%d}.tar
# cp /StoreOnce_vol/${hostname}/${date +%Y%m%d}/etc/origin/node/node-config.yaml
/etc/origin/node/node-config.yaml
# systemctl restart atomic-openshift-node
```

Note

If you are restoring files to a running worker node, restarting services may cause downtime.

3. If an etcd node is corrupted, replace the `/etc/etcd/etcd.conf` file by restoring it from the NAS share mounted from HPE StoreOnce and restart the service:

```
# mount 10.0.20.xx:/nas/OCP_backup /StoreOnce_vol
# gunzip /StoreOnce_vol/${hostname}-${date +%Y%m%d}.tar.gz
# tar -xvf /StoreOnce_vol/${hostname}-${date +%Y%m%d}.tar
```



```
# cp /StoreOnce_vol/backup/${hostname}/${date +%Y%m%d}/etcd-config-${date +%Y%m%d}/etcd.conf
/etc/etcd/etcd.conf

# restorecon -Rv /etc/etcd/etcd.conf

# systemctl restart etcd.service
```

4. For restoring etcd v3 data, perform the following commands:

```
# systemctl stop etcd.service

# rm -rf /var/lib/etcd

# ETCDCCTL_API=3 etcdctl snapshot restore /etcdsnap/snapshot.db --data-dir /var/lib/etcd --
endpoints=https://10.0.62.24:2379 --cacert=/etc/etcd/ca.crt --cert=/etc/etcd/server.crt --
key=/etc/etcd/server.key

# chown -R etcd:etcd /var/lib/etcd/

# restorecon -Rv /var/lib/etcd

# systemctl start etcd
```

5. When complete, check the cluster health.

Restoring persistent volumes

To restore the persistent volume (PV) to a point in time, perform the following tasks:

1. Log in to the Recovery Manager Central (RMC) portal.
2. Navigate to the **Recovery Manager Central** drop-down list, and select **Volumes > Volume Set > Clone / Restore** as in Figure A3.

The screenshot displays the HPE Recovery Manager Central (RMC) portal. The top navigation bar includes the 'HPE Recovery Manager Central' logo, a search bar, and user profile icons. Below the navigation bar, the 'Volumes' section is active, showing a list of volume sets. The 'test' volume set is selected, and its details are shown on the right. The 'test' volume set is associated with the 'sstor01-020P' storage system and has a 'Copy Policy' of '-'. The 'Job Status' is 'Last 24 Hrs.' and the 'Snapshot' and 'Express Protect' counts are both 0, indicating they have failed out of 1. The 'Copies' section shows a table with columns for 'Copy Type', 'Snapshot', and 'Express Protect'. The 'Snapshot' row shows 1 successful copy, 0 failed copies, and 0 pending copies. The 'Express Protect' row shows 1 successful copy, 0 failed copies, and 0 pending copies. A 'Clone / Restore' button is visible in the bottom right corner.

Copy Type	Snapshot	Express Protect
Snapshot	1	0
Express Protect	1	0

Figure A3. HPE RMC clone/restore option for the volume set

3. Once you select the **Clone / Restore** option, a window is displayed. Select **Restore**. Before you proceed with this step, make sure the replica count is set to zero using the deployment API object in the OCP console.

Note

Setting the replica count to zero for the pods causes downtime for the application.

4. At the OCP console, navigate to **Project**, select **Deployment Configuration**, select the down arrow on the right side to scale down the replica count, and accept the scale down confirmation as in Figure A4.

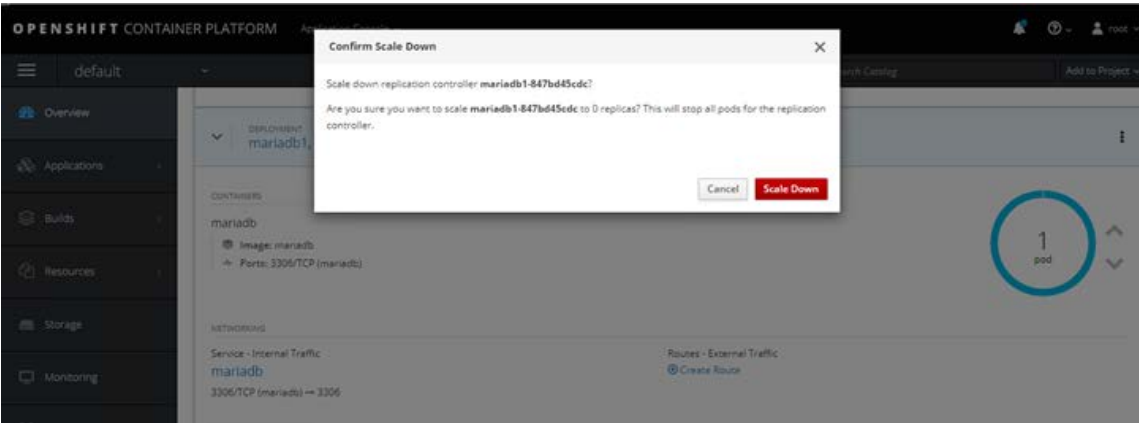


Figure A4. RH-OCP console scaling down the replica set to zero

5. Proceed with the **Restore** from the RMC GUI using the console as in Figure A5.

Clone / Restore 3par-expressprotect

Virtual Clone from a new Snapshot

This selection makes it possible to create a new snapshot of volume(s) in the current state. The clone of the volume(s) is created from the newly created snapshot.

Virtual Clone from an existing Copy

This selection makes it possible to choose an existing RMC Copy and then create a clone of the volume(s).

Physical Clone

This selection makes it possible to restore data from an Express Protect or a Catalyst Copy to a different set of volumes.

Restore

Restores data back to the parent volume. The existing data will be lost and it will be overwritten with the data from the chosen RMC Copy.

Figure A5. HPE RMC Clone/Restore option for the volume set



6. Select the date/time and the backup either from snapshot or from the Express Protect Backup and click **Next** as in Figure A6.

Restore 3par-expressprotect?

26	27	28	29	30	31	1
2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30	1	2	3	4	5	6

Available Copies

Search:

	Point in Time	Target Device	Copy Type	Name
<input type="radio"/>	2019-06-06 09:52:52	sstor01-020P	Snapshot	3par-expressprotect_1559832623
<input type="radio"/>	2019-06-06 09:52:52	so	Express Protect	3par-expressprotect_1559832630

Back

Next

Cancel

Figure A6. HPE RMC Restore available Copies with selected date



7. Select the volume to be restored and click **Restore** as in Figure A7.

Restore test?

Select Volume(s)

<input type="checkbox"/>	▲	Restore From	size(MiB)	Restore To	WWN
<input checked="" type="checkbox"/>		-	16384	dcv-dk7F9PdWS3uqVQiX0qftKw	60002AC00000000000000035A00018EEC
<input checked="" type="checkbox"/>		-	81920	Dockerlocal03.r	60002AC00000000000000002DD00018EEC
<input type="checkbox"/>		-	16384	dcv-WayFggaSIGkSmndgFMDUw	60002AC000000000000000031900018EEC
<input type="checkbox"/>		-	102400	express-protect-clone	60002AC00000000000000001170001D8E7
<input type="checkbox"/>		-	3072	dcv-eWYo-S0qTnqS4qBgIRqhKg	60002AC00000000000000002EF00018EEC

Back

Restore

Cancel

Figure A7. HPE RMC clone/restore option for the volume set

This operation overwrites the data on the volume being restored. After the restore is completed, change the replica count to 1 and check the PV and the data inside the pod.

Appendix B: HPE 3PAR StoreServ Volume Express Protection

Array-based snapshots and replication provide fast, non-disruptive point-in-time copies of your data. But snapshots alone cannot deliver comprehensive backup as they have retention limitations and a dependency on the underlying storage system. Simply put, your snapshots will be lost if the storage system fails.

The RMC Express Protect feature allows you to back up snapshots directly from HPE 3PAR StoreServ Storage to HPE StoreOnce. In this case, either the RMC appliance or the HPE StoreOnce itself will act as a data mover, reducing the cost and complexity of the architecture.

RMC leverages the snapshot differential technology in HPE 3PAR StoreServ Storage ensuring that only changed blocks are sent to the StoreOnce backup system. This reduces both network traffic and storage usage which lowers costs. Every backup completes at the speed of an incremental data transfer but is stored as a synthetic full backup. This makes application recovery faster and more efficient.

RMC Express Protect

Before you configure Express Protect Backup in RMC, you need to add the HPE 3PAR StoreServ Storage and StoreOnce to RMC. Make sure to have proper zoning configured on the SAN switch between the HPE 3PAR StoreServ Storage and the data mover to mount the volumes. In this case, the data mover can either be HPE RMC or HPE StoreOnce.



- 1. Log in to the RMC portal.
- 2. Navigate to the **Recovery Manager Central** drop-down.
- 3. On the **Configure** tab, select **Storage Devices** . A list of available storage devices appears as in Figure B1.

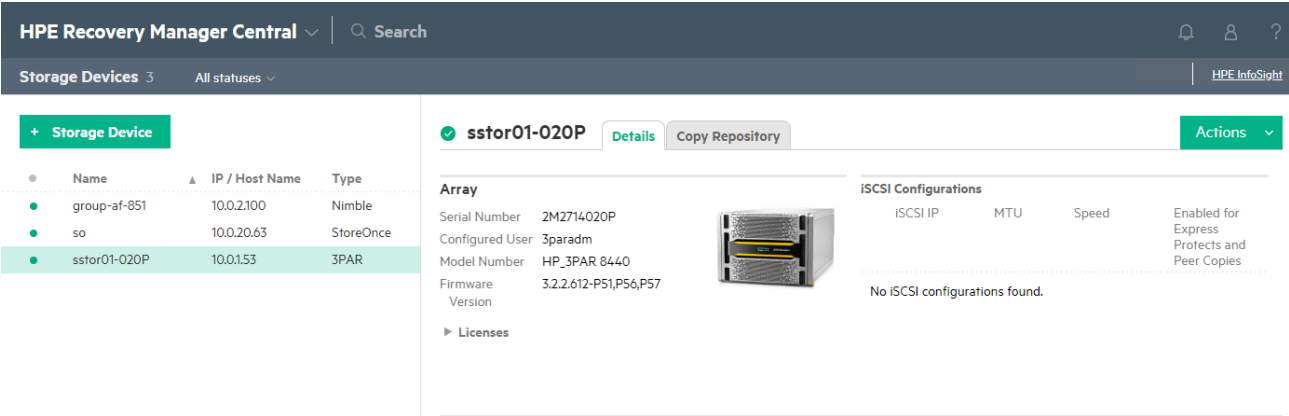


Figure B1. HPE Recovery Manager Central - Storage Devices window

- 4. Click **Storage Device** and add HPE 3PAR StoreServ Storage and HPE StoreOnce as in Figure B2.

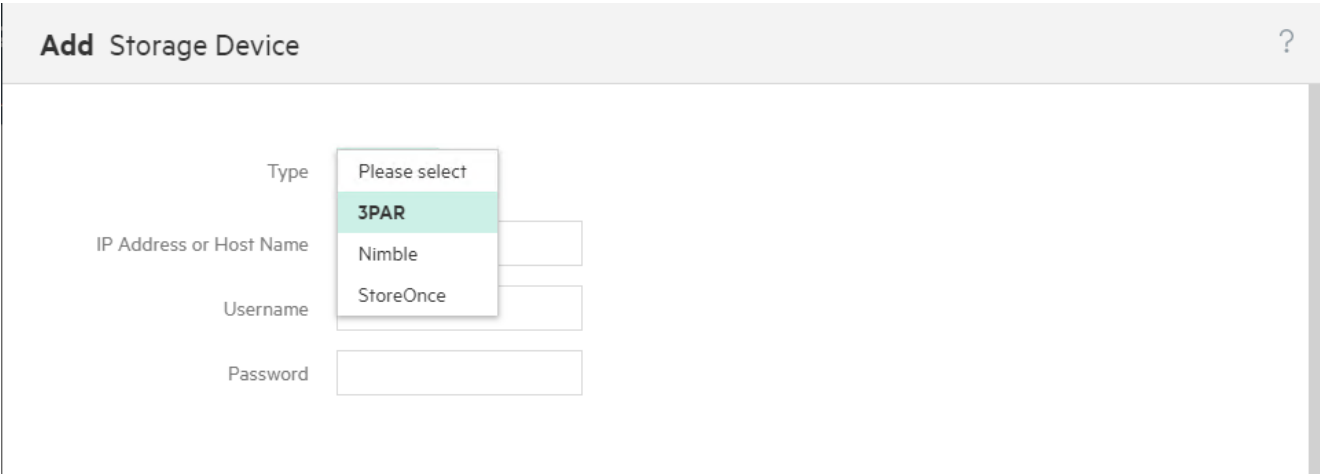


Figure B2. Adding storage devices in HPE Recovery Manager Central

- 5. Navigate to the **Recovery Manager Central** drop-down.



6. On the **Protect** tab, select **Volumes**.

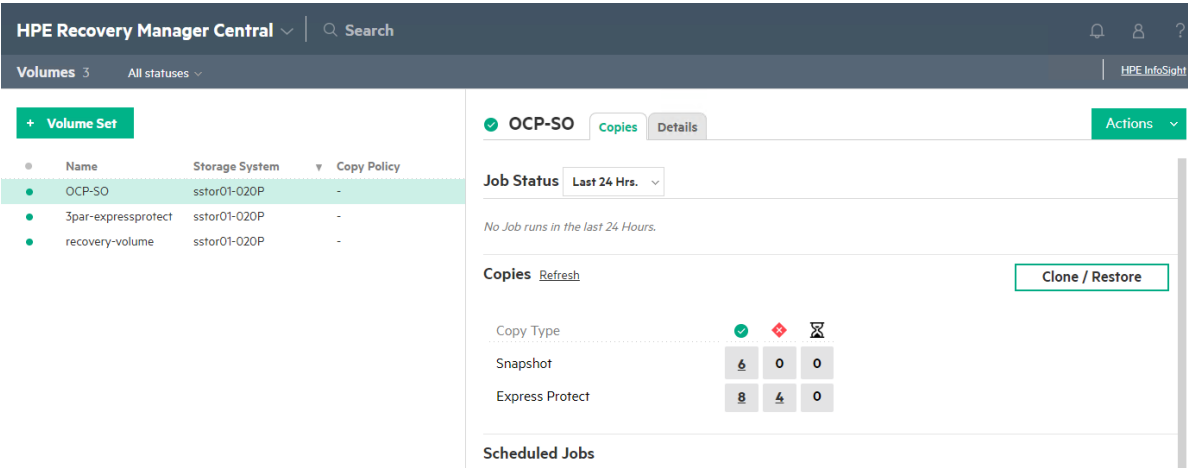


Figure B3. HPE Recovery Manager Central - Volume Set window

7. Click **Volume Set**.

A new window to create a volume set appears.

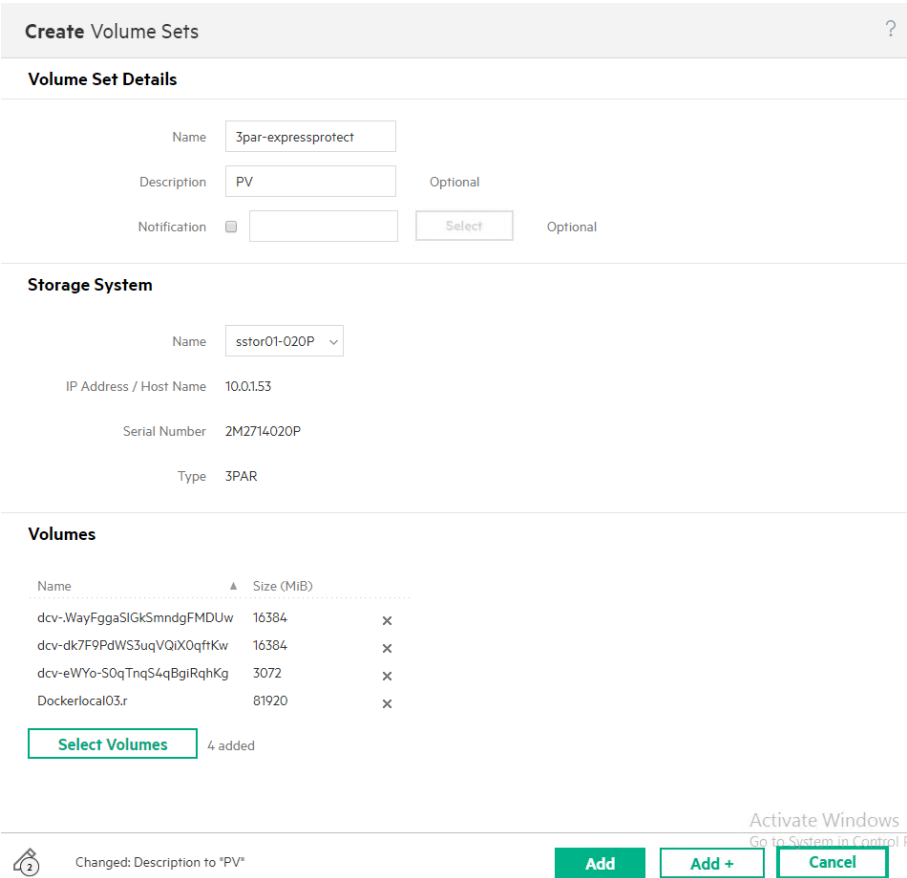


Figure B4. HPE Recovery Manager Central - Volume Set creation screen



8. Provide the details of the volume set, select the storage system, and then select the volumes. Then click **Add** to create volume set as in Figure B4.
9. Before protecting the volume set, add the **copy policies** under the **HPE Recovery Manager Central** drop-down and select the appropriate Snapshot and Express Protect schedules and HPE StoreOnce repository. Create the copy policy based on the business SLA requirement.

Create Copy Policy ?

Application Type Volumes Policy Name RMC-3par-SO Description

Protection Rules

+ Snapshot

+ Remote Snapshot

+ Express Protect

+ Catalyst Copy

+ Cloud Copy

+ Peer Copy

Snapshot Every 8 hours

Express Protect Source Snapshot Target so\OCP_Store

+ Schedule

Schedule Every Auto Delete After Prevent Deletion Until Verify Copy

Daily 1 Days 31 Days Never Hours

Source Snapshot Copy Repository so\OCP_Store Select Clear

View advanced option

Click on tiles to add more rules

Figure B5. HPE Recovery Manager Central - creating copy policy

10. Select the volume set and from the **Actions** drop-down list, select **Add Protection** as in Figure B6.

HPE Recovery Manager Central Search

Volumes 3 All statuses HPE InfoSight

+ Volume Set

Name	Storage System	Copy Policy
OCP-SO	ssstor01-020P	-
3par-expressprotect	ssstor01-020P	-
recovery-volume	ssstor01-020P	-

3par-expressprotect Copies Details

Job Status Last 24 Hrs. No Job runs in the last 24 Hours.

Copies Refresh

Copy Type

Snapshot	2	0	0
Express Protect	2	0	0
Clone	2	0	0

Clone /

Actions Edit Delete Attach Refresh Clone / Restore Add Protection Protect Once

Figure B6. HPE Recovery Manager Central - adding protection



11. From the Add Protection window, select the previously created copy policy and click **Select** as in Figure B7.

Add Protection?

▶ Volume(s) 3par-expressprotect selected for protection.

Copy Policy

Select

Effective From

Now

Select Copy Policy?

Total 7
















Name	
RMC-3par-SO	 
For_native_backup	 
RMC Gold Policy	  
RMC Peer Copy Policy	 
RMC Bronze Policy	
RMC Silver Policy	 
RMC Platinum Policy	  

Figure B7. HPE Recovery Manager Central - Selecting Protection policy

12. Select the **Protect Once** option for the created volume set. From the **Protection Type** drop-down list, select Express Protect as in Figure B8.

Protect Once?

▶ Volume(s) 3par-expressprotect selected for protection.

Protection Type

Express Protect

Choose any of the custom protection template required for the resources. You can also select protection from the predefined policies.

▶ Snapshot

▶ Express Protect

Source Snapshot

Target (Auto)

Start Date and Time

Now




Figure B8. HPE Recovery Manager Central - protecting the volume set



13. View the progress/status on the **Activities** tab under the **HPE Recovery Manger Central** drop-down as in Figure B9.

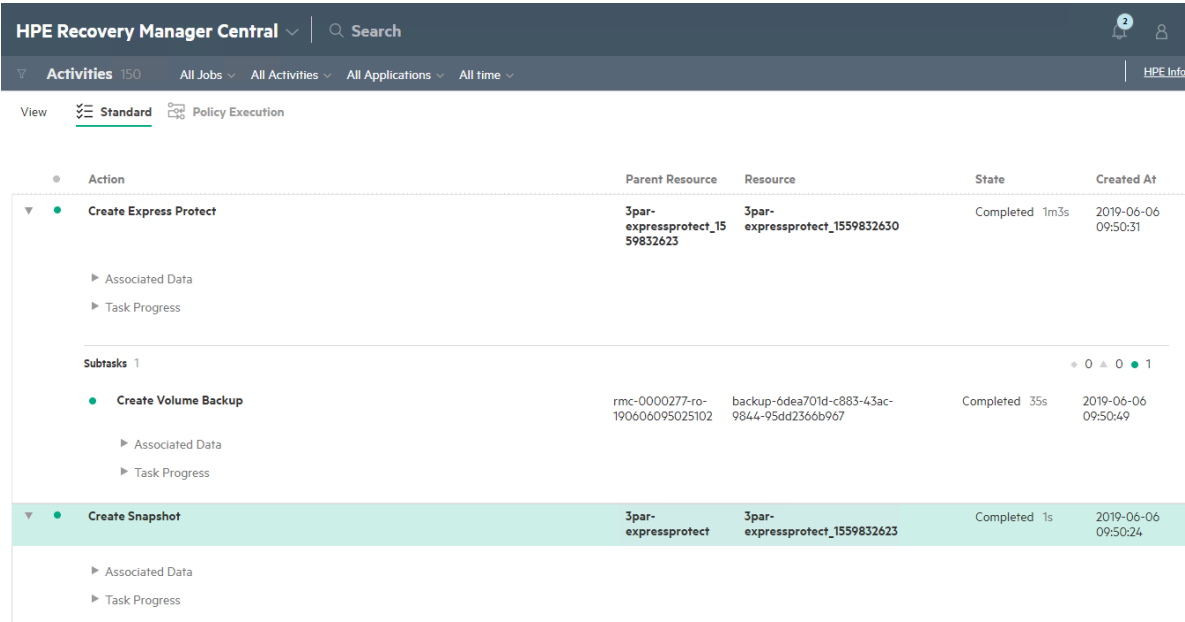


Figure B9. HPE Recovery Manager Central Activities windows

14. Figure B10 shows the backup/restore report from the HPE StoreOnce Management GUI.

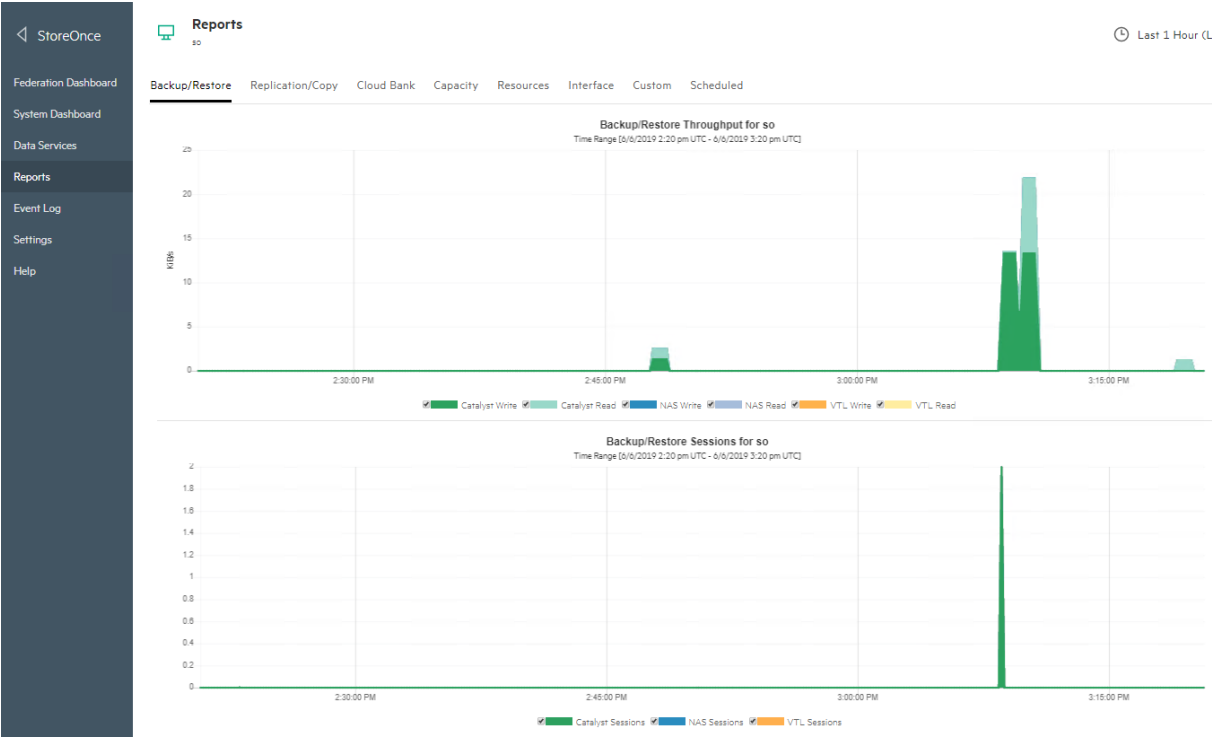


Figure B10. HPE StoreOnce backup/restore reports tab



Restore volume

RMC allows multiple options to restore the express protected volume. This is shown in Figure A5.

- 1. Select **Restore** to demonstrate the restore process. Prior to proceeding with this step, ensure the replica set count is set to zero using the Deployment API object in the OCP console as shown in Figure A4.
- 2. As shown in Figure B11, select the date, select the available copies of backup from the list and then click **Next**.

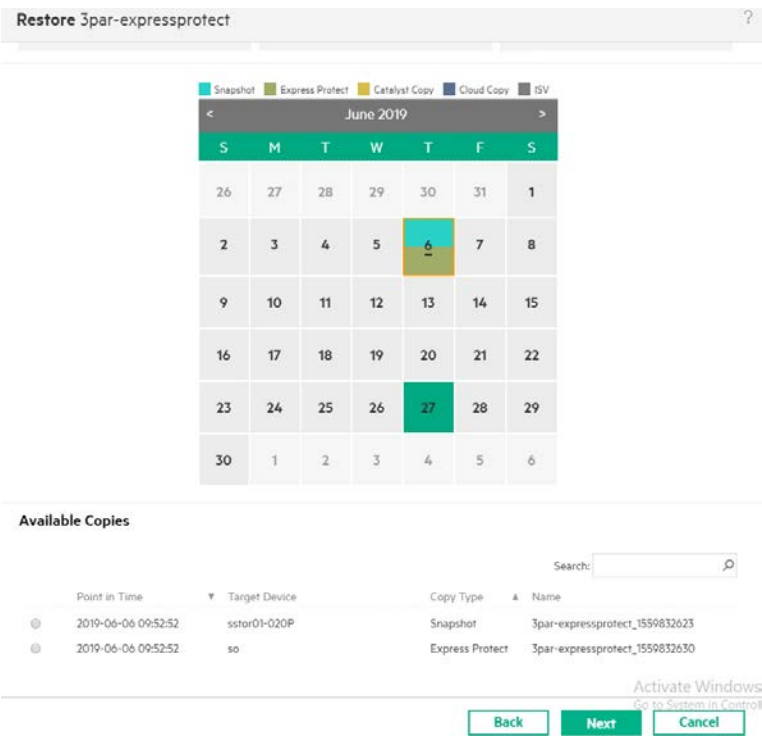


Figure B11. HPE Recovery Manager Central physical clone details

- 3. View the progress/status on the **Activities** tab under the **HPE Recovery Manger Central** drop-down as in Figure B12.

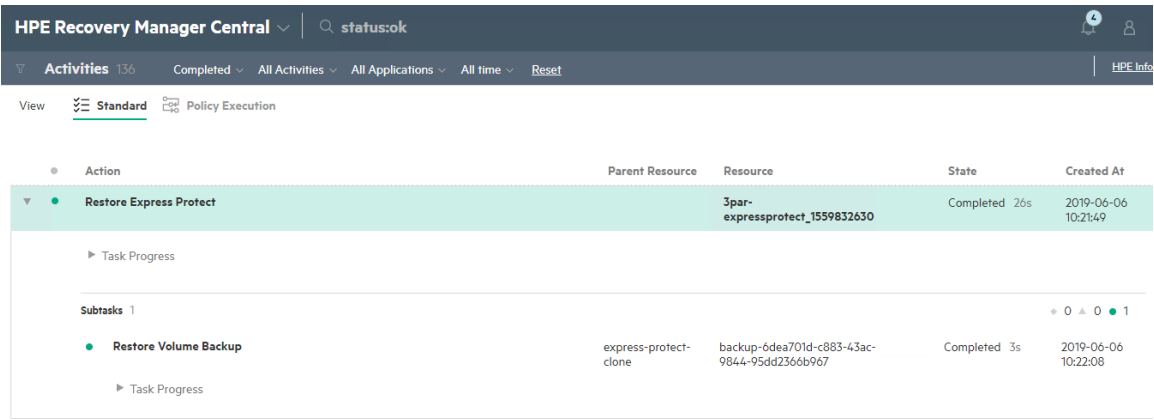


Figure B12. HPE Recovery Manager Central activities

- 4. Once again, view the backup/restore report from the HPE StoreOnce Management GUI as shown Figure B10.



Appendix C: HPE 3PAR StoreServ Storage replication protection

The HPE 3PAR Volume Plugin for Docker supports two storage replication methods:

1. Active/Passive replication
2. Peer Persistence replication

Active/Passive replication (Disruptive)

With Active/Passive replication, only one array is in an active state serving the VLUNs of a given replicated volume at any point in time. With this scenario, the remote copy group (RCG) is failed over manually via the HPE 3PAR CLI to the secondary array at which point the secondary array becomes active. However, the VLUNs of the failed over volumes are still not exported by the secondary array to the host. In order to trigger that, the container/POD running on the host needs to be restarted.

Configuring Active/Passive replication within the HPE 3PAR Volume Plugin for Docker

Refer to <https://github.com/hpe-storage/python-hpedockerplugin/blob/master/docs/active-passive-based-replication.md>, to learn how to configure replication using the HPE 3PAR Volume Plugin for Docker.

Peer Persistence replication (Non-Disruptive)

HPE 3PAR Peer Persistence software enables HPE 3PAR systems located at metropolitan distances to act as peers to each other. This presents a nearly continuous storage system to the hosts that are connected to them. This capability allows for the configuration of a high-availability solution between two sites or data centers where failover and failback are completely transparent to the hosts and applications running on those hosts. Compared to the traditional failover models where the hosts must be restarted upon failover, the Peer Persistence software allows hosts to remain online serving their business applications even when they switch from their original site to the disaster recovery (DR) site. This results in an improved recovery time compared to other methods. The Peer Persistence software achieves this key enhancement by taking advantage of the Asymmetric Logical Unit Access (ALUA) capability that allows paths to a SCSI device to be marked as having different characteristics.

Using Peer Persistence, an OpenShift user mounts a replicated volume(s) and then the HPE 3PAR Docker Plugin creates VLUNs corresponding to the replicated volume(s) on both the arrays. However, the VLUN(s) is served only by the active array with the other array being in standby mode. When the corresponding RCG is switched over or the primary array goes down, the secondary array takes over and makes the VLUN(s) available. After switchover, the active array goes to standby mode while the secondary array becomes active. To use Peer Persistence replication, the following prerequisites should be met:

- Remote copy is configured, up and running.
- Quorum Witness is running with the primary and secondary arrays registered.
- The multipath daemon is running so that non-disruptive, seamless mounting of VLUN(s) on the host is possible.

Configuring Peer Persistence (Non-Disruptive) replication within the HPE 3PAR Volume Plugin for Docker

Compared to Active/Passive configuration, the only discriminator with Peer Persistence is the presence of a `quorum_witness_ip` sub-field under the `replication_device` field.

Refer to the following document to learn how to configure replication using the HPE 3PAR Volume Plugin for Docker: <https://github.com/hpe-storage/python-hpedockerplugin/blob/master/docs/peer-persistence-based-replication.md>.



Change Tracker

Version	Release Date	Changes
1.0	07/03/2019	Initial release
1.0.1	07/11/2019	Title change to include HPE Solution Architecture and removing capitalization of backup and recovery
1.0.2	07/12/2019	Swapped out Figure 1 (removed etcd) and moved etcd sub-heading to main heading



Resources and additional links

Validated deployment architecture, <https://h20195.www2.hpe.com/V2/GetDocument.aspx?docname=a00056101enw>

HPE Reference Architectures, hpe.com/info/ra

HPE Servers, hpe.com/servers

HPE Storage, hpe.com/storage

HPE Networking, hpe.com/networking

HPE Technology Consulting Services, hpe.com/us/en/services/consulting.html

HPE Information Library, <http://h17007.www1.hpe.com/us/en/enterprise/integrated-systems/info-library/index.aspx?cat=convergedsystems&subcat=cs750>

HPE StoreOnce, <https://h20195.www2.hpe.com/v2/gethtml.aspx?docname=c04328820>

HPE RMC, https://support.hpe.com/hpsc/doc/public/display?docId=emr_na-c04552659

To help us improve our documents, please provide feedback at hpe.com/contact/feedback.

Share 

© Copyright 2019 Hewlett Packard Enterprise Development LP. The information contained herein is subject to change without notice. The only warranties for Hewlett Packard Enterprise products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. Hewlett Packard Enterprise shall not be liable for technical or editorial errors or omissions contained herein.

Red Hat is a registered trademark of Red Hat, Inc. in the United States and other countries. Linux is the registered trademark of Linus Torvalds in the U.S. and other countries. Intel is a trademark of Intel Corporation in the U.S. and other countries.

OCP3793 Version 1.0.2, July 2019