# Red Hat OpenShift Data Foundation 4.16

## 4.16 Release Notes

Release notes for features and enhancements, known issues, and other important release information.

Last Updated: 2024-08-09

# Red Hat OpenShift Data Foundation 4.16 4.16 Release Notes

Release notes for features and enhancements, known issues, and other important release information.

## Legal Notice



## Abstract

The release notes for Red Hat OpenShift Data Foundation 4.16 summarizes all new features and enhancements, notable technical changes, and any known bugs upon general availability.

# Table of Contents

MAKING OPEN SOURCE MORE INCLUSIVE ..... 4

CHAPTER 1. OVERVIEW ..... 5
1.1. ABOUT THIS RELEASE 5
1.1.1. Limitation in this release 6

CHAPTER 2. NEW FEATURES ..... 7
2.1. DISASTER RECOVERY SOLUTION 7
2.1.1. User interface support for discovered applications in Disaster Recovery 7
2.1.2. Disaster recovery solution for Applications that require Kube resource protection with labels 7
2.1.3. Expand discovered application DR support to multi-namespace Applications 7
2.1.4. OpenShift virtualization workloads for Regional-DR 7
2.1.5. OpenShift virtualization in a stretch cluster 7
2.1.6. Recovering to a replacement cluster for Regional-DR 7
2.1.7. Enable monitoring support for ACM Subscription application type 8
2.1.8. Hub recovery support for co-situated and neutral site Regional-DR deployments 8
2.2. WEEKLY CLUSTER-WIDE ENCRYPTION KEY ROTATION 8
2.3. SUPPORT CUSTOM TAINTS 8
2.4. SUPPORT FOR SELINUX MOUNT FEATURE WITH READWRITEONCEPOD ACCESS MODE 9
2.5. SUPPORT FOR READWRITEONCEPOD ACCESS MODE 9
2.6. FASTER CLIENT IO OR RECOVERY IO DURING OSD BACKFILL 9
2.7. SUPPORT FOR GENERIC EPHEMERAL STORAGE FOR PODS 9
2.8. CROSS STORAGE CLASS CLONE 9
2.9. OVERPROVISION LEVEL POLICY CONTROL 9

CHAPTER 3. ENHANCEMENTS ..... 10
3.1. NEW ELEMENTS FOR BUCKET POLICIES 10
3.2. ADDITION OF A NEW AWS REGION TO THE MULTICLOUD OBJECT GATEWAY OPERATOR 10
3.3. INCREASE IN RESOURCE ALLOCATION FOR OPENSHIFT DATA FOUNDATION MULTICLOUD OBJECT GATEWAY BACKINGSTORE 10
3.4. MULTICLOUD OBJECT GATEWAY CREATED ROUTES TO WORK WITH HTTPS ONLY 10
3.5. ADDITION OF PROTECTED CONDITION TO DR PROTECTED WORKLOADS WITH METRICS AND ALERTS FOR MONITORING 10
3.6. SUPPORT FOR LISTING MULTIPLE UPLOADS IN NAMESPACESTORE FILESYSTEM 10
3.7. OPTION TO MODIFY THRESHOLDS FOR CEPH FULL, NEARFULL, AND BACKFILLFULL ATTRIBUTES 11

CHAPTER 4. TECHNOLOGY PREVIEWS ..... 12
4.1. AZURE KEY VAULT OPTION FOR CLUSTER WIDE/PV ENCRYPTION 12
4.2. BUCKET LOGGING AND LOG BASED REPLICATION OPTIMIZATION FOR MULTICLOUD OBJECT GATEWAY BUCKETS 12
4.3. MULTI NETWORK PLUG-IN (MULTUS) SUPPORT FOR IPV6 12

CHAPTER 5. DEVELOPER PREVIEWS ..... 13
5.1. SUPPORT FOR TOPOLOGY AWARENESS AND REPLICA-1 IN OPENSHIFT DATA FOUNDATION EXTERNAL MODE 13
5.2. MULTI-VOLUME CONSISTENCY FOR BACKUP - CEPHFS 13
5.3. STORAGE CLASS WITH REPLICA 2 FOR CEPHFS 13
5.4. AUTOMATIC SCALING OF RGW 13
5.5. CEPH REPLICA-2 POOL WITH BOTH DISKS IN THE SAME ZONE 13
5.6. RGW ERASURE CODING IN INTERNAL MODE 14

CHAPTER 6. BUG FIXES ..... 15
6.1. DISASTER RECOVERY 15

6.2. MULTICLOUD OBJECT GATEWAY 15
6.3. CEPH CONTAINER STORAGE INTERFACE (CSI) DRIVER 17
6.4. OCS OPERATOR 17
6.5. OPENSHIFT DATA FOUNDATION CONSOLE 17
6.6. ROOK 18
6.7. CEPH MONITORING 19

**CHAPTER 7. KNOWN ISSUES** ..... **21**
7.1. DISASTER RECOVERY 21
7.2. MULTICLOUD OBJECT GATEWAY 24
7.3. CEPH 25
7.4. CSI DRIVER 25
7.5. OPENSHIFT DATA FOUNDATION CONSOLE 26
7.6. OCS OPERATOR 26
7.7. NON-AVAILABILITY OF IBM Z PLATFORM 27

# MAKING OPEN SOURCE MORE INCLUSIVE

Red Hat is committed to replacing problematic language in our code, documentation, and web properties. We are beginning with these four terms: master, slave, blacklist, and whitelist. Because of the enormity of this endeavor, these changes will be implemented gradually over several upcoming releases. For more details, see our CTO Chris Wright's message.

# CHAPTER 1. OVERVIEW

Red Hat OpenShift Data Foundation is software-defined storage that is optimized for container environments. It runs as an operator on OpenShift Container Platform to provide highly integrated and simplified persistent storage management for containers.

Red Hat OpenShift Data Foundation is integrated into the latest Red Hat OpenShift Container Platform to address platform services, application portability, and persistence challenges. It provides a highly scalable backend for the next generation of cloud-native applications, built on a technology stack that includes Red Hat Ceph Storage, the Rook.io Operator, and NooBaa's Multicloud Object Gateway technology.

Red Hat OpenShift Data Foundation is designed for FIPS. When running on RHEL or RHEL CoreOS booted in FIPS mode, OpenShift Container Platform core components use the RHEL cryptographic libraries submitted to NIST for FIPS Validation on only the x86_64, ppc64le, and s390X architectures. For more information about the NIST validation program, see Cryptographic Module Validation Program. For the latest NIST status for the individual versions of the RHEL cryptographic libraries submitted for validation, see Compliance Activities and Government Standards.

Red Hat OpenShift Data Foundation provides a trusted, enterprise-grade application development environment that simplifies and enhances the user experience across the application lifecycle in a number of ways:

- Provides block storage for databases.
- Shared file storage for continuous integration, messaging, and data aggregation.
- Object storage for cloud-first development, archival, backup, and media storage.
- Scale applications and data exponentially.
- Attach and detach persistent data volumes at an accelerated rate.
- Stretch clusters across multiple data-centers or availability zones.
- Establish a comprehensive application container registry.
- Support the next generation of OpenShift workloads such as Data Analytics, Artificial Intelligence, Machine Learning, Deep Learning, and Internet of Things (IoT).
- Dynamically provision not only application containers, but data service volumes and containers, as well as additional OpenShift Container Platform nodes, Elastic Block Store (EBS) volumes and other infrastructure services.

## 1.1. ABOUT THIS RELEASE

Red Hat OpenShift Data Foundation 4.16 (RHSA-2024:4591) is now available. New enhancements, features, and known issues that pertain to OpenShift Data Foundation 4.16 are included in this topic.

Red Hat OpenShift Data Foundation 4.16 is supported on the Red Hat OpenShift Container Platform version 4.16. For more information, see Red Hat OpenShift Data Foundation Supportability and Interoperability Checker.

For Red Hat OpenShift Data Foundation life cycle information, refer to the layered and dependent products life cycle section in Red Hat OpenShift Container Platform Life Cycle Policy .

### 1.1.1. Limitation in this release

In this release, upgrading from OpenShift Data Foundation version 4.15 to 4.16 is blocked due to a known issue. Only fresh deployments of OpenShift Data Foundation 4.16 version are currently supported. This limitation will be removed in an upcoming OpenShift Data Foundation 4.16.z release.

# CHAPTER 2. NEW FEATURES

This section describes new features introduced in Red Hat OpenShift Data Foundation 4.16.

## 2.1. DISASTER RECOVERY SOLUTION

### 2.1.1. User interface support for discovered applications in Disaster Recovery

For discovered applications not deployed using RHACM (discovered applications), the OpenShift Data Foundation Disaster Recovery solution extends protection with a new user experience for failover and failback operations that are managed using RHACM.

For more information, see Metro-DR protection for discovered applications and Regional-DR protection for discovered applications.

### 2.1.2. Disaster recovery solution for Applications that require Kube resource protection with labels

The OpenShift Data Foundation Disaster Recovery solution supports applications that are developed or deployed using an imperative model. The cluster resources for these discovered applications are protected and restored at the secondary cluster using OpenShift APIs for Data Protection (OADP).

For instructions on how to enroll discovered applications, see Enrolling discovered applications for Metro-DR and Enrolling discoverd applications for Regional-DR .

### 2.1.3. Expand discovered application DR support to multi-namespace Applications

The OpenShift Data Foundation Disaster Recovery solution now extends protection to discovered applications that span across multiple namespaces.

### 2.1.4. OpenShift virtualization workloads for Regional-DR

Regional disaster recovery (Regional-DR) solution can be easily set up for OpenShift Virtualization workloads using OpenShift Data Foundation.

For more information, see the knowledgebase article, Use OpenShift Data Foundation Disaster Recovery to Protect Virtual Machines.

### 2.1.5. OpenShift virtualization in a stretch cluster

Disaster recovery with stretch clusters for workloads based on OpenShift Virtualization technology using OpenShift Data Foundation can now be easily set up.

For more information, see the OpenShift Virtualization in OpenShift Container Platform guide.

### 2.1.6. Recovering to a replacement cluster for Regional-DR

When a primary or a secondary cluster of Regional-DR fails, the cluster can be either repaired or wait for the recovery of the existing cluster, or replace the cluster entirely if the cluster is irredeemable. OpenShift Data Foundation provides the ability to replace a failed primary or a secondary cluster with a new cluster and enable failover (relocate) to the new cluster.

For more information, see Recovering to a replacement cluster .

### 2.1.7. Enable monitoring support for ACM Subscription application type

The disaster recovery dashboard on Red Hat Advanced Cluster Management (RHACM) console is extended to display monitoring data for Subscription type applications in addition to ApplicationSet type applications.

Data such as the following can be monitored:

- Volume replication delays
- Count of protected Subscription type applications with or without replication issues
- Number of persistent volumes with replication healthy and unhealthy
- Application-wise data like the following:
  - Recovery Point Objective (RPO)
  - Last sync time
  - Current DR activity status (Relocating, Failing over, Deployed, Relocated, Failed Over)
- Application-wise persistent volume count with replication healthy and unhealthy

### 2.1.8. Hub recovery support for co-situated and neutral site Regional-DR deployments

The Regional disaster recovery solutions of OpenShift Data Foundation now support neutral site deployments and hub recovery of co-situated managed clusters using Red Hat Advanced Cluster Management. For configuring hub recovery setup, a 4th cluster is required which acts as the passive hub. The passive hub cluster can be set up in either one of the following ways:

- The primary managed cluster (Site-1) can be co-situated with the active RHACM hub cluster while the passive hub cluster is situated along with the secondary managed cluster (Site-2).
- The active RHACM hub cluster can be placed in a neutral site (Site-3) that is not impacted by the failures of either of the primary managed cluster at Site-1 or the secondary cluster at Site-2. In this situation, if a passive hub cluster is used it can be placed with the secondary cluster at Site-2.

For more information, see Regional-DR chapter on Hub recovery using Red Hat Advanced Cluster Management.

## 2.2. WEEKLY CLUSTER-WIDE ENCRYPTION KEY ROTATION

Security common practices require periodic encryption key rotation. OpenShift Data Foundation automatically rotates the encryption keys stored in kubernetes secret (non-KMS) on a weekly basis.

For more information, see Cluster-wide encryption.

## 2.3. SUPPORT CUSTOM TAINTS

Custom taints can be configured using the storage cluster CR by directly adding tolerations under the placement section of the CR. This helps to simplify the process of adding custom taints.

For more information, see the knowledgebase article, How to add toleration for the "non-ocs" taints to the OpenShift Data Foundation pods?

## 2.4. SUPPORT FOR SELINUX MOUNT FEATURE WITH READWRITEONCEPOD ACCESS MODE

OpenShift Data Foundation now supports SELinux mount feature with ReadWriteOncePod access mode. This feature helps to reduce the time taken to change the SELinux labels of the files and folders in a volume, especially when the volume has many files and is on a remote filesystem such as CephFS.

## 2.5. SUPPORT FOR READWRITEONCEPOD ACCESS MODE

OpenShift Data Foundation provides ReadWriteOncePod (RWOP) access mode to ensure that only one pod across the whole cluster can read the persistent volume claim (PVC) or write to it.

## 2.6. FASTER CLIENT IO OR RECOVERY IO DURING OSD BACKFILL

Client IO or recovery IO can be set to be favored during a maintenance window. Favoring recovery IO over client IO significantly reduces OSD recovery time.

For more information in setting the recovery profile, see Enabling faster client IO or recovery IO during OSD backfill.

## 2.7. SUPPORT FOR GENERIC EPHEMERAL STORAGE FOR PODS

OpenShift Data Foundation provides support for generic ephemeral volume. This support enables a user to specify generic ephemeral volumes in its pod specification and tie the lifecycle of the PVC with the pod.

## 2.8. CROSS STORAGE CLASS CLONE

OpenShift Data Foundation provides an ability to move from a storage class with replica 3 to replica 2 or replica 1 while cloning. This helps to reduce storage footprint.

For more information, see Creating a clone .

## 2.9. OVERPROVISION LEVEL POLICY CONTROL

Overprovision control mechanism enables defining a quota on the amount of persistent volume claims (PVCs) consumed from a storage cluster, based on the specific application namespace.

When this overprovision control mechanism is enabled, overprovisioning the PVCs consumed from the storage cluster is prevented.

For more information, see Overprovision level policy control .

# CHAPTER 3. ENHANCEMENTS

This section describes the major enhancements introduced in Red Hat OpenShift Data foundation 4.16.

## 3.1. NEW ELEMENTS FOR BUCKET POLICIES

OpenShift Data Foundation now has the bucket policy elements, **NotPrincipal**, **NotAction**, and **NotResource**. For more information on these elements, see IAM JSON policy elements reference.

## 3.2. ADDITION OF A NEW AWS REGION TO THE MULTICLOUD OBJECT GATEWAY OPERATOR

A new AWS region, **ca-west-1**, is added to the supported regions of the Multicloud Object Gateway (MCG) operator for the creation of default backingstore.

## 3.3. INCREASE IN RESOURCE ALLOCATION FOR OPENSHIFT DATA FOUNDATION MULTICLOUD OBJECT GATEWAY BACKINGSTORE

The default resources for PV pool CPU and memory are increased to 999m and 1Gi respectively to enable more resource allocation for OpenShift Data Foundation MCG BackingStore.

## 3.4. MULTICLOUD OBJECT GATEWAY CREATED ROUTES TO WORK WITH HTTPS ONLY

For deployments that need to disable HTTP and use only HTTPS, an option is added to set **DenyHTTP** to the storage cluster CR **spec.multiCloudGateway.denyHTTP**. This causes the Multicloud Object Gateway created routes to use HTTPS only.

## 3.5. ADDITION OF PROTECTED CONDITION TO DR PROTECTED WORKLOADS WITH METRICS AND ALERTS FOR MONITORING

Protected condition to DR protected workloads is added by summarizing various conditions regarding the DR protected workload from the **ManagedCluster**, and metrics and alerts are generated based on the same.

DR protected workload health at the hub is reflected based only on the time when contents of the respective PVCs are synced. This applies only to RegionalDR use cases and not to MetroDR use cases. On the **ManagedCluster**, the workload DR protection health is expanded into several conditions. This makes it non-trivial for a user to monitor the workload DR health across these conditions. This added protected condition and alerts helps better workload DR protection monitoring.

## 3.6. SUPPORT FOR LISTING MULTIPLE UPLOADS IN NAMESPACESTORE FILESYSTEM

You can now list the files that are still uploading or list incomplete multipart uploads in NamespaceStore Filesystem by using the following command:

```
$ s3api list--multipart-uploads --bucket <bucket_name>
```

## 3.7. OPTION TO MODIFY THRESHOLDS FOR CEPH FULL, NEARFULL, AND BACKFILLFULL ATTRIBUTES

Depending on the cluster requirements, the **full**, **nearfull**, and **backfillfull** threshold values can be updated by using the **odf-cli** CLI command.

For example:

**odf set full <val>**

**odf set nearful <val>**

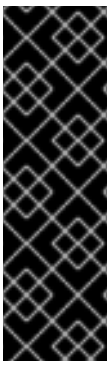
**odf set backfillfull <val>**

NOTE

The value must be in the range of 0.0 to 1.0 and you need to ensure that the value is not very close to 1.0.

# CHAPTER 4. TECHNOLOGY PREVIEWS

This section describes the technology preview features introduced in Red Hat OpenShift Data Foundation 4.16 under Technology Preview support limitations.



**IMPORTANT**

Technology Preview features are not supported with Red Hat production service level agreements (SLAs), might not be functionally complete, and Red Hat does not recommend using them for production. These features provide early access to upcoming product features, enabling customers to test functionality and provide feedback during the development process.

Technology Preview features are provided with a limited support scope, as detailed on the Customer Portal: Technology Preview Features Support Scope .

## 4.1. AZURE KEY VAULT OPTION FOR CLUSTER WIDE/PV ENCRYPTION

OpenShift Data Foundation supports Azure Key Vault for cluster-wide PV encryption on Microsoft Azure platform. This helps to use a local native solution for Azure instead of a third party tool.

For more information, see Creating an OpenShift Data Foundation cluster .

## 4.2. BUCKET LOGGING AND LOG BASED REPLICATION OPTIMIZATION FOR MULTICLOUD OBJECT GATEWAY BUCKETS

Supports replication of large amounts of data between Multicloud Object Gateway (MCG) and Amazon Web Services (AWS) or MCG and MCG. The support for log-based replication for AWS S3 using bucket logging is extended to MCG bucket for optimization. The log-based replication optimization also supports object prefix filtering.

For more information, see Enabling bucket logging for Multicloud Object Gateway .

## 4.3. MULTI NETWORK PLUG-IN (MULTUS) SUPPORT FOR IPV6

Multus networks can be configured to use either IPv4 or IPv6. OpenShift Data Foundation supports Multus networks that are pure IPv4 or pure IPv6. Networks cannot be in a mixed mode. For more information on Mutlus, see Multi network plug-in (Multus) support .

# CHAPTER 5. DEVELOPER PREVIEWS

This section describes the developer preview features introduced in Red Hat OpenShift Data Foundation 4.16.

**IMPORTANT**

Developer preview feature is subject to Developer preview support limitations. Developer preview releases are not intended to be run in production environments. The clusters deployed with the developer preview features are considered to be development clusters and are not supported through the Red Hat Customer Portal case management system. If you need assistance with developer preview features, reach out to the ocs-devpreview@redhat.com mailing list and a member of the Red Hat Development Team will assist you as quickly as possible based on availability and work schedules.

## 5.1. SUPPORT FOR TOPOLOGY AWARENESS AND REPLICA-1 IN OPENSHIFT DATA FOUNDATION EXTERNAL MODE

OpenShift Data Foundation supports provisioning storage from replica-1 using a storage class. The storage class is created to provision storage from replica 1 Ceph pools that are located in separate zones. This helps applications that have multiple running instances with each service instance creating a new claim, which is expected to be located in a different zone. As these applications have their own redundant instances, they do not require redundancy at the data layer. For more information, see the knowledgebase article, OpenShift Data Foundation external mode support for RBD/block Replica-1 pools and topology awareness.

## 5.2. MULTI-VOLUME CONSISTENCY FOR BACKUP - CEPHFS

To provide better support for applications and to support OpenShift Virtualization, crash consistent multi volume consistency groups are provided for backup solutions. This helps applications that are deployed over multiple volumes.

## 5.3. STORAGE CLASS WITH REPLICA 2 FOR CEPHFS

OpenShift Data Foundation provides a way to create a new CephFS based storage class with 2 replicas. This helps to provide better storage efficiency in case only CephFs is being used. Also, this provides a solution for unbalanced usage of either RBD or CephFS. For more information, see the knowledgebase article, Using replica 2 for CephFS .

## 5.4. AUTOMATIC SCALING OF RGW

OpenShift Data Foundation provides the ability to enable auto scaling to automatically adjust the serviceability and performance of RADOS Gateway (RGW) to serve as per the S3 load. For more information, see the knowledgebase article, Autoscaling for RGW in OpenShift Data Foundation via HPA using KEDA.

## 5.5. CEPH REPLICA-2 POOL WITH BOTH DISKS IN THE SAME ZONE

To protect applications in case of an OSD disk failure, replica-2 pool with both disks in a single zone can be created.

For more information, see the knowledgebase article RBD Replica-2 with both disks in same zone .

## 5.6. RGW ERASURE CODING IN INTERNAL MODE

Erasure coding deployment for RGW using CLI is supported to help reduce cost by increasing storage efficiency.

For more information, see the knowledgebase article Support for RGW Erasure Coding in Internal Mode .

# CHAPTER 6. BUG FIXES

This section describes the notable bug fixes introduced in Red Hat OpenShift Data Foundation 4.16.

## 6.1. DISASTER RECOVERY

**FailOver of applications are hung in FailingOver state**

Previously, applications were not DR protected successfully because of the errors in protecting required resources to the provided S3 stores. So, failing over such applications resulted in FailingOver state.

With this fix, a metric and a related alert is added to the application DR protection health that shows an alert to rectify protection issues after DR protects the applications. As a result, the applications that are successfully protected are failed over.

(BZ#2248723)

**Post hub recovery, applications which were in FailedOver state consistently report FailingOver**

Previously, the Ramen hub operator on a recovered hub cluster reported the status of a managed cluster that survived a loss of both the hub and its peer managed cluster as **Ready** for the future failover actions without ensuring if the surviving cluster is reporting such a status.

With this fix, Ramen hub operator ensures if the target cluster is ready for a failover operation before initiating the action. As a result, any failover initiated is successful or if stale resources still exist on the failover target cluster, the operator stalls the failover till the stale resources are cleaned up.

(BZ#2270259)

## 6.2. MULTICLOUD OBJECT GATEWAY

**Multicloud Object Gateway (MCG) DB PVC consumption more than 400GB**

Previously, the Multicloud Object Gateway (MCG) Database (DB) showed increased DB size when it was not necessary as activity logs were being saved to the DB.

With this fix, the object activity logs are converted to regular debug logs. As a result, NooBaa DB no longer shows increased DB size.

(BZ#2141422)

**Log based replication works even after removing the replication policy from the OBC**

Previously, it was not possible to remove a log based replication policy from the object bucket claims (OBCs) as replication policy evaluation resulted in an error when presented with an empty string.

With this fix, replication policy evaluation method is modified to enable removal of replication policy from the OBCs.

(BZ#2266805)

**Multicloud Object Gateway (MCG) component security context fix**

Previously, when the default security context constraint (SCC) for the Multicloud Object Gateway (MCG) pods was updated to avoid the defaults set by OpenShift, the security scanning process failed.

With this fix, when the SCC is updated to override the defaults, MCG's behaviour does not change so that the security scan passes.

(BZ#2273670)

**NooBaa operator logs expose AWS secret**

Previously, the NooBaa operator logs exposed the AWS secret as plain text, which caused potential risk that anyone with logs could access the buckets.

With this fix, noobaa-operator logs no longer expose AWS secret.

(BZ#2277186)

**AWS S3 list takes a long time**

Previously, AWS S3 took a long time to list the objects as two database queries were used instead of a single one.

With this fix, the queries are restructured into a single one, thereby reducing the calls to the database and time to complete the list objects operation.

(BZ#2277990)

**After upgrade to OpenShift Data Foundation the standalone MCG backing store gets rejected**

Previously, when trying to use a persistent volume (PV) pool, **xattr** is used to save metadata of objects. However, updates to that metadata fails as Filesystem on PV does not support **xattr**.

With this fix, there is a failback if Filesystem does not support **xattr** and metadata is saved in a file.

(BZ#2278389)

**Multicloud Object Gateway database persistent volume claim (PVC) consumption rising continuously**

Previously, object bucket claim (OBC) deletion, which resulted in deletion of all its objects took time to free up from the database. This was because of the limited work by MCG's database cleaner causing a slow and limited deletion of entries from the database.

With this fix, updating the DB Cleaner configuration for MCG is possible. The DB Cleaner is a process that removes old deleted entries from the MCG Database. The exposed configurations are frequency of runs and age of entries to be deleted.

(BZ#2279742)

**Multicloud Object Gateway bucket lifecycle policy does not delete all objects**

Previously, the velocity of deletion of expired objects was very low.

With this fix, the batch size and number of runs per day to delete expired objects is increased.

(BZ#2279964) (BZ#2283753)

**HEAD-request returns the HTTP 200 Code for the prefix path instead of 404 from the API**

Previously, when trying to read or head an object which is a directory on the NamespaceStore Filesystem bucket of Multicloud Object Gateway, if the trailing / character is missing, the request returned HTTP 200 code for the prefix path instead of 404.

With this fix, **ENOENT** is returned when the object is a directory but the key is missing the trailing /.

(BZ#2280664)

**Multicloud Object Gateway Backingstore In Phase: "Connecting" with "Invalid URL"**

Previously, the operator failed to get the system information in the reconciliation loop which prevented the successful completion of the reconciliation. This was due to a bug in the URL parsing that caused the parsing to fail when the address was IPv6.

With this fix, the case of IPv6 address is handled as the URL host. As a result, the operator successfully completes the system reconciliation.

(BZ#2284652)

## 6.3. CEPH CONTAINER STORAGE INTERFACE (CSI) DRIVER

**PVC cloning failed with an error "RBD image not found"**

Previously, restore of volume snapshot failed when the parent of the snapshot did not exist as the CephCSI driver falsely identified an RBD image in trash to exist due to a bug in the driver.

With this fix, the CephCSI driver bug is fixed to identify the images in trash appropriately and as a result, the volume snapshot is restored successfully even when the parent of the snapshot does not exist.

(BZ#2264900)

**Incorrect warning logs from fuserecovery.go even when FUSE mount is not used**

Previously, the warning logs from the fuse recovery functions, such as **fuserecovery.go** were logged even when the kernel mounter was chosen and that was misleading.

With this fix, the fuse recovery functions are attempted or called only when the fuse mounter is chosen and as a result, the logs from **fuesreovery.go** are not logged when the kernel mounter is chosen.

(BZ#2266237)

## 6.4. OCS OPERATOR

**StorageClasses are not created if the RGW endpoint is not reachable**

Previously, storage classes were dependent on RADOS gateway (RGW) storage class creation as RADOS Block Device (RBD) and CephFS storage classes were not created if the RGW endpoint was not reachable.

With this fix, the storage class creation is made independent and as a result storage classes are no longer dependent on RGW storage class creation.

(BZ#2213757)

## 6.5. OPENSHIFT DATA FOUNDATION CONSOLE

**Status card reflects the status of standalone MCG deployment**

Previously, Multicloud Object Gateway (MCG) standalone mode was not showing any health status in OpenShift cluster Overview dashboard and an unknown icon was seen for Storage.

With this fix, MCG health metrics are pushed when the cluster is deployed in standalone mode and as a result the storage health is shown in cluster Overview dashboard.

(BZ#2256563)

**Create StorageSystem wizard overlaps Project dropdown**

Previously, the unused **Project** dropdown on top of the **Create StorageSystem** page caused confusion and was not used in any scenario.

With this fix, the **Project** dropdown is removed and as a result the StorageSystem creation namespace is populated in the header of the page.

(BZ#2271593)

**Capacity and Utilization cards do not include custom storage classes**

Previously, the Requested capacity and Utilization cards displayed only data for the default storage classes created by the OCS operator as part of the storage system creation. The cards do not include any custom storage classes that were created later. This was due to the refactoring of the prometheus to support multiple storage clusters.

With this fix, the queries are updated and the cards now show report capacity for both default and custom created storage classes.

(BZ#2284090)

## 6.6. ROOK

**Rook-Ceph operator deployment fail when storage class device sets are deployed with duplicate names**

Previously, when StorageClassDeviceSets were added into the StorageCluster CR with duplicate names, the OSDs failed leaving Rook confused about the OSD configuration.

With this fix, if the duplicate device set names are found in the CR, Rook refuses to reconcile the OSDs until it is fixed. An error is seen in the rook operator log about failing to reconcile the OSDs.

(BZ#2259209)

**Rook-ceph-mon pods listen to both 3300 and 6789 port**

Previously, when a cluster was deployed with MSGRv2, the mon pods were listening unnecessarily on port 6789 for MSGR1 traffic.

With this fix, the mon daemons start with flags to suppress listening on the v1 port 6789 and only listen exclusively on the v2 port 3300 thereby reducing the attack surface area.

(BZ#2262134)

**Legacy LVM-based OSDs are in crashloop state**

Previously, starting from OpenShift Data Foundation 4.14, the legacy OSDs were crashing in the init container that resized the OSD. This was because the legacy OSDs that were created in OpenShift Container Storage 4.3 and since upgraded to a future version might have failed.

With this fix, the crashing resize init container was removed from the OSD pod spec. As a result, the legacy OSD starts, however it is recommended that the legacy OSDs are replaced soon.

(BZ#2273398) (BZ#2274757)

## 6.7. CEPH MONITORING

**Quota alerts overlapping**

Previously, redundant alerts were fired when object bucket claim (OBC) quota limit was reached. This is because when OBC quota reached 100%, both **ObcQuotaObjectsAlert** (when OBC object quota crosses 80% of its limit) and **ObcQuotaObjectsExhausedAlert** (when quota reaches 100%) alerts were fired.

With this fix, the queries of the alerts were changed to make sure that only one alert is triggered at a time indicating the issue. As a result, when the quota crosses 80%, **ObcQuotaObjectsAlert** is triggered and when quota is at 100%, **ObcQuotaObjectsExhausedAlert** is triggered.

(BZ#2257949X)

**PrometheusRule evaluation failing for pool-quota rule**

Previously, none of the Ceph pool quota alerts were displayed because in a multi-cluster setup, **PrometheusRuleFailures** alert was fired due to **pool-quota** rules. The queries in the **pool-quota** section were unable to distinguish the cluster from which the alert was fired in a multi-cluster setup.

With this fix, a **managedBy** label was added to all the queries in the **pool-quota** to generate unique results from each cluster. As a result, **PrometheusRuleFailures** alert is no longer seen and all the alerts in **pool-quota** work as expected.

(BZ#2262943)

**Wrong help text shown in runbooks for some alerts**

Previously, wrong help text was shown in the runbooks for some alerts as there was wrong text in runbook markdown files of those alerts.

With this fix, the text in the runbook markdown files is corrected so that the alerts show the correct help text.

(BZ#2265492)

**PrometheusRuleFailures alert after installation or upgrade**

Previously, Ceph quorum related alerts were not seen as prometheus failure alert, **PrometheusRuleFailures** was fired, which is usually fired when the queries produced ambiguous results. In a multi-cluster scenario, queries in **quorum-alert** rules were giving indistinguishable results, as it could not identify from which cluster the quorum alerts were fired.

With this fix, a unique **managedBy** label was added to each query in quorum rules so that the query results contained the data about the cluster name from which the result was received. As a result, prometheus failure is not fired and the clusters are able to trigger all the Ceph mon quorum related alerts.

(BZ#2266316)

**Low default interval duration for two ServiceMonitors,rook-ceph-exporter and rook-ceph-mgr**

Previously, the exporter data collected by prometheus added load to the system as the prometheus scrapePVC interval provided for service monitors, **rook-ceph-exporter** and **rook-ceph-mgr** was only 5 seconds.

With this fix, the interval is increased to 30 seconds to balance the prometheus scrapping, thereby reducing the system load.

(BZ#2269354)

**Alert when there are LVM backed legacy OSDs during upgrade**

Previously, when OpenShift Data Foundation with legacy OSDs was upgraded from version 4.12 to 4.14, it was noticed that all the OSDs were stuck in a crash loop and down. This led to potential data unavailability and service disruption.

With this fix, a check is included to detect legacy OSDs based on local volume manager (LVM) and to alert if such OSDs are present during the upgrade process. As a result, a warning is displayed during upgrade to indicate about the legacy OSDs so that appropriate actions can be taken.

(BZ#2279928)

# CHAPTER 7. KNOWN ISSUES

This section describes the known issues in Red Hat OpenShift Data Foundation 4.16.

## 7.1. DISASTER RECOVERY

- **Creating an application namespace for the managed clusters**
  Application namespace needs to exist on RHACM managed clusters for disaster recovery (DR) related pre-deployment actions and hence is pre-created when an application is deployed at the RHACM hub cluster. However, if an application is deleted at the hub cluster and its corresponding namespace is deleted on the managed clusters, they reappear on the managed cluster.

  Workaround: **openshift-dr** maintains a namespace **manifestwork** resource in the managed cluster namespace at the RHACM hub. These resources need to be deleted after the application deletion. For example, as a cluster administrator, execute the following command on the hub cluster:

  ```
  $ oc delete manifestwork -n <managedCluster namespace> <drPlacementControl name>-<namespace>-ns-mw
  ```

  (BZ#2059669)

- **ceph df reports an invalid MAX AVAIL value when the cluster is in stretch mode**
  When a crush rule for a Red Hat Ceph Storage cluster has multiple "take" steps, the **ceph df** report shows the wrong maximum available size for the map. The issue will be fixed in an upcoming release.

  (BZ#2100920)

- **Both the DRPCs protect all the persistent volume claims created on the same namespace**
  The namespaces that host multiple disaster recovery (DR) protected workloads, protect all the persistent volume claims (PVCs) within the namespace for each DRPlacementControl resource in the same namespace on the hub cluster that does not specify and isolate PVCs based on the workload using its **spec.pvcSelector** field.

  This results in PVCs that match the DRPlacementControl **spec.pvcSelector** across multiple workloads. Or, if the selector is missing across all workloads, replication management to potentially manage each PVC multiple times and cause data corruption or invalid operations based on individual DRPlacementControl actions.

  Workaround: Label PVCs that belong to a workload uniquely, and use the selected label as the DRPlacementControl **spec.pvcSelector** to disambiguate which DRPlacementControl protects and manages which subset of PVCs within a namespace. It is not possible to specify the **spec.pvcSelector** field for the DRPlacementControl using the user interface, hence the DRPlacementControl for such applications must be deleted and created using the command line.

  Result: PVCs are no longer managed by multiple DRPlacementControl resources and do not cause any operation and data inconsistencies.

  (BZ#2128860)

- **MongoDB pod is in CrashLoopBackoff because of permission errors reading data in cephrbd volume**

The OpenShift projects across different managed clusters have different security context constraints (SCC), which specifically differ in the specified UID range and/or **FSGroups**. This leads to certain workload pods and containers failing to start post failover or relocate operations within these projects, due to filesystem access errors in their logs.

Workaround: Ensure workload projects are created on all managed clusters with the same project-level SCC labels, allowing them to use the same filesystem context when failed over or relocated. Pods will no longer fail post-DR actions on filesystem-related access errors.

(BZ#2081855)

- **Disaster recovery workloads remain stuck when deleted**
  When deleting a workload from a cluster, the corresponding pods might not terminate with events such as **FailedKillPod**. This might cause delay or failure in garbage collecting dependent DR resources such as the **PVC**, **VolumeReplication**, and **VolumeReplicationGroup**. It would also prevent a future deployment of the same workload to the cluster as the stale resources are not yet garbage collected.

  Workaround: Reboot the worker node on which the pod is currently running and stuck in a terminating state. This results in successful pod termination and subsequently related DR API resources are also garbage collected.

  (BZ#2159791)

- **Regional DR CephFS based application failover show warning about subscription**
  After the application is failed over or relocated, the hub subscriptions show up errors stating, "Some resources failed to deploy. Use View status YAML link to view the details." This is because the application persistent volume claims (PVCs) that use CephFS as the backing storage provisioner, deployed using Red Hat Advanced Cluster Management for Kubernetes (RHACM) subscriptions, and are DR protected are owned by the respective DR controllers.

  Workaround: There are no workarounds to rectify the errors in the subscription status. However, the subscription resources that failed to deploy can be checked to make sure they are PVCs. This ensures that the other resources do not have problems. If the only resources in the subscription that fail to deploy are the ones that are DR protected, the error can be ignored.

  (BZ-2264445)

- **Disabled PeerReady flag prevents changing the action to Failover**
  The DR controller executes full reconciliation as and when needed. When a cluster becomes inaccessible, the DR controller performs a sanity check. If the workload is already relocated, this sanity check causes the **PeerReady** flag associated with the workload to be disabled, and the sanity check does not complete due to the cluster being offline. As a result, the disabled **PeerReady** flag prevents you from changing the action to Failover.

  Workaround: Use the command-line interface to change the DR action to Failover despite the disabled **PeerReady** flag.

  (BZ-2264765)

- **Ceph becomes inaccessible and IO is paused when connection is lost between the two data centers in stretch cluster**
  When two data centers lose connection with each other but are still connected to the Arbiter node, there is a flaw in the election logic that causes an infinite election between the monitors. As a result, the monitors are unable to elect a leader and the Ceph cluster becomes unavailable. Also, IO is paused during the connection loss.

Workaround: Shut down the monitors in one of the data centers where monitors are out of quorum (you can find this by running **ceph -s** command) and reset the connection scores of the remaining monitors.

As a result, monitors can form a quorum and Ceph becomes available again and IOs resume.

(Partner BZ#2265992)

- **RBD applications fail to Relocate when using stale Ceph pool IDs from replacement cluster**
  For the applications created before the new peer cluster is created, it is not possible to mount the RBD PVC because when a peer cluster is replaced, it is not possible to update the CephBlockPoolID's mapping in the CSI configmap.

  Workaround: Update the **rook-ceph-csi-mapping-config** configmap with cephBlockPoolID's mapping on the peer cluster that is not replaced. This enables mounting the RBD PVC for the application.

  (BZ#2267731)

- **Information about lastGroupSyncTime is lost after hub recovery for the workloads which are primary on the unavailable managed cluster**
  Applications that are previously failed over to a managed cluster do not report a **lastGroupSyncTime**, thereby causing the trigger of the alert **VolumeSynchronizationDelay**. This is because when the ACM hub and a managed cluster that are part of the DRPolicy are unavailable, a new ACM hub cluster is reconstructed from the backup.

  Workaround: If the managed cluster to which the workload was failed over is unavailable, you can still failover to a surviving managed cluster.

  (BZ#2275320)

- **MCO operator reconciles the veleroNamespaceSecretKeyRef and CACertificates fields**
  When the OpenShift Data Foundation operator is upgraded, the **CACertificates** and **veleroNamespaceSecretKeyRef** fields under **s3StoreProfiles** in the Ramen config are lost.

  Workaround: If the Ramen config has the custom values for the **CACertificates** and **veleroNamespaceSecretKeyRef** fields, then set those custom values after the upgrade is performed.

  (BZ#2277941)

- **Instability of the token-exchange-agent pod after upgrade**
  The **token-exchange-agent** pod on the managed cluster is unstable as the old deployment resources are not cleaned up properly. This might cause application failover action to fail.

  Workaround: Refer the knowledgebase article, "token-exchange-agent" pod on managed cluster is unstable after upgrade to ODF 4.16.0.

  Result: If the workaround is followed, "token-exchange-agent" pod is stabilized and failover action works as expected.

  (BZ#2293611)

- **virtualmachines.kubevirt.io resource fails restore due to mac allocation failure on relocate**
  When a virtual machine is relocated to the preferred cluster, it might fail to complete relocation due to unavailability of the mac address. This happens if the virtual machine is not fully cleaned up on the preferred cluster when it is failed over to the failover cluster.

Ensure that the workload is completely removed from the preferred cluster before relocating the workload.

(BZ#2295404)

- **Post hub recovery, subscription app pods are not coming up after Failover**
  Post hub recovery, the subscription application pods do not come up after failover from primary to the secondary managed clusters. RBAC error occurs in **AppSub** subscription resource on managed cluster. This is due to a timing issue in the backup and restore scenario. When application-manager pod is restarted on each managed cluster, the hub subscription and channel resources are not recreated in the new hub. As a result, the child **AppSub** subscription resource is reconciled with an error.

  Workaround:

  Fetch the name of the **appsub** using the following command:

  ```
  % oc get appsub -n <namespace of sub app>
  ```

  Add a new label with any value to the AppSub on the hub using the following command:

  ```
  % oc edit appsub -n <appsub-namespace> <appsub>-subscription-1
  ```

  In case the child appsub error still exists showing unknown certificate issue, restart the application-manager pod on the managed cluster to which the workloads are failedover.

  ```
  % oc delete pods -n open-cluster-management-agent-addon application-manager-<>-<>
  ```

  (BZ#2295782)

## 7.2. MULTICLOUD OBJECT GATEWAY

- **Multicloud Object Gateway instance fails to finish initialization**
  Due to a race in timing between the pod code run and OpenShift loading the Certificate Authority (CA) bundle into the pod, the pod is unable to communicate with the cloud storage service. As a result, the default backing store cannot be created.

  Workaround: Restart the Multicloud Object Gateway (MCG) operator pod:

  ```
  $ oc delete pod noobaa-operator-<ID>
  ```

  With the workaround the backing store is reconciled and works.

  (BZ#2271580)

- **Upgrade to OpenShift Data Foundation 4.16 results in noobaa-db pod CrashLoopBackOff state**
  Upgrading to OpenShift Data Foundation 4.16 from OpenShift Data Foundation 4.15 fails when the PostgreSQL upgrade fails in Multicloud Object Gateway which always start with PostgresSQL version 15. If there is a PostgreSQL upgrade failure, the **NooBaa-db-pg-0** pod fails to start.

  Workaround: Refer to the knowledgebase article Recover NooBaa's PostgreSQL upgrade failure in OpenShift Data Foundation 4.16.

(BZ#2298152)

## 7.3. CEPH

- **Poor performance of the stretch clusters on CephFS**
  Workloads with many small metadata operations might exhibit poor performance because of the arbitrary placement of metadata server (MDS) on multi-site Data Foundation clusters.

  (BZ#1982116)

- **SELinux relabelling issue with a very high number of files**
  When attaching volumes to pods in Red Hat OpenShift Container Platform, the pods sometimes do not start or take an excessive amount of time to start. This behavior is generic and it is tied to how SELinux relabelling is handled by the Kubelet. This issue is observed with any filesystem based volumes having very high file counts. In OpenShift Data Foundation, the issue is seen when using CephFS based volumes with a very high number of files. There are different ways to workaround this issue. Depending on your business needs you can choose one of the workarounds from the knowledgebase solution https://access.redhat.com/solutions/6221251.

  (Jira#3327)

- **Ceph reports no active mgr after workload deployment**
  After workload deployment, Ceph manager loses connectivity to MONs or is unable to respond to its liveness probe.

  This causes the OpenShift Data Foundation cluster status to report that there is "no active mgr". This causes multiple operations that use the Ceph manager for request processing to fail. For example, volume provisioning, creating CephFS snapshots, and others.

  To check the status of the OpenShift Data Foundation cluster, use the command **oc get cephcluster -n openshift-storage**. In the status output, the **status.ceph.details.MGR_DOWN** field will have the message "no active mgr" if your cluster has this issue.

  Workaround: Restart the Ceph manager pods using the following commands:

  ```
  # oc scale deployment -n openshift-storage rook-ceph-mgr-a --replicas=0
  ```

  ```
  # oc scale deployment -n openshift-storage rook-ceph-mgr-a --replicas=1
  ```

  After running these commands, the OpenShift Data Foundation cluster status reports a healthy cluster, with no warnings or errors regarding **MGR_DOWN**.

  (BZ#2244873)

## 7.4. CSI DRIVER

- **Automatic flattening of snapshots is not working**
  When there is a single common parent RBD PVC, if volume snapshot, restore, and delete snapshot are performed in a sequence more than 450 times, it is further not possible to take volume snapshot or clone of the common parent RBD PVC.

  To workaround this issue, instead of performing volume snapshot, restore, and delete snapshot in a sequence, you can use PVC to PVC clone to completely avoid this issue.

If you hit this issue, contact customer support to perform manual flattening of the final restored PVCs to continue to take volume snapshot or clone of the common parent PVC again.

(BZ#2232163)

## 7.5. OPENSHIFT DATA FOUNDATION CONSOLE

- **Optimize DRPC creation when multiple workloads are deployed in a single namespace**
When multiple applications refer to the same placement, then enabling DR for any of the applications enables it for all the applications that refer to the placement.

  If the applications are created after the creation of the DRPC, the PVC label selector in the DRPC might not match the labels of the newer applications.

  Workaround: In such cases, disabling DR and enabling it again with the right label selector is recommended.

  (BZ#2294704)

- **Last snapshot synced is missing for appset based applications on the DR monitoring dashboard**
ApplicationSet type applications do not display last volume snapshot sync time on the monitoring dashboard.

  Workaround: Go to **Applications** navigation under ACM perspective and filter the desired application from the list. Then from the **Data policy** column (popover) note the "Sync status".

  (BZ#2295324)

## 7.6. OCS OPERATOR

- **Incorrect unit for the`ceph_mds_mem_rss` metric in the graph**
When you search for the **ceph_mds_mem_rss** metrics in the OpenShift user interface (UI), the graphs show the y-axis in Megabytes (MB), as Ceph returns **ceph_mds_mem_rss** metric in Kilobytes (KB). This can cause confusion while comparing the results for the **MDSCacheUsageHigh** alert.

  Workaround: Use **ceph_mds_mem_rss * 1000** while searching this metric in the OpenShift UI to see the y-axis of the graph in GB. This makes it easier to compare the results shown in the **MDSCacheUsageHigh** alert.

  (BZ#2261881)

- **Increasing MDS memory is erasing CPU values when pods are in CLBO state**
When the metadata server (MDS) memory is increased while the MDS pods are in a crash loop back off (CLBO) state, CPU request or limit for the MDS pods is removed. As a result, the CPU request or the limit that is set for the MDS changes.

  Workaround: Run the **oc patch** command to adjust the CPU limits.

  For example:

  ```
  $ oc patch -n openshift-storage storagecluster ocs-storagecluster \
    --type merge \
    --patch '{"spec": {"resources": {"mds": {"limits": {"cpu": "3"},
  ```

```
"requests": {"cpu": "3"}}}}}'
```

(BZ#2265563)

## 7.7. NON-AVAILABILITY OF IBM Z PLATFORM

IBM Z platform is not available with OpenShift Data foundation 4.16 release. IBM Z will be available with full features and functionality in an upcoming release.

(BZ#2279527)