Red Hat OpenShift Data Foundation 4.15

4.15 Release Notes

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

Last Updated: 2024-05-13

Red Hat OpenShift Data Foundation 4.15 4.15 Release Notes

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

Legal Notice

Abstract

The release notes for Red Hat OpenShift Data Foundation 4.15 summarizes all new features and enhancements, notable technical changes, and any known bugs upon general availability.

Table of Contents

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 4.15 (RHSA-2024:1383) is now available. New enhancements, features, and known issues that pertain to OpenShift Data Foundation 4.15 are included in this topic.

Red Hat OpenShift Data Foundation 4.15 is supported on the Red Hat OpenShift Container Platform version 4.15. 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 .

CHAPTER 2. NEW FEATURES

This section describes new features introduced in Red Hat OpenShift Data Foundation 4.15.

2.1. SUPPORT FOR MULTIPLE STORAGE CLUSTERS

Red Hat OpenShift Data Foundation provides the ability to deploy two storage clusters, one in internal mode and the other in external mode. The first cluster must be installed in internal mode in the **openshift-storage** namespace and the second cluster in external mode in the **openshift-storage-extended** namespace. Vice-versa is currently not supported.

For more information, see the Deploying multiple OpenShift Data Foundation storage clusters .

2.2. NON RESILIENT STORAGE CLASS

OpenShift Data Foundation allows the addition and use of a new non resilient replica-1 storage class. This helps to avoid redundant data copies and enables resilient management at the application level.

For more information, see the deployment guide for your platform in the Storage class with single replica.

2.3. RECOVERING TO A REPLACEMENT CLUSTER FOR METRO-DR

When a primary or a secondary cluster of Metro-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 .

2.4. OPENSHIFT VIRTUALIZATION WORKLOADS FOR METRO-DR

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

For more information, see the knowledgebase article, Use ODF Metro DR to protect ACM applications containing Virtual Machines in OpenShift.

2.5. SUPPORT SETTING OF RBD STORAGE CLASS AS THE DEFAULT

The Ceph RADOS block device (RBD) storage class can be set as the default storage class during the deployment of OpenShift Data Foundation on bare metal and IBM Power platforms. This helps to avoid manual annotation of the storage cluster when it is required to set Ceph RBD as the default storage class. In addition, it helps to avoid the confusion of selecting the correct storage class.

For more information, see Creating OpenShift Data Foundation cluster on bare metal and Creating OpenShift Data Foundation cluster on IBM Power.

2.6. PERFORMANCE PROFILES

OpenShift Data Foundation provides an option to choose a resource profile based on the availability of resources during deployment. The performance profile helps to obtain enhanced performance levels. The following performance profiles can be configured both during deployment and post deployment:

- Lean - To be used in a resource constrained environment with minimum resources that are lower than the recommended. This profile minimizes resource consumption by allocating fewer CPUs and less memory.
- Balanced - To be used when recommended resources are available. This profile provides a balance between resource consumption and performance for diverse workloads.
- Performance - To be used in an environment with sufficient resources to get the best performance. This profile is tailored for high performance by allocating ample memory and CPUs to ensure optimal execution of demanding workloads.

For more information, see the deployment guide for your platform in the OpenShift Data Foundation documentation.

2.7. ABILITY TO CREATE BACKING STORES IN OPENSHIFT CLUSTER THAT USE AWS SECURITY TOKEN SERVICE

OpenShift Data Foundation can be deployed on an OpenShift cluster that has the Amazon Web Services security token service (AWS STS) enabled and then backing stores of type **aws-sts-s3** can be created using the Multicloud Object Gateway command-line interface.

For more information, see Creating an AWS-STS-backed backingstore .

2.8. RUNBOOKS FOR OPENSHIFT DATA FOUNDATION ALERTS

OpenShift Data Foundation alerts include runbooks that provide guidance to fix problems on clusters that are surfaced by alerts. Alerts displayed in OpenShift Data Fondation have links to the corresponding runbooks.

2.9. ALLOW EXPANSION OF ENCRYPTED RBD VOLUMES

With this release, the expansion of the encrypted RADOS block device (RBD) volume feature is generally available. This feature provides resize capability for encrypted RBD persistent volume claims (PVCs).

For more information, see the knowledgebase article Enabling resize for encrypted RBD PVC.

2.10. IMPROVED CLUSTER AVAILABILITY WITH ADDITIONAL MONITOR DAEMON COMPONENTS

OpenShift Data Foundation provides the ability to configure up to five Ceph monitor daemon components in an internal mode deployment based on the number of racks or zones when there are three, five, or more number of failure domains present in the deployment. Ceph monitor count can be increased to improve the availability of the cluster.

For more information, see Resolving low Ceph monitor count alert .

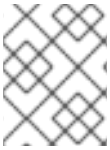
2.11. ALERTS FOR MONITORING SYSTEM OVERLOAD

OpenShift Data Foundation 4.15 introduces three new alerts to monitor the system that is getting overloaded. The new alerts are **OSDCPULoadHigh**, **MDSCPUUsageHigh**, and **MDSCacheUsageHigh**. These alerts improve the visibility to the current system performance and suggest tuning it when needed.

For more information, see Resolving cluster alerts.

2.12. SHALLOW VOLUMES SUPPORT FOR SNAPSHOT OR CLONE

With this release, PVC creation from snapshot functionality in OpenShift Data Foundation supports shallow volumes. These shallow volumes act as a reference to the source subvolume snapshot with no actual new subvolume being created in CephFS. The supported access mode for the shallow volume is **ReadOnlyMany**. When such PVCs are mounted, it means that the respective CephFS subvolume snapshot is exposed to the workloads. These shallow volumes help to reduce the time and resources to create clones.



NOTE

It is not possible to take a snapshot of the ROX PVC and creating a ROX PVC clone from ROX PVC results in a pending state. This is an expected behavior.

2.13. SUPPORT FOR LOGICAL PARTITION (LPAR) DEPLOYMENT

OpenShift Data Foundation on IBM Z supports Logical Partition (LPAR) as one of the additional deployment methods.

CHAPTER 3. ENHANCEMENTS

This section describes the major enhancements introduced in Red Hat OpenShift Data foundation 4.15.

3.1. DEPLOYMENT OF ONE ACTIVE AND ONE STANDBY MGR PODS BY OCS OPERATOR

The **ocs-operator** now deploys two MGR pods by default, one active and one standby. This enhancement does not impact cluster resource requirements.

3.2. SUPPORT FOR CUSTOM TIMEOUTS FOR RECLAIM SPACE OPERATION

Custom timeout values can be set for the reclaim space operation to avoid the failure of the operation with the error **context deadline exceeded**. The error would occur depending on the RBD volume size and its data pattern.

For more information, see Enabling reclaim space operation using ReclaimSpaceJob .

3.3. MODULARIZED MUST-GATHER UTILITY

OpenShift Data Foundation **must-gather** utility can be run in a modular mode and collect only the resources that are required. This enhancement helps to avoid long duration of time taken to run **must-gather** in some environments as well as focus on the inspected components faster.

For more information, see Downloading log files and diagnostic information using must-gather .

3.4. PREHOOK TO MCG'S DATABASE POD TO GRACEFULLY FLUSH CACHES WHEN THE POD IS GOING DOWN

A prehook to Multicloud Object Gateway's database pod (DB pod) is added to gracefully flush the cache when the pod is going down. This graceful shutdown reduces the risk of corruption in the journal file of the DB when the DB pod is taken down in a planned manner. However, this is not applicable for the shutdowns through OpenShift node crash or such.

3.5. ALL CONTROLLER OPERATIONS TO REACH ONE CONTROLLER

When a CSI-driver provides the **CONTROLLER_SERVICE** capability, the sidecar tries to become the leader by obtaining a lease based on the name of the CSI-driver.

The Kubernetes CSI-Addons Operator tries to connect to the random CSI-Addons sidecar that is registered and try to make the RPC calls to the random sidecar. This can create a problem if the CSI-driver has implemented some internal locking mechanism or has some local cache for the lifetime of that instance.

The NetworkFence (and other CSI-Addons) operations are only sent to a CSI-Addons sidecar that has the **CONTROLLER_SERVICE** capability. There is a single leader for the CSI-Addons sidecars that support that, and the leader can be identified by the Lease object for the **CSI-drivername**.

3.6. ENHANCED DATA DISTRIBUTION FOR CEPHFS STORAGE CLASS

This feature enables the default subvolume groups of Container Storage Interface (CSI) to be **automatically** pinned to the ranks according to the default pinning configuration. This is useful when you have multiple active CephFS metadata servers (MDSs) in the cluster. This helps to better distribute the load across MDS ranks in stable and predictable ways.

3.7. ABILITY TO USE BLUESTORE-RDR AS OBJECT STORAGE DEVICE BACKING STORE

OpenShift Data Foundation provides the ability to use **bluestore-rdr** as the object storage device (OSD) backing store for the Brownfield customers. This **bluestore-rdr** has improved performance over bluestore backend store, which is important when the cluster is required to be used for Regional Disaster Recovery (RDR). Also, it is possible to migrate the OSDs to **bluestore-rdr** from the user interface.

CHAPTER 4. TECHNOLOGY PREVIEWS

This section describes the technology preview features introduced in Red Hat OpenShift Data Foundation 4.15 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. MULTICLOUD OBJECT GATEWAY TO USE EXTERNAL POSTGRESQL

Multicloud Object Gateway (MCG) is allowed to use an external PostgreSQL to get a high availability solution where the PostgreSQL pod is a single point of failure. This solution provides a way to use external PostgreSQL independent of OpenShift Data Foundation.

For more information, see the deployment guide for your platform in the OpenShift Data Foundation documentation.

4.2. DISASTER RECOVERY FOR BROWNFIELD DEPLOYMENTS

With this release, disaster recovery can be enabled for an existing cluster which was deployed using OpenShift Data Foundation 4.13 or earlier or using OpenShift Data Foundation 4.14 without the Disaster Recovery flag enabled.

For more information, see the knowledgebase article, OpenShift Data Foundation cluster migration: Optimizing the cluster for Regional-DR.

4.3. 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.

4.4. SUPPORT READWRITEONCEPOD ACCESS MODE

OpenShift Data Foundation introduces the new **ReadWriteOncePod** (RWOP) access mode. With this RWOP access mode, it can be ensured that only one pod across the whole cluster can read that PVC or write to it. This access mode can be used either from the YAML file or the command-line interface only.

4.5. SUPPORT FOR EFFICIENT SELINUX VOLUME RELABELING

OpenShift Data Foundation has introduced an efficient SELinux relabeling for **ReadWriteOncePod** access mode. This helps to reduce the time consumption when the volume is mounted on a remote filesystem such as CephFS and when there are many files on the volume. This operation is faster than labeling all the files and folders individually.

4.6. HUB RECOVERY SUPPORT FOR CO-SITUATED AND NEUTRAL SITE DEPLOYMENTS

The Metropolitan and 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 respective Metro-DR and Regional-DR chapters on *Hub recovery using Red Hat Advanced Cluster Management*.

CHAPTER 5. DEVELOPER PREVIEWS

This section describes the developer preview features introduced in Red Hat OpenShift Data Foundation 4.15.

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. MULTICLOUD OBJECT GATEWAY SUPPORT STS FOR CLIENTS

Multicloud Object Gateway (MCG) provides support to a security token service (STS) similar to the one provided by Amazon Web Services. To allow other users to assume the role of a certain user, it is necessary to assign a role configuration to the user using MCG command line interface.

For more information, see the knowledgebase article, Use the Multi-Cloud Object Gateway's Security Token Service to assume the role of another user.

5.2. SUPPORT RADOS NAMESPACE FOR EXTERNAL MODE

The RADOS block device (RBD) storage class created in the OpenShift Data Foundation cluster uses a namespace for provisioning storage instead of the complete pool. The newly created namespace has restricted permissions.

For more information, see the knowledgebase article, Adding RADOS namespace for external mode cluster.

5.3. OPENSHIFT DATA FOUNDATION DEPLOYED ACROSS THREE VSPHERE CLUSTERS WITH VSPHERE IPI

OpenShift Data Foundation supports OpenShift deployment stretched across vSphere installer provisioned infrastructure clusters managed by one vCenter. This support enables you to deploy OpenShift Container Platform and OpenShift Data Foundation across Availability Zones (AZ) with each replica having affinity to an AZ. This helps to survive failure of any single zone as a minimum of three zones are required for the deployment.

For more information, see Installing a cluster on vSphere with customizations .

5.4. USER CAPABILITIES FOR CEPHOBJECTSTOREUSER

With this release, user capabilities (caps) for RADOS gateway (RGW) by using **CephObjectStore** CRD is supported. Enabling these caps such as user, bucket, and so on gives administrator like capabilities through REST API similar to **radosgw-admin** commands.

For more information, see the knowledgeabase article, User capabilities in CephObjectStoreUser .

5.5. CEPH-CSI BUILT-IN CAPABILITY TO FIND AND CLEAN STALE SUBVOLUMES

OpenShift Data Foundation 4.15 introduces an inbuilt script to delete stale volumes on a OpenShift Data Foundation cluster that have RADOS block device (RBD) images or CephFS subvolumes without the parent PVC.

For more information, see the knowledgebase article, Listing and cleaning stale subvolumes.

5.6. COMPLETE BUCKET POLICY ELEMENTS IN MULTICLOUD OBJECT GATEWAY

With this release, bucket policies can be updated to allow lists in Multicloud Object Gateway. For example, policy definition created for a bucket can be such that *read* access is granted to all the directories whereas only one specific directory has the *write* access.

For more information, see the knowledgebase article, Support for additional elements to the S3 BucketPolicy in Multicloud Object Gateway.

5.7. RECOVERY TO REPLACEMENT CLUSTER WITH REGIONAL DR

When there is a failure with the primary cluster, options are either to repair or wait for the recovery of the existing cluster or replace the cluster entirely if the cluster is irredeemable. The failed primary cluster can be replaced with a new cluster and fallback (relocate) can be enabled to this new cluster.

For more information, contact Red Hat Customer Support.

5.8. SUPPORT IPV6 FOR EXTERNAL MODE

With this release, IPv6 is supported in Openshift Data foundation external mode deployments.

CHAPTER 6. BUG FIXES

This section describes the notable bug fixes introduced in Red Hat OpenShift Data Foundation 4.15.

6.1. DISASTER RECOVERY

Fencing takes more time than expected

Previously, fencing operations took more time than expected. This was due to reconcile of Ramen hub controller a couple of times and requeue with delay as extra checks were added to ensure that the fencing operation was complete on the managed cluster.

With this fix, the hub controller is registered for the updates in fencing state. As a result, the updates of the fencing status change is received immediately and it takes less time to finish fencing operation.

(BZ#2249462)

6.2. MULTICLOUD OBJECT GATEWAY

Multicloud Object Gateway failing to use the new internal certificate after rotation

Previously, Multicloud Object Gateway (MCG) client was not able to connect to S3 using the new certificate unless the MCG endpoint pods were restarted. Even though the MCG endpoint pods were loading the certificate for the S3 service at the start of the pod, the changes in the certificate were not watched, which means that rotating a certificate was not affecting the endpoint till the pods were restarted.

With this fix, a watch to check for the changes in certificate of the endpoint pods are added. As a result, the pods load the new certificate without the need for a restart.

(BZ#2237903)

Regenerating S3 credentials for OBC in all namespaces

Previously, the Multicloud Object Gateway command for **obc regenerate** did not have the flag **app-namespace**. This flag is available for the other object bucket claim (OBC) operations such as creation and deletion of OBC. With this fix, the **app-namespace** flag is added to the **obc generate** command. As a result, OBC regenerates S3 credentials in all namespaces.

(BZ#2242414)

Signature validation failure

Previously, in Multicloud Object Gateway, there was failure to verify signatures when operations fail as AWS's C++ software development kit (SDK) does not encode the "=" sign in signature calculations when it appears as a part of the key name.

With this fix, MCG's decoding of the path in the HTTP request is fixed to successfully verify the signature.

(BZ#2265288)

6.3. CEPH

Metadata server run out of memory and reports over-sized cache

Previously, metadata server (MDS) would run out of memory as the standby-replay MDS daemons would not trim their caches.

With this fix, the MDS trims its cache when in standby-replay. As a result MDS would not run out of memory.

(BZ#2141422)

Ceph is inaccessible after crash or shutdown tests are run

Previously, in a stretch cluster, when a monitor is revived and is in the probing stage for other monitors to receive the latest information such as **MonitorMap** or **OSDMap**, it is unable to enter **stretch_mode**. This prevents it from correctly setting the elector's **disallowed_leaders** list, which leads to the Monitors getting stuck in **election** and Ceph eventually becomes unresponsive.

With this fix, the marked-down monitors are unconditionally added to the **disallowed_leaders** list. This fixes the problem of newly revived monitors having different **disallowed_leaders** set and getting stuck in an election.

(BZ#2241937)

6.4. CEPH CONTAINER STORAGE INTERFACE (CSI)

Snapshot persistent volume claim in pending state

Previously, creation of readonlymany (ROX) CephFS persistent volume claim (PVC) from snapshot source failed when a pool parameter was present in the storage class due to a bug.

With this fix, the check for the pool parameter is removed as it is not required. As a result, creation of ROX CephFS PVC from a snapshot source will be successful.

(BZ#2248117)

6.5. OPENSHIFT DATA FOUNDATION CONSOLE

Incorrect tooltip message for the raw capacity card

Previously, the tooltip for the raw capacity card in the block pool page showed an incorrect message. With this fix, the tooltip content for the raw capacity card has been changed to display an appropriate message, "Raw capacity shows the total physical capacity from all the storage pools in the StorageSystem".

(BZ#2237895)

System raw capacity card not showing external mode StorageSystem

Previously, the System raw capacity card did not display Ceph external StorageSystem as the Multicloud Object Gateway (MCG) standalone and Ceph external StorageSystems were filtered out from the card.

With this fix, only the StorageSystems that do not report the total capacity as per the information reported by the **odf_system_raw_capacity_total_bytes** metric is filtered out. As a result, any StorageSystem that reports the total raw capacity is displayed on the System raw capacity card and only the StorageSystems that do not report the total capacity is not displayed in the card.

(BZ#2257441)

6.6. ROOK

Provisioning object bucket claim with the same bucket name

Previously, for the green field use case, creation of two object bucket claims (OBCs) with the same bucket name was successful from the user interface. Even though two OBCs were created, the second one pointed to invalid credentials.

With this fix, creation of the second OBC with the same bucket name is blocked and it is no longer possible to create two OBCs with the same bucket name for green field use cases.

(BZ#2228785)

Change of the parameter name for the Python script used in external mode deployment

Previously, while deploying OpenShift Data Foundation using Ceph storage in external mode, the Python script used to extract Ceph cluster details had a parameter name, **--cluster-name**, which could be misunderstood to be the name of the Ceph cluster. However, it represented the name of the OpenShift cluster that the Ceph administrator provided.

With this fix, the **--cluster-name** flag is changed to **--k8s-cluster-name`**. The legacy flag **--cluster-name** is also supported to cater to the upgraded clusters used in automation.

(BZ#2244609)

Incorrect pod placement configurations while detecting Multus Network Attachment Definition CIDRS

Previously, some OpenShift Data Foundation clusters failed where the network "canary" pods were scheduled on nodes without Multus cluster networks, as OpenShift Data Foundation did not process pod placement configurations correctly while detecting Multus Network Attachment Definition CIDRS.

With this fix, OpenShift Data Foundation was fixed to process pod placement for Multus network "canary" pods. As a result, network "canary" scheduling errors are no longer experienced.

(BZ#2249678)

Deployment strategy to avoid rook-ceph-exporter pod restart

Previously, the **rook-ceph-exporter** pod restarted multiple times on a freshly installed HCI cluster that resulted in crashing of the exporter pod and the Ceph health showing the WARN status. This was because restarting the exporter using **RollingRelease** caused a race condition resulting in crash of the exporter.

With this fix, the deployment strategy is changed to **Recreate**. As a result, exporter pods no longer crash and there is no more health WARN status of Ceph.

(BZ#2250995)

rook-ceph-rgw-ocs-storagecluster-cephobjectstore-a pod stuck in CrashLoopBackOff state

Previously, the **rook-ceph-rgw-ocs-storagecluster-cephobjectstore-a** pod was stuck in **CrashLoopBackOff** state as the RADOS Gateway (RGW) multisite zonegroup was not getting created and fetched, and the error handling was reporting wrong text.

With this release, the error handling bug in multisite configuration is fixed and fetching the zonegroup is improved by fetching it for a particular **rgw-realm** that was created earlier. As a result, the multisite

configuration and **rook-ceph-rgw-ocs-storagecluster-cephobjectstore-a** pod gets created successfully.

(BZ#2253185)

6.7. CEPH MONITORING

TargetDown alert reported for ocs-metrics-exporter

Previously, metrics endpoint of the **ocs-metrics-exporter** used to be unresponsive as persistent volume resync by **ocs-metrics-exporter** was blocked indefinitely.

With this fix, the blocking operations from persistent volume resync in **ocs-metrics-exporter** is removed and the metrics endpoint is responsive. Also, the **TargetDown** alert for **ocs-metrics-exporter** no longer appears.

(BZ#2168042)

Label references of object bucket claim alerts

Previously, label for the object bucket claim alerts was not displayed correctly as the format for the **label-template** was wrong. Also, a blank object bucket claim name was displayed and the description text was incomplete.

With this fix, the format is corrected. As a result, the description text is correct and complete with appropriate object bucket claim name.

(BZ#2188032)

Discrepancy in storage metrics

Previously, the capacity of a pool was reported incorrectly as a wrong metrics query was used in the Raw Capacity card in the Block Pool dashboard.

With this fix, the metrics query in the user interface is updated. As a result, the metrics of the total capacity of a block pool is reported correctly.

(BZ#2252035)

Add managedBy label to rook-ceph-exporter metrics and alerts

Previously, the metrics generated by **rook-ceph-exporter** did not have the **managedBy** label. So, it was not possible for the OpenShift console user interface to identify from which StorageSystem the metrics are generated.

With this fix, the **managedBy** label, which has the name of the StorageSystem as a value, is added through the OCS operator to the storage cluster's **Monitoring** spec. This spec is read by the Rook operator and it relabels the ceph-exporter's **ServiceMonitor** endpoint labels. As a result, all the metrics generated from this exporter will have the new label **managedBy**.

(BZ#2255491)

6.8. MUST GATHER

Must gather logs not collected after upgrade

Previously, the **must-gather** tool failed to collect logs after the upgrade as **Collection started <time>** was seen twice.

With this fix, the **must-gather** tool is updated to run the pre-install script only once. As a result, the tool is able to collect the logs successfully after upgrade.

(BZ#2255240)

CHAPTER 7. KNOWN ISSUES

This section describes the known issues in Red Hat OpenShift Data Foundation 4.15.

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)

- **When DRPolicy is applied to multiple applications under same namespace, volume replication group is not created**
  When a DRPlacementControl (DRPC) is created for applications that are co-located with other applications in the namespace, the DRPC has no label selector set for the applications. If any subsequent changes are made to the label selector, the validating admission webhook in the OpenShift Data Foundation Hub controller rejects the changes.

  Workaround: Until the admission webhook is changed to allow such changes, the DRPC **validatingwebhookconfigurations** can be patched to remove the webhook:

  ```
  $ oc patch validatingwebhookconfigurations vdrplacementcontrol.kb.io-lq2kz --type=json --patch='[{"op": "remove", "path": "/webhooks"}]'
  ```

  (BZ#2210762)

- **Application failover hangs in FailingOver state when the managed clusters are on different versions of OpenShift Container Platform and OpenShift Data Foundation**
  Disaster Recovery solution with OpenShift Data Foundation protects and restores persistent volume claim (PVC) data in addition to the persistent volume (PV) data. If the primary cluster is on an older OpenShift Data Foundation version and the target cluster is updated to 4.15 then the failover will be stuck as the S3 store will not have the PVC data.

  Workaround: When upgrading the Disaster Recovery clusters, the primary cluster must be upgraded first and then the post-upgrade steps must be run.

  (BZ#2215462)

- **Failover of apps from c1 to c2 cluster hang in FailingOver**
  The failover action is not disabled by Ramen when data is not uploaded to the s3 store due to s3 store misconfiguration. This means the cluster data is not available on the failover cluster during the failover. Therefore, failover cannot be completed.

Workaround: Inspect the Ramen logs after initial deployment to ensure there are no s3 configuration errors reported.

```
$ oc get drpc -o yaml
```

(BZ#2248723)

- **Potential risk of data loss after hub recovery**
  A potential data loss risk exists following hub recovery due to an eviction routine designed to clean up orphaned resources. This routine identifies and marks **AppliedManifestWorks** instances lacking corresponding **ManifestWorks** for collection. A hardcoded grace period of one hour is provided. After this period elapses, any resources associated with the **AppliedManifestWork** become subject to garbage collection.

  If the hub cluster fails to regenerate corresponding **ManifestWorks** within the initial one hour window, data loss could occur. This highlights the importance of promptly addressing any issues that might prevent the recreation of **ManifestWorks** post-hub recovery to minimize the risk of data loss.

- **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)

- **Cleanup and data synchronization for ApplicationSet workloads remain stuck after older primary managed cluster is recovered post the failover**
  ApplicationSet based workload deployments to the managed clusters are not garbage collected in cases when the hub cluster fails. It is recovered to a standby hub cluster while the workload has been failed over to a surviving managed cluster. The cluster that the workload failed over from, rejoins the new recovered standby hub.

  ApplicationSets that are disaster recovery (DR) protected and with a regional DRPolicy starts firing the **VolumeSynchronizationDelay** alert. Further such DR protected workloads cannot be failed over to the peer cluster or relocated to the peer cluster as data is out of sync between the two clusters.

  For a workaround, see the Troubleshooting section for Regional-DR in Configuring OpenShift Data Foundation Disaster Recovery for OpenShift Workloads.

  (BZ#2268594)

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, 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#2269379) and (BZ#2268429)

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)

- **CephBlockPool creation fails when custom deviceClass is used in StorageCluster**
  Due to a known issue, CephBlockPool creation fails when custom deviceClass is used in StorageCluster.

  (BZ#2248487)

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 restore PVCs to continue to take volume snapshot or clone of the common parent PVC again.

  (BZ#2232163)

7.5. OPENSHIFT DATA FOUNDATION CONSOLE

- **Missing NodeStageVolume RPC call blocks new pods from going into Running state**
  NodeStageVolume RPC call is not being issued blocking some pods from going into **Running** state. The new pods are stuck in **Pending** forever.

  To workaround this issue, scale down all the affected pods at once or do a node reboot. After applying the workaround, all pods should go into Running state.

  (BZ#2244353)

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)

CHAPTER 8. ASYNCHRONOUS ERRATA UPDATES

8.1. RHBA-2024:1708 OPENSHIFT DATA FOUNDATION 4.15.1 BUG FIXES AND SECURITY UPDATES

OpenShift Data Foundation release 4.15.1 is now available. The bug fixes that are included in the update are listed in the RHBA-2024:1708 advisory.

8.1.1. Documentation updates

Added new sections on related to hub recovery in *Configuring OpenShift Data Foundation Disaster Recovery for OpenShift Workloads* guide.

The section covers how to configure the 4th cluster for hub recovery to failover or relocate the disaster recovery protected workloads using Red Hat Advanced Cluster Management for Kubernetes (RHACM) in case where the active hub is down or unreachable. The hub recovery solution is a Technology Preview feature and is subject to Technology Preview support limitations. For more information, see Hub recovery support for co-situated and neutral site deployments.