Red Hat Ceph Storage 8

8.0 Release Notes

Release notes for Red Hat Ceph Storage 8.0

Last Updated: 2024-12-11

Red Hat Ceph Storage 8 8.0 Release Notes

Release notes for Red Hat Ceph Storage 8.0

Legal Notice

Abstract

The release notes describes the major features, enhancements, known issues, and bug fixes implemented for the Red Hat Ceph Storage 8.0 product release. 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.

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.

PROVIDING FEEDBACK ON RED HAT CEPH STORAGE DOCUMENTATION

We appreciate your input on our documentation. Please let us know how we could make it better. To do so, create a Bugzilla ticket:

+ . Go to the *Bugzilla* website. . In the Component drop-down, select **Documentation**. . In the Sub-Component drop-down, select the appropriate sub-component. . Select the appropriate version of the document. . Fill in the **Summary** and **Description** field with your suggestion for improvement. Include a link to the relevant part(s) of documentation. . Optional: Add an attachment, if any. . Click **Submit Bug**.

CHAPTER 1. INTRODUCTION

Red Hat Ceph Storage is a massively scalable, open, software-defined storage platform that combines the most stable version of the Ceph storage system with a Ceph management platform, deployment utilities, and support services.

The Red Hat Ceph Storage documentation is available at https://access.redhat.com/documentation/en-us/red_hat_ceph_storage/8.

CHAPTER 2. ACKNOWLEDGMENTS

Red Hat Ceph Storage version 8.0 contains many contributions from the Red Hat Ceph Storage team. In addition, the Ceph project is seeing amazing growth in the quality and quantity of contributions from individuals and organizations in the Ceph community. We would like to thank all members of the Red Hat Ceph Storage team, all of the individual contributors in the Ceph community, and additionally, but not limited to, the contributions from organizations such as:

- Intel®
- Fujitsu ®
- UnitedStack
- Yahoo ™
- Ubuntu Kylin
- Mellanox ®
- CERN ™
- Deutsche Telekom
- Mirantis ®
- SanDisk ™
- SUSE ®

CHAPTER 3. NEW FEATURES

This section lists all major updates, enhancements, and new features introduced in this release of Red Hat Ceph Storage.

3.1. THE CEPHADM UTILITY

High Availability can now be deployed for the Grafana, Prometheus, and Alertmanager monitoring stacks

With this enhancement, the cephadm **mgmt-gateway** service offers better reliability and ensures uninterrupted monitoring by allowing these critical services to function seamlessly, even during the event of an individual instance failure. High availability is crucial for maintaining visibility into the health and performance of the Ceph cluster and responding promptly to any issues.

Use High Availability for continuous, uninterrupted operations to improve the stability and resilience of the Ceph cluster.

For more information, see Using the Ceph Management gateway .

Bugzilla:1878777

New streamlining deployment for EC pools for Ceph Object Gateway

The Ceph Object Gateway manager module can now create pools for the **rgw** service. Within the pool, data pools can receive attributes, based on the provided specification.

This enhancement streamlines deployment for users who want Ceph Object Gateway pools used for Ceph Object Gateway to use EC instead of replica.

To create a data pool with the specified attributes, use the following command:

```
ceph rgw realm bootstrap -i <path-to-spec-file> --start-radosgw
```

Currently, the EC profile fields of this specification only make use of the **k**, **m**, **pg_num**, and **crush-device-class** attributes. If other attributes are set, or if the pool type is replicated, the key value pairs pass to the **ceph osd pool create** command. The other pools for the Ceph Object Gateway zone, for example, the buckets index pool, are all created as replicated pools with default settings.

Bugzilla:2249537

A self-signed certificate can be generated by cephadm within the Ceph Object Gateway service specification

With this enhancement, adding **generate_cert: true** into the Ceph Object Gateway service specification file, enables cephadm to generate a self-signed certificate for the Ceph Object Gateway service. This can be done instead of manually creating the certificate and inserting into the specification file.

Using **generate_cert: true** works for the Ceph Object Gateway service, including SAN modifications based on the **zonegroup_hostnames** parameter included in the Ceph Object Gateway specification file.

The following is an example of Ceph Object Gateway specification file:

```
service_type: rgw
service_id: bar
```

```
service_name: rgw.bar
placement:
  hosts:
  - vm-00
  - vm-02
spec:
  generate_cert: true
  rgw_realm: bar_realm
  rgw_zone: bar_zone
  rgw_zonegroup: bar_zonegroup
  ssl: true
  zonegroup_hostnames:
  - s3.example.com
  - s3.foo.com
```

This specification file would generate a self-signed certificate that includes the following output:

```
X509v3 Subject Alternative Name:
    DNS:s3.example.com, DNS:s3.foo.com
```

Bugzilla:2249984

Setting rgw_run_sync_thread to 'false' for Ceph Object gateway daemon users is now automated

With this enhancement, by setting **disable_multisite_sync_traffic** to 'true' under the **spec** section of an Ceph Object Gateway specification, Cephadm will handle setting the **rgw_run_sync_thread** setting to 'false' for Ceph Object Gateway daemons under that service. That will stop the Ceph Object Gateway daemons from spawning threads to handle the sync of data and metadata. The process of setting **rgw_run_sync_thread** to 'false' for Ceph Object Gateway daemon users is now automated through the Ceph Object Gateway specification file.

Bugzilla:2252263

Cephadm can now deploy ingress over Ceph Object Gateway with the ingress service's haproxy daemon in TCP rather than HTTP mode

Setting up **haproxy** in TCP mode allows encrypted messages to be passed through **haproxy** directly to Ceph Object Gateway without **haproxy** needing to understand the message contents. This allows end-to-end SSL for ingress and Ceph Object Gateway setups.

With this enhancement, users can now specify a certificate for the **rgw** service and not ingress service. Specify the **use_tcp_mode_over_rgw** as **True** in the ingress specification to get the **haproxy** daemons deployed for that service in TCP mode, rather than in HTTP mode.

Bugzilla:2260522

New cmount_path option with a unique user ID generated for CephFS

With this enhancement, you can add the optional **cmount_path** option and generate a unique user ID for each Ceph File System. Unique user IDs allow sharing CephFS clients across multiple Ganesha exports. Reducing clients across exports also reduces memory usage for a single CephFS client.

Bugzilla:2300273

Exports sharing the same FSAL block have a single Ceph user client linked to them

Previously, on an upgraded cluster, the export creation failed with the "Error EPERM: Failed to update caps" message.

With this enhancement, the user key generation is modified when creating an export so that any exports that share the same Ceph File System Abstraction Layer (FSAL) block will have only a single Ceph user client linked to them. This enhancement also prevents memory consumption issues in NFS Ganesha.

Bugzilla:2302911

3.2. CEPH DASHBOARD

Added health warnings for when daemons are down

Previously, there were no health warnings and alerts to notify if the **mgr**, **mds**, and **rgw** daemons were down.

With this enhancement, health warnings are emitted when any of the **mgr**, **mds**, and **rgw** daemons are down.

Bugzilla:2138386

Ceph Object Gateway NFS export management is now available through the Ceph Dashboard

Previously, Ceph Object Gateway NFS export management was only available through the command-line interface.

With this enhancement, the Ceph Dashboard also supports managing exports that were created based on selected Ceph Object Gateway users.

For more information about editing a bucket, see Managing NFS Ganesha exports on the Ceph dashboard.

Bugzilla:2221195

Enhanced multi-site creation with default realms, zones, and zone groups

Previously, a manual restart was required for Ceph Object Gateway services after creating a multi-site with default realms, zones, or zone groups.

With this enhancement, due to the introduction of the new multi-site replication wizard, any necessary service restarts are done automatically.

Bugzilla:2243175

Ceph Dashboard now supports an EC 8+6 profile

With this enhancement, the dashboard supports an erasure coding 8+6 profile.

Bugzilla:2293055

Enable or disable replication for a bucket in a multi-site configuration during bucket creation

A new Replication checkbox is added in the Ceph Object Gateway bucket creation form. This enhancement allows enabling or disabling replication from a specific bucket in a multi-site configuration.

Bugzilla:2298099

New sync policy management through the Ceph Dashboard

Previously, there was no way to manage sync policies from the Ceph Dashboard.

With this enhancement, you can now manage sync policies directly from the Ceph Dashboard, by going to Object>Multi-site.

Bugzilla:2298101

Improved experience for server-side encryption configuration on the Ceph Dashboard

With this enhancement, server-side encryption can easily be found by going to Objects>Configuration from the navigation menu.

Bugzilla:2298396

New option to enable mirroring on a pool during creation

Previously, there was no option to enable mirroring on a pool during pool creation.

With this enhancement, mirroring can be enabled on a pool directly from the Create Pool form.

Bugzilla:2298553

Enhanced output for Ceph Object Gateway ops and audit logs in the centralized logging

With this enhancement, you can now see Ceph Object Gateway ops and audit log recollection in the Ceph Dashboard centralized logs.

Bugzilla:2298554

Improved experience when creating an erasure coded pool with the Ceph Dashboard

Previously, devices in the Ceph cluster were automatically selected when creating an erasure coded (EC) profile, such as HDD, SSD, and so on. When a device class is specified with EC pools, pools are created with only one placement group and the autoscaler did not work.

With this enhancement, a device class has to be manually selected and all devices are automatically selected and available.

Bugzilla:2305949

Enhanced multi-cluster views on the Ceph Dashboard

Previously, the Ceph Cluster Grafana dashboard was not visible for a cluster that was connected in a multi-cluster setup and multi-cluster was not fully configurable with mTLS through the dashboard.

With these enhancements, users can connect a multi-cluster setup with mTLS enabled on both clusters. Users can also see individual cluster Grafana dashboards by expanding a particular cluster row, when going to Multi-Cluster > Manage Clusters.

Bugzilla:2311466

CephFS subvolume groups and subvolumes can now be selected directly from the Create NFS export form

Previously, when creating a CephFS NFS export, you would need to know the existing subvolume and subvolume groups prior to creating the NFS export, and manually enter the information into the form.

With the enhancement, once a volume is selected, the relevant subvolume groups and subvolumes are available to select seamlessly from inside the Create NFS export form.

Bugzilla:2312627

Non-default realm sync status now visible for Ceph Object Gateways

Previously, only the default realm sync status was visible in the Object>Overview sync status on the Ceph Dashboard.

With this enhancement, the sync status of any selected Object Gateway is displayed, even if it is in a non-default realm.

Bugzilla:2317967

New RGW Sync overview dashboard in Grafana

With this release, you can now track replication differences over a time per shard from within the new RGW Sync overview dashboard in Grafana.

Bugzilla:2298621

New S3 bucket lifecycle management through the Ceph Dashboard

With this release, a bucket lifecycle can be managed through the Edit Bucket form in the Ceph Dashboard.

For more information about editing a bucket, see Editing Ceph Object Gateway buckets on the dashboard.

Bugzilla:2248808

3.3. CEPH FILE SYSTEM

snapdiff API now only syncs the difference of files between two snapshots

With this enhancement, the **snapdiff** API is used to sync only the difference of files between two snapshots. Syncing only the difference avoids bulk copying during an incremental snapshot sync, providing performance improvement, as only **snapdiff** delta is being synced.

Bugzilla:2301912

New metrics for data replication monitoring logic

This enhancement provides added labeled metrics for the replication start and end notifications.

The new labeled metrics are: **last_synced_start**, **last_synced_end**, **last_synced_duration**, and **last_synced_bytes**.

Bugzilla:2303452

Enhanced output remote metadata information in peer status

With this enhancement, the peer status output shows **state**, **failed**, and 'failure_reason' when there is invalid metadata in a remote snapshot.

Bugzilla:2271767

New support for NFS-Ganesha async FSAL

With this enhancement, the non-blocking Ceph File System Abstraction Layer (FSAL), or async, is introduced. The FSAL reduces thread utilization, improves performance, and lowers resource utilization.

Bugzilla:2232674

New support for earmarking subvolumes

Previously, the Ceph storage system did not support a mixed protocol being used within the same subvolume. Attempting to use a mixed protocol could lead to data corruption.

With this enhancement, subvolumes have protocol isolation. The isolation prevents data integrity issues and reduces the complexity of managing multi-protocol environments, such as SMB and NFS.

Bugzilla:2312545

3.4. CEPH OBJECT GATEWAY

The CopyObject API can now be used to copy the objects across storage classes

Previously, objects could only be copied within the same storage class. This limited the scope of the **CopyObject** function. Users would have to download the objects and then reupload them to another storage class.

With this enhancement, the objects can be copied to any storage class within the same Ceph Object Gateway cluster from the server-side.

Bugzilla:1470874

Improved read operations for Ceph Object Gateway

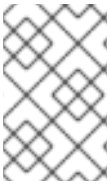
With this enhancement, read affinity is added to the Ceph Object Gateway. The read affinity allows read calls to the nearest OSD by adding the flags and setting the correct CRUSH location.

Bugzilla:2252291

S3 requests are no longer cut off in the middle of transmission during shutdown

Previously, a few clients faced issues with the S3 request being cut off in the middle of transmission during shutdown without waiting.

With this enhancement, the S3 requests can be configured (off by default) to wait for the duration defined in the **rgw_exit_timeout_secs** parameter for all outstanding requests to complete before exiting the Ceph Object Gateway process unconditionally. Ceph Object Gateway will wait for up to 120 seconds (configurable) for all on-going S3 requests to complete before exiting unconditionally. During this time, new S3 requests will not be accepted.



NOTE

In containerized deployments, an additional **extra_container_args** parameter configuration of **--stop-timeout=120** (or the value of **rgw_exit_timeout_secs** parameter, if not default) is also necessary.

Bugzilla:2298701

Copying of encrypted objects using copy-object APIs is now supported

Previously, in Ceph Object gateway, copying of encrypted objects using copy-object APIs was unsupported since the inception of its server-side encryption support.

With this enhancement, copying of encrypted objects using copy-object APIs is supported and workloads that rely on copy-object operations can also use server-side encryption.

Bugzilla:2300284

New S3 additional checksums

With this release, there is added support for S3 additional checksums. This new support provides improved data integrity for data in transit and at rest. The additional support enables use of strong checksums of object data, such as SHA256, and checksum assertions in S3 operations.

Bugzilla:2303759

New support for S3 **GetObjectAttributes** API

The **GetObjectAttributes** API returns a variety of traditional and non-traditional metadata about S3 objects. The metadata returns include S3 additional checksums on objects and on the parts of objects that were originally stored as multipart uploads. The **GetObjectAttributes** is exposed in the AWS CLI.

Bugzilla:2266243

Improved efficiency of Ceph Object Gateway clusters over multiple locations

With this release, if possible, data is now read from the nearest physical OSD instance in a placement group.

As a result, the local read improves the efficiency of Ceph Object Gateway clusters that span over multiple physical locations.

Bugzilla:2298523

Format change observed for tenant owner in the event record: **ownerIdentity –> principalId**

With this release, in bucket notifications, the **principalId** inside **ownerIdentity** now contains complete user ID, prefixed with tenant ID.

Bugzilla:2309013

Client IDs can be added and thumbprint lists can be updated in an existing OIDC Provider within Ceph Object Gateway

Previously, users were not able to add a new client ID or update the thumbprint list within the OIDC Provider.

With this enhancement, users can add a new client ID or update the thumbprint list within the OIDC Provider and any existing thumbprint lists are replaced.

Bugzilla:2237887

3.5. MULTI-SITE CEPH OBJECT GATEWAY

New multi-site configuration header

With this release, GetObject and HeadObject responses for objects written in a multi-site configuration include the **x-amz-replication-status: PENDING** header. After the replication succeeds, the header's value changes to **COMPLETED**.

Bugzilla:2212311

New notification_v2 zone feature for topic and notification metadata

With this enhancement, bucket notifications and topics that are saved in fresh installation deployments (Greenfield) have their information synced between zones.

When upgrading to Red Hat Ceph Storage 8.0, this enhancement needs to be added by enabling the notification_v2 feature.

Bugzilla:1945018

3.6. RADOS

Balanced primary placement groups can now be observed in a cluster

Previously, users could only balance primaries with the offline **osdmaptool**.

With this enhancement, autobalancing is available with the **upmap** balancer. Users can now choose between either the **upmap-read`or `read** mode. The **upmap-read** mode offers simultaneous upmap and read optimization. The **read** mode can only be used to optimize reads.

For more information, see Using the Ceph manager module .

Bugzilla:1870804

New MSR CRUSH rules for erasure encoded pools

Multi-step-retry (MSR) is a type of CRUSH rule in the Ceph cluster that defines how data is distributed across storage devices. MSR ensures efficient data retrieval, balancing, and fault tolerance.

With this enhancement, **crush-osds-per-failure-domain** and **crush-num-failure-domains** can now be specified for erasure-coded (EC) pools during their creation. These pools use newly introduced MSR crush rules to place multiple OSDs within each failure domain. For example, 14 OSDs split across 4 hosts.

For more information, see Ceph erasure coding.

Bugzilla:2227600

New generalized stretch cluster configuration for three availability zones

Previously, there was no way to apply a stretch peer rule to prevent placement groups (PGs) from becoming active when there weren't enough OSDs in the acting set from different buckets without enabling the stretch mode.

With a generalized stretch cluster configuration for three availability zones, three data centers are supported, with each site holding two copies of the data. This helps ensure that even during a data center outage, the data remains accessible and writeable from another site. With this configuration, the pool replication size is 6 and the pool min_size is 3.

For more information, see Generalized stretch cluster configuration for three availability zones

Bugzilla:2300170

3.7. RADOS BLOCK DEVICES (RBD)

Added support for live importing of an image from another cluster

With this enhancement, you can now migrate between different image formats or layouts from another Ceph cluster. When live migration is initiated, the source image is deep copied to the destination image, pulling all snapshot history while preserving the sparse allocation of data wherever possible.

For more information, see Live migration of images.

Bugzilla:2219736

New support for cloning images from non-user type snapshots

With this enhancement, there is added support for cloning Ceph Block Device images from snapshots of non-user types. Cloning new groups from group snapshots that are created with the **rbd group snap create** command is now supported with the added **--snap-id** option for the **rbd clone** command.

For more information, see Cloning a block device snapshot.

Bugzilla:2298566

New commands are added for Ceph Block Device

Two new commands were added for enhanced Ceph Block Device usage. The **rbd group info** command shows information about a group. The **rbd group snap info** command shows information about a group snapshot.

Bugzilla:2302137

New support for live importing an image from an NBD export

With this enhancement, images with encryption support live migration from an NBD export.

For more information, see Streams.

Bugzilla:2303651

3.8. RBD MIRRORING

New optional --remote-namespace argument for the rbd mirror pool enable command

With this enhancement, Ceph Block Device has the new optional **--remote-namespace** argument for the **rbd mirror pool enable** command. This argument provides the option for a namespace in a pool to be mirrored to a different namespace in a pool of the same name on another cluster.

Bugzilla:2307280

CHAPTER 4. BUG FIXES

This section describes bugs with significant impact on users that were fixed in this release of Red Hat Ceph Storage. In addition, the section includes descriptions of fixed known issues found in previous versions.

4.1. THE CEPHADM UTILITY

The original_weight field is added as an attribute for the OSD removal queue

Previously, cephadm osd removal queue did not have a parameter for original_weight. As a result, the cephadm module would crash during OSD removal. With this fix, the original_weight field is added as an attribute for the osd removal queue and the cephadm no longer crashes during OSD removal.

Bugzilla:2305678

Cephadm no longer randomly marks hosts offline during large deployments

Previously, in some cases, when Cephadm had a short command timeout on larger deployments, hosts would randomly be marked offline during the host checks.

With this fix, the short timeout is removed. Cephadm now relies on the timeout specified by **mgr/cephadm/default_cephadm_command_timeout** setting.

The **ssh_keepalive_interval** interval and **ssh_keepalive_count_max** settings are also now configurable through the **mgr/cephadm/ssh_keepalive_interval** and **mgr/cephadm/ssh_keepalive_count_max** settings.

These settings give users better control over how hosts are marked offline in their Cephadm managed clusters and Cephadm no longer randomly marks hosts offline during larger deployments.

Bugzilla:2308688

Custom webhooks are now specified under the custom-receiver receiver.

Previously, custom Alertmanager webhooks were being specified within the **default** receiver in the Alertmanager configuration file. As a result, custom alerts were not being sent to the specified webhook unless the alerts did not match any other receiver.

With this fix, custom webhooks are now specified under the **custom-receiver** receiver. Alerts are now sent to custom webhooks even if the alerts match another receiver.

Bugzilla:2313614

4.2. CEPH DASHBOARD

cherrypy no longer gets stuck during network security scanning

Previously, due to a bug in the cheroot package, **cherrypy** would get stuck during some security scans that were scanning the network. As a result, the Ceph Dashboard became unresponsive and needed to restart the mgr module.

With this fix, the cheroot package is updated and the issue is resolved.

Bugzilla:2184612

Zone storage class details now display the correct compression information

Previously, the wrong compression information was being set for zone details. As a result, the zone details under the storage classes section were showing incorrect compression information.

With this fix, the information is corrected for storage classes and the zone details now show the correct compression information.

Bugzilla:2252712

Zone storage class detail values are now set correctly

Previously, wrong data pool values were set for storage classes in zone details. As a result, data pool values were incorrect in the user interface when multiple storage classes were created.

With this fix, the correct values are being set for storage classes in zone details.

Bugzilla:2252732

_nogroup is now listed in the Subvolume Group list even if there are no subvolumes in _nogroup

Previously, while cloning subvolume, **_nogroup** subvolume group was not listed if there were no subvolumes in **_nogroup**. As a result, users were not able to select **_nogroup** as a subvolume group.

With this fix, while cloning a subvolume, **_nogroup** is listed in the **Subvolume Group** list, even if there are no subvolumes in **_nogroup**.

Bugzilla:2269104

Correct UID containing $ in the name is displayed in the dashboard

Previously, when a user was created through the CLI, the wrong UID containing **$** in the name was being displayed in the Ceph Dashboard.

With this fix, the correct UID is displayed, even if the user containing **$** in the name is created by using the CLI.

Bugzilla:2271812

NFS in File and Object now have separate routing

Previously, the same route was being used for both NFS in **File** and **NFS** in Object. This caused issues from a usability perspective, as both navigation links of NFS in **File** and **Object** got highlighted. The user was also required to select the storage backend for both views, **File** and **Object**.

With this fix, NFS in **File** and **Object** have separate routing and do not ask users to enter the storage backend, thereby improving usability.

Bugzilla:2301988

Validation for pseudo path and CephFS path is now added during NFS export creation

Previously, during the creation of NFS export, the psuedo path had to be entered manually. As a result, CephFS path could not be validated.

With this fix, the pseudo path field is left blank for the user to input a path and CephFS path gets the updated path from the selected subvolume group and subvolume. A validation is now also in place for invalid values added for CephFS path. If a user attempts to change the CephFS path to an invalid value,

the export creation fails.

Bugzilla:2302756

Users are now prompted to enter a path when creating an export

Previously, creating an export was not prompting for a path and by default / was entered.

With this fix, when attempting to create the export directly on the file system, it prompts for a path. If an invalid path is entered, creation is not permitted. Additionally, when entering the path of the CephFS file system directly, a warning appears stating "Export on CephFS volume '/' not allowed".

Bugzilla:2303196

Bugzilla:2303655

Snapshots containing "." and "/" in the name cannot be deleted

When a snapshot is created with "." as a name, it cannot be deleted.

As a workaround, users must avoid creating the snapshot name with "." and "/".

Bugzilla:2306684

A period update commit is now added after migrating to a multi-site

Previously, a period commit after completing the migration to multi-site form was not being done. As a result, a warning about the master zone not having an endpoint would display despite the endpoint being configured.

With this fix, a period update commit is added after migrating to multi-site form and no warning is emitted.

Bugzilla:2309409

Performance statistics latency graph now displays the correct data

Previously, the latency graph in Object > Overview > Performance statistics was not displaying data because of the way the NaN values were handled in the code.

With this fix, the latency graph displays the correct data, as expected.

Bugzilla:2312818

”Delete realm” dialog is now displayed when deleting a realm

Previously, when clicking “Delete realm”, the delete realm dialog was not being displayed as it was broken.

With this fix, the delete realm dialog loads properly and users can delete the realm.

Bugzilla:2314892

Configuration values, such as rgw_realm, rgw_zonegroup, and rgw_zone are now set before deploying the Ceph Object Gateway daemons

Previously, configuration values like **rgw_realm**, **rgw_zonegroup**, and **rgw_zone** were being set after deploying the Ceph Object Gateway daemons. This would cause the Ceph Object Gateway daemons to deploy in the default realm, zone group, and zone configurations rather than the specified

configurations. This would require a restart to deploy them under the correct realm, zone group, and zone configuration.

With this fix, the configuration values are now set before deploying the Ceph Object Gateway daemons and they are deployed in the specified realm, zone group, and zone in the specification.

Bugzilla:2317492

4.3. CEPH FILE SYSTEM

Exceptions during **cephfs-top** are fixed

Previously, in some cases where there was insufficient space on the terminal, the **cephfs-top** command would not have enough space to run and would throw an exception.

With this fix, the exceptions that arise during the running of the **cephfs-top** command on large and small sized windows are fixed.

Bugzilla:2272580

The path restricted **cephx** credential no longer fails permission checks on a removed snapshot of a directory

Previously, path restriction checks were constructed on anonymous paths for unlinked directories accessed through a snapshot. As a result, a path restricted **cephx** credential would fail permission checks on a removed snapshot of a directory.

With this fix, the path constructed for the access check is constructed from the original path for the directory at the time of snapshot and the access checks are successfully passed.

Bugzilla:2293353

MDS no longer requests unnecessary authorization PINs

Previously, MDS would unnecessarily acquire remote authorization PINs for some workloads causing slower metadata operations.

With this fix, MDS no longer requests unnecessary authorization PINs, resulting in normal metadata performance.

Bugzilla:2299658

The erroneous patch from the kernel driver is processed appropriately and MDS no longer enters an infinite loop

Previously, an erroneous patch to the kernel driver caused the MDS to enter an infinite loop processing an operation due to which MDS would become largely unavailable.

With this fix, the erroneous message from the kernel driver is processed appropriately and the MDS no longer enters an infinite loop.

Bugzilla:2303693

Mirror daemon can now restart when blocklisted or fails

Previously, the time difference taken would lead to negative seconds and never reach the threshold interval. As a result, the mirror daemon would not restart when blocklisted or failed.

With this fix, the time difference calculation is corrected.

Bugzilla:2303865

JSON output of the ceph fs status command now correctly prints the rank field

Previously, due to a bug in the JSON output of the **ceph fs status** command, the rank field for standby-replay MDS daemons were incorrect. Instead of the format **{rank}-s**, where {rank} is the active MDS which the standby-replay is following, it displayed a random {rank}.

With this fix, the JSON output of **ceph fs status** command correctly prints the rank field for the standby-replay MDS in the format '{rank}-s'.

Bugzilla:2307231

sync_duration is now calculated in seconds

Previously, the sync duration was calculated in milliseconds. This would cause usability issues, as all other calculations were in seconds.

With this fix, **sync_duration** is now displayed in seconds.

Bugzilla:2308300

A lock is now implemented to guard the access to the shared data structure

Previously, access to a shared data structure without a lock caused the applications using the CephFS client library to throw an error.

With this fix, a lock, known as a mutex is implemented to guard the access to the shared data structure and applications using Ceph client library work as expected.

Bugzilla:2310155

The snap-schedule manager module correctly enforces the global mds_max_snaps_per_dir configuration option

Previously, the configuration value was not being correctly retrieved from the MDS. As a result, **snap-schedule manager** module would not enforce the **mds_max_snaps_per_dir** setting and would enforce a default limit of 100.

With this fix, the configuration item is correctly fetched from the MDS. The **snap-schedule manager** module correctly enforces the global **mds_max_snaps_per_dir** configuration option.

Bugzilla:2311030

CephFS FUSE clients can now correctly access the specified mds auth caps path

Previously, when parsing a path while validating **mds auth caps** FUSE clients were not able to access a specific path, even when the path was specified as **rw** for the **mds auth caps**.

With this fix, the parsing issue of the path while validating **mds auth caps** is fixed and the paths can be accessed as expected.

Bugzilla:2307931

4.4. CEPH OBJECT GATEWAY

SQL queries on a JSON statement no longer confuse key with array or object

SQL queries on a JSON statement no longer confuse `key` with `array` or `object`

Previously, in some cases, a result of an SQL statement on a JSON structure would confuse **key** with **array** or **object**. As a result, there was no **venue.id**, as defined, with **id** as the **key** value in the `venue object and it keeps traversing the whole JSON object.

With this fix, the SQL engine is fixed to avoid wrong results for mixing between **key** with **array** or **object** and returns the correct results, according to the query.

Bugzilla:2242089

Error code of local authentication engine is now returned correctly

Previously, incorrect error codes were returned when the local authentication engine was specified last in the authentication order and when the previous authentication engines were not applicable. As a result, incorrect error codes were returned.

With this fix, the code returns the error code of the local authentication engine in case the previous external authentication engines are not applicable for authenticating a request and correct error codes are returned.

Bugzilla:2268234

Lifecycle transition now works for the rules containing "Date"

Previously, due to a bug in the lifecycle transition code, rules containing "Date" did not get processed causing the objects meeting the criteria to not get transitioned to other storage classes.

With this fix, the lifecycle transition works for the rules containing "Date".

Bugzilla:2269490

Notifications are now sent on lifecycle transition

Previously, logic to dispatch on transition (as distinct from expiration) was missed. Due to this, notifications were not seen on transition.

With this fix, new logic is added and notifications are now sent on lifecycle transition.

Bugzilla:2281421

Batch object deleting is now allowed, with IAM policy permissions

Previously, during a batch delete process, also known as multi object delete, due to the incorrect evaluation of IAM policies returned **AccessDenied** output if no explicit or implicit deny were present. The **AccessDenied** occurred even if there were Allow privileges. As a result, batch deleting fails with the **AccessDenied** error.

With this fix, the policies are evaluated as expected and batch deleting succeeds, when IAM policies are enabled.

Bugzilla:2284154

Removing an S3 object now properly frees storage space

Previously, in some cases when removing the CopyObject and the size was larger than 4 MB, the object did not properly free all storage space that was used by that object. With this fix the source and destination handles are passed into various RGWRados call paths explicitly and the storage frees up, as expected.

Bugzilla:2294620

Quota and rate limit settings for assume-roles are properly enforced for S3 requests with temporary credentials

Previously, information of a user using an assume-role was not loaded successfully from the backend store when temporary credentials were being used to serve an S3 request. As a result, the user quota or rate limit settings were not applied with the temporary credentials.

With this fix, the information is loaded from the backend store, even when authenticating with temporary credentials and all settings are applied successfully.

Bugzilla:2298710

Pre-signed URLs are now accepted with Keystone EC2 authentication

Previously, a properly constructed pre-signed HTTP PUT URLs failed unexpectedly, with a **403**/**Access Denied** error. This happened because of a change in processing of HTTP OPTIONS requests containing CORS changed the implied AWSv4 request signature calculation for some pre-signed URLs when authentication was through Keystone EC2 (Swift S3 emulation).

With this fix, a new workflow for CORS HTTP OPTIONS is introduced for the Keystone EC2 case and pre-signed URLs no longer unexpectedly fail.

Bugzilla:2299642

Malformed JSON of radosgw-admin notification output is now corrected

Previously, when bucket notifications were configured with metadata and tag filters, the output of **radosgw-admin** notification for the get/list output was malformed JSON. As a result, any JSON parser, such as jquery, reading the output would fail.

With this fix, the JSON output for **radosgw-admin** is corrected.

Bugzilla:2303947

Clusters can now be configured with both QAT and non-QAT Ceph Object Gateway daemons

Previously, QAT could only be configured on new setups (Greenfield only). As a result, QAT Ceph Object Gateway daemons could not be configured in the same cluster as non-QAT (regular) Ceph Object Gateway daemons.

With this fix, both QAT and non-QAT Ceph Object Gateway daemons can be configured in the same cluster.

Bugzilla:2307218

Ceph Object Gateway now tolerates minio SDK with checksums and other hypothetical traffic

Previously, some versions of the minio client SDK would be missing an appended part number for multipart objects. This would result in unexpected errors for multipart uploads.

With this fix, checksums, both with and without part number suffixes, are accepted. The fix also permits the checksum type to be inferred from the init-multipart when a checksum is not asserted in part uploads.

Bugzilla:2310424

Lifecycle transition no longer fails for non-current object with an empty instance

Previously, when bucket versioning was enabled, old plain object entries would get converted to versioned by updating its instance as "null" in its raw head/old object. Due to this, the lifecycle transition would fail for a non-current object with an instance empty.

With this fix, the code is corrected to keep the instance empty while updating bucket index entries and the lifecycle transition works for all plain entries which are converted to versioned.

Bugzilla:2317891

The AST structure SQL statement no longer causes a crash

Previously, in some cases, due to an erroneous semantic combined with the Parquet flow, the AST creation that is produced by the SQL engine was wrong and a crash would occur.

With this fix, more safety checks are in place for the AST structure and the statement processing time is fixed and a crash is avoided.

Bugzilla:2290775

Bucket policy authorizations now work as expected

Previously, only a bucket owner was able to set, get, and delete the configurations for bucket notifications from a bucket. This was the case even if the bucket policy authorized another user for running these operations.

With this fix, authorization for configuring bucket notifications works as expected.

Bugzilla:2306898

Bucket policy evaluations now work as expected and allow cross tenant access for actions that are allowed by the policy

Previously, due to an incorrect value bucket tenant, during a bucket policy evaluation access was defined for S3 operations, even if they were explicitly allowed in the bucket policies. As a result, the bucket policy evaluation failed and S3 operations which were marked as allowed by the bucket policy were denied.

With this fix, the requested bucket tenant name is correctly passed when getting the bucket policy from the backend store. The tenant is then matched against the bucket tenant which was passed in as part of the S3 operation request, and S3 operations work as expected.

Bugzilla:2302940

SSL sessions can now reuse connections for uploading multiple objects

Previously, during consecutive object uploads using SSL, the cipher negotiations occurred for each object. As a result, there would be a low performance of objects per second transfer rate.

With this fix, the SSL session reuse mechanism is activated, allowing supporting clients to reuse existing SSL connections for uploading multiple objects. This avoids the performance penalty of renegotiating the SSL connection for each object.

Bugzilla:2236510

4.5. MULTI-SITE CEPH OBJECT GATEWAY

Objects with null version IDs delete on the second site

Previously, objects with null version IDs were not deleted on the second site. In a multi-site environment, deleting an object with a null version ID on one of the sites did not delete the object on the second site.

With this fix, objects with null version IDs delete on the second site.

Bugzilla:1967576

Bucket creations in a secondary zone no longer fail

Previously, when a secondary zone forwarded a **create_bucket** request with a location constraint, the bucket would set the **content_length** to a non-zero value. However, the **content_length** was not parsed on the primary zone when forwarded from the secondary zone. As a result, when running a **create_bucket** operation and the **content_length** is **0** with an existing payload hash, the bucket failed to replicate.

With this fix a request body is included when the **CreateBucket** operation is forwarded to the primary zone and the bucket is created as expected.

Bugzilla:2312161

CopyObject requests now replicate as expected

Previously, a **copy_object** would retain the source attributes by default. As a result, during a check for **RGW_ATTR_OBJ_REPLICATION_TRACE**, a **NOT_MODIFIED** error would emit if the destination zone was already present in the trace. This would cause a failure to replicate the copied object.

With this fix, the source object **RGW_ATTR_OBJ_REPLICATION_TRACE** attribute is removed during a **copy_object** and the **CopyObject** requests replicate as expected.

Bugzilla:2291166

4.6. RADOS

Newly added capacity no longer marked as allocated

Previously, newly added capacity was marked automatically as allocated. As a result, added disk capacity did not add available space.

With this fix, added capacity is marked as free and available, and after restarted OSD, newly added capacity is recognized as added space, as expected.

Bugzilla:2296247

BlueStore now works as expected with OSDs

Previously, the **ceph-bluestore-tool show-label** would not work on mounted OSDs and the **ceph-volume lvm zap** command was not able to erase the identity of an OSD. With this fix, the **show-label** attribute does not require exclusive access to the disk. In addition, the **ceph-volume** command now uses **ceph-bluestore-tool zap** to clear OSD devices.

Bugzilla:2311904

BlueStore no longer overwrites labels

Previously, BlueFS would write over the location reserved for a label. As a result, the OSD would not start as expected.

With this fix, the label location is marked as reserved and does not get overwritten. BlueStore now mounts and the OSD starts as expected.

Bugzilla:2314687

RocksDB files now only take as much space as needed

Previously, RocksDB files were generously preallocated but never truncated. This resulted in wasted disk space that was assigned to files that would never be used.

With this fix, proper truncation is implemented, moving unused allocations back to the free pool.

Bugzilla:2317027

Monitors no longer get stuck in elections during crash or shutdown tests

Previously, the **disallowed_leaders** attribute of the MonitorMap was conditionally filled only when entering **stretch_mode**. However, there were instances wherein Monitors that just got revived would not enter **stretch_mode** right away because they would be in a **probing** state. This led to a mismatch in the **disallowed_leaders** set between the monitors across the cluster. Due to this, Monitors would fail to elect a leader, and the election would be stuck, resulting in Ceph being unresponsive.

With this fix, Monitors do not have to be in **stretch_mode** to fill the **disallowed_leaders** attribute. Monitors no longer get stuck in elections during crash or shutdown tests.

Bugzilla:2249962

4.7. RADOS BLOCK DEVICES (RBD)

librbd no longer crashes when handling discard I/O requests

Previously, due to an implementation defect, **librbd** would crash when handling discard I/O requests that straddle multiple RADOS objects on an image with journaling feature enabled. The work around was to set the **setting rbd_skip_partial_discard** option to 'false' (the default is 'true').

With this fix, the implementation defect is fixed and **librbd** no longer crashes and the workaround is no longer necessary.

Bugzilla:2272517

rbd du command no longer crashes if a 0-sized block device image is encountered

Previously, due to an implementation defect, **rbd du** command would crash if it encountered a 0-sized RBD image.

With this fix, the implementation defect is fixed and the **rbd du** command no longer crashes if a 0-sized RBD image is encountered.

Bugzilla:2291447

rbd_diff_iterate2() API returns correct results for block device images with LUKS encryption loaded

Previously, due to an implementation defect, **rbd_diff_iterate2()** API returned incorrect results for RBD images with LUKS encryption loaded.

With this fix, **rbd_diff_iterate2()** API returns correct results for RBD images with LUKS encryption loaded.

Bugzilla:2292562

Encrypted image decryption is no longer skipped after importing or live migration

Previously, due to an implementation defect, when reading from an encrypted image that was live migrated or imported, decryption was skipped. As a result, an encrypted buffer (ciphertext), instead of the actual stored data (plaintext) was returned to the user.

With this fix, decryption is no longer skipped when reading from an encrypted image that is being live migrated or imported and the actual stored data (plaintext) is returned to the user, as expected.

Bugzilla:2303528

Encryption specification now always propagates to the migration source

Previously, due to an implementation defect, when opening an encrypted clone image that is being live migrated or imported, the encryption specification wouldn't be propagated to the migration source. As a result, the encrypted clone image that is being live migrated or imported would not open. The only workaround for the user was to pass the encryption specification twice by duplicating it.

With this fix, the encryption specification always propagates to the migration source.

Bugzilla:2308345

4.8. RBD MIRRORING

rbd-mirror daemon now properly disposes of outdated **PoolReplayer** instances

Previously, due to an implementation defect, **rbd-mirror** daemon did not properly dispose of outdated **PoolReplayer** instances in particular when refreshing the mirror peer configuration. Due to this there was unnecessary resource consumption and number of **PoolReplayer** instances competed against each other causing **rbd-mirror** daemon health to be reported as ERROR and replication would hang in some cases. To resume replication the administrator had to restart the **rbd-mirror** daemon.

With this fix, the implementation defect is corrected and rbd-mirror daemon now properly disposes of outdated **PoolReplayer** instances.

Bugzilla:2279528

4.9. NFS GANESHA

All memory consumed by the configuration reload process is now released

Previously, reload exports would not release all the memory consumed by the configuration reload process causing the memory footprint to increase.

With this fix, all memory consumed by the configuration reload process is released resulting in reduced memory footprint.

Bugzilla:2280364

The **reap_expired_client_list** no longer causes a deadlock

Previously, in some cases, the **reap_expired_client_list** would create a deadlock. This would occur due two threads waiting on each other to acquire a lock.

With this fix, the lock order is resolved and no deadlock occurs.

Bugzilla:2311296

File parsing and startup time is significantly reduced

Previously, due to poor management of parsed tokens, configuration file parsing was too slow.

With this fix, token lookup is replaced with AVL tree resulting in the reduction of parsing time and startup time.

Bugzilla:2314531

CHAPTER 5. TECHNOLOGY PREVIEWS

This section provides an overview of Technology Preview features introduced or updated in this release of Red Hat Ceph Storage.

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 to use them for production. These features provide early access to upcoming product features, enabling customers to test functionality and provide feedback during the development process.

For more information on Red Hat Technology Preview features support scope, see https://access.redhat.com/support/offerings/techpreview/.

Users can archive older data to an AWS bucket

With this release, users can enable data transition to a remote cloud service, such as Amazon Web Services (AWS), as part of the lifecycle configuration. See the *Transitioning data to Amazon S3 cloud service* for more details.

Expands the application of S3 select to Apache Parquet format

With this release, there are now two S3 select workflows, one for CSV and one for Parquet, that provide S3 select operations with CSV and Parquet objects. See the *S3 select operations* in the *Red Hat Ceph Storage Developer Guide* for more details.

Bucket granular multi-site sync policies is now supported

Red Hat now supports bucket granular multi-site sync policies. See the *Using multi-site sync policies* section in the *Red Hat Ceph Storage Object Gateway Guide* for more details.

Server-Side encryption is now supported

With this release, Red Hat provides the support to manage Server-Side encryption. This enables S3 users to protect data at rest with a unique key through Server-Side encryption with Amazon S3-managed encryption keys (SSE-S3).

Users can use the PutBucketEncryption S3 feature to enforce object encryption

Previously, to enforce object encryption in order to protect data, users were required to add a header to each request which was not possible in all cases.

With this release, Ceph Object Gateway is updated to support **PutBucketEncryption S3** action. Users can use the **PutBucketEncryption S3** feature with the Ceph Object Gateway without adding headers to each request. This is handled by the Ceph Object Gateway.

5.1. THE CEPHADM UTILITY

New Ceph Management gateway and the OAuth2 Proxy service for unified access and high availability

With this enhancement, the Ceph Dashboard introduces the Ceph Management gateway (**mgmt-gateway**) and the OAuth2 Proxy service (**oauth2-proxy**). With the Ceph Management gateway (**mgmt-gateway**) and the OAuth2 Proxy (**oauth2-proxy**) in place, **nginx** automatically directs the user through the **oauth2-proxy** to the configured Identity Provider (IdP), when single sign-on (SSO) is configured.

Bugzilla:2298666

5.2. CEPH DASHBOARD

New OAuth2 SSO

OAuth2 SSO uses the **oauth2-proxy** service to work with the Ceph Management gateway (**mgmt-gateway**), providing unified access and improved user experience.

Bugzilla:2312560

5.3. CEPH OBJECT GATEWAY

New bucket logging support for Ceph Object Gateway

Bucket logging provides a mechanism for logging all access to a bucket. The log data can be used to monitor bucket activity, detect unauthorized access, get insights into the bucket usage and use the logs as a journal for bucket changes. The log records are stored in objects in a separate bucket and can be analyzed later. Logging configuration is done at the bucket level and can be enabled or disabled at any time. The log bucket can accumulate logs from multiple buckets. The configured **prefix** may be used to distinguish between logs from different buckets.

For performance reasons, even though the log records are written to persistent storage, the log object appears in the log bucket only after a configurable amount of time or when reaching the maximum object size of 128 MB. Adding a log object to the log bucket is done in such a way that if no more records are written to the object, it might remain outside of the log bucket even after the configured time has passed.

There are two logging types: **standard** and **journal**. The default logging type is **standard**.

When set to **standard** the log records are written to the log bucket after the bucket operation is completed. As a result the logging operation can fail with no indication to the client.

When set to **journal** the records are written to the log bucket before the bucket operation is complete. As a result, the operation does not run if the logging action fails and an error is returned to the client.

You can complete the following bucket logging actions: enable, disable, and get.

Bugzilla:2308169

Support for user accounts through Identity and Access Management (IAM)

With this release, Ceph Object Gateway supports user accounts as an optional feature to enable the self-service management of Users, Groups, and Roles similar to those in AWS Identity and Access Management(IAM).

Restore objects transitioned to remote cloud endpoint back into Ceph Object gateway using the cloud-restore feature

With this release, the **cloud-restore** feature is implemented. This feature allows users to restore objects transitioned to remote cloud endpoint back into Ceph Object gateway, using either S3 restore-object API or by rehydrating using read-through options.

Bugzilla:2293539

CHAPTER 6. KNOWN ISSUES

This section documents known issues found in this release of Red Hat Ceph Storage.

6.1. THE CEPHADM UTILITY

Using the haproxy_qat_support setting in ingress specification causes the haproxy daemon to fail deployment

Currently, the **haproxy_qat_support** is present but not functional in the ingress specification. This was added to allow haproxy to offload encryption operations on machines with QAT hardware, intending to improve performance. The added function does not work as intended, due to an incomplete code update. If the **haproxy_qat_support** setting is used, then the haproxy daemon fails to deploy.

To avoid this issue, do not use this setting until it is fixed in a later release.

Bugzilla:2308344

PROMETHEUS_API_HOST may not get set when Cephadm initially deploys Prometheus

Currently, **PROMETHEUS_API_HOST** may not get set when Cephadm initially deploys Prometheus. This issue is seen most commonly when bootstrapping a cluster with **--skip-monitoring-stack**, then deploying Prometheus at a later time. Due to this, a few monitoring information may be unavailable.

As a workaround, use the command **ceph orch redeploy prometheus** to set the **PROMETHEUS_API_HOST** as it redeploys the Prometheus daemon(s). Additionally, the value can be set manually with the **ceph dashboard set-prometheus-api-host <value>** command.

Bugzilla:2315072

6.2. CEPH MANAGER PLUGINS

Sometimes ceph-mgr modules are temporarily unavailable and their commands fail

Occasionally, the balancer module takes a long time to load after a **ceph-mgr** restart. As a result, other **ceph-mgr** modules can become temporarily unavailable and their commands fail.

For example,

```
[ceph: root@host01 /]# ceph crash ls
Error ENOTSUP: Warning: due to ceph-mgr restart, some PG states may not be up to date
Module 'crash' is not enabled/loaded (required by command 'crash ls'): use `ceph mgr module enable crash` to enable it
```

As a workaround, in cases after a **ceph-mgr** restart, commands from certain **ceph-mgr** modules fail, check the status of the balancer, using the **ceph balancer status** command. This might occur, for example, during an upgrade. * If the balancer was previously active **"active": true** but now it is marked as **“active": false**, check its status until it is active again, then rerun the other ceph-mgr module commands. * In other cases, try to turn off the balancer **ceph-mgr** module. **ceph balancer off** After turning off the balancer, rerun the other **ceph-mgr** module commands.

Bugzilla:2314146

6.3. CEPH DASHBOARD

Ceph Object Gateway page does not load after a multi-site configuration

The Ceph Object Gateway page does not load because the dashboard cannot find the correct access key and secret key for the new realm during multi-site configuration.

As a workaround, use the **ceph dashboard set-rgw-credentials** command to manually update the keys.

Bugzilla:2231072

CephFS path is updated with the correct subvolume path when navigating through the subvolume tab

In the **Create NFS Export** form for CephFS, the CephFS path is updating the subvolume group path instead of the subvolume.

Currently, there is no workaround.

Bugzilla:2303247

Multi-site automation wizard mentions multi-cluster for both Red Hat and IBM Storage Ceph products

Within the multi-site automation wizard both Red Hat and IBM Storage Ceph products are mentioned in reference to multi-cluster. Only IBM Storage Ceph supports multi-cluster.

Bugzilla:2322398

Deprecated iSCSI feature is displayed in the Ceph Dashboard

Currently, although iSCSI is a deprecated feature, it is displayed in the Ceph Dashboard.

The UI for iSCSI feature is not usable.

Bugzilla:2331648

6.4. CEPH OBJECT GATEWAY

Objects uploaded as Swift SLO cannot be downloaded by anonymous users

Objects that are uploaded as Swift SLO cannot be downloaded by anonymous users.

Currently, there is no workaround for this issue.

Bugzilla:2272648

Not all apparently eligible reads can be performed locally

Currently, if a RADOS object has been recently created and in some cases, modified, it is not immediately possible to make a local read. Even when correctly configured and operating, not all apparently eligible reads can be performed locally. This is due to limitations of the RADOS protocol. In test environments, many objects are created and it is easy to create an unrepresentative sample of read-local I/Os.

Bugzilla:2309383

6.5. MULTI-SITE CEPH OBJECT GATEWAY

Buckets created by tenanted users do not replicate correctly

Currently, buckets that are created by tenanted users do not replicate correctly.

To avoid this issue, bucket owners should avoid using tenanted users to create buckets on secondary zone but instead only create them on master zone.

Bugzilla:2325018

When a secondary zone running Red Hat Ceph Storage 8.0 replicates user metadata from a pre-8.0 metadata master zone, access keys of those users are erroneously marked as "inactive".

Currently, when a secondary zone running Red Hat Ceph Storage 8.0 replicates user metadata from a pre-8.0 metadata master zone, the access keys of those users are erroneously marked as "inactive". Inactive keys cannot be used to authenticate requests, so those users are denied access to the secondary zone.

As a workaround, the current primary zone must be upgraded before other sites.

Bugzilla:2327402

6.6. RADOS

Placement groups are not scaled down in **upmap-read** and **read** balancer modes

Currently, **pg-upmap-primary** entries are not properly removed for placement groups (PGs) that are pending merge. For example, when the bulk flag is removed on a pool, or any case where the number of PGs in a pool decreases. As a result, the PG scale-down process gets stuck and the number of PGs in the affected pool do not decrease as expected.

As a workaround, remove the **pg_upmap_primary** entries in the OSD map of the affected pool. To view the entries, run the **ceph osd dump** command and then run **ceph osd rm-pg-upmap-primary PG_ID** for reach PG in the affected pool.

After using the workaround, the PG scale-down process resumes as expected.

Bugzilla:2302230

CHAPTER 7. ASYNCHRONOUS ERRATA UPDATES

This section describes the bug fixes, known issues, and enhancements of the z-stream releases.

7.1. RED HAT CEPH STORAGE 8.0Z1

Red Hat Ceph Storage release 8.0z1 is now available. The security updates and bug fixes that are included in the update are listed in the RHSA-2024:10956 and RHSA-2024:10957 advisories.

CHAPTER 8. SOURCES

The updated Red Hat Ceph Storage source code packages are available at the following location:

- For Red Hat Enterprise Linux 9:
  https://ftp.redhat.com/redhat/linux/enterprise/9Base/en/RHCEPH/SRPMS/