



Red Hat Ceph Storage 2.0 Release Notes

Release notes for Red Hat Ceph Storage 2.0

Red Hat Ceph Storage Documentation
Team

Red Hat Ceph Storage 2.0 Release Notes

Release notes for Red Hat Ceph Storage 2.0

Legal Notice

Copyright © 2017 Red Hat, Inc.

The text of and illustrations in this document are licensed by Red Hat under a Creative Commons Attribution–Share Alike 3.0 Unported license ("CC-BY-SA"). An explanation of CC-BY-SA is available at

<http://creativecommons.org/licenses/by-sa/3.0/>

. In accordance with CC-BY-SA, if you distribute this document or an adaptation of it, you must provide the URL for the original version.

Red Hat, as the licensor of this document, waives the right to enforce, and agrees not to assert, Section 4d of CC-BY-SA to the fullest extent permitted by applicable law.

Red Hat, Red Hat Enterprise Linux, the Shadowman logo, JBoss, OpenShift, Fedora, the Infinity logo, and RHCE are trademarks of Red Hat, Inc., registered in the United States and other countries.

Linux ® is the registered trademark of Linus Torvalds in the United States and other countries.

Java ® is a registered trademark of Oracle and/or its affiliates.

XFS ® is a trademark of Silicon Graphics International Corp. or its subsidiaries in the United States and/or other countries.

MySQL ® is a registered trademark of MySQL AB in the United States, the European Union and other countries.

Node.js ® is an official trademark of Joyent. Red Hat Software Collections is not formally related to or endorsed by the official Joyent Node.js open source or commercial project.

The OpenStack ® Word Mark and OpenStack logo are either registered trademarks/service marks or trademarks/service marks of the OpenStack Foundation, in the United States and other countries and are used with the OpenStack Foundation's permission. We are not affiliated with, endorsed or sponsored by the OpenStack Foundation, or the OpenStack community.

All other trademarks are the property of their respective owners.

Abstract

The Release Notes document describes the major features and enhancements implemented in Red Hat Ceph Storage in a particular release. The document also includes known issues and bug fixes.

Table of Contents

CHAPTER 1. INTRODUCTION	3
CHAPTER 2. ACKNOWLEDGMENTS	4
CHAPTER 3. MAJOR UPDATES	5
CHAPTER 4. TECHNOLOGY PREVIEWS	10
CHAPTER 5. DEPRECATED FUNCTIONALITY	11
CHAPTER 6. KNOWN ISSUES	12
CHAPTER 7. NOTABLE BUG FIXES	21
CHAPTER 8. SOURCES	22

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 the Red Hat Storage Console deployment and management utility, and support services.

CHAPTER 2. ACKNOWLEDGMENTS

Red Hat Ceph Storage version 2.0 contains many contributions from the Red Hat Ceph Storage team. Additionally, 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:

- ✧ Mirantis
- ✧ XSky
- ✧ SUSE
- ✧ DreamHost
- ✧ Intel
- ✧ ZTE
- ✧ Gentoo

CHAPTER 3. MAJOR UPDATES

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

Installation by using Ansible is now supported

With this release, the Ansible automation application can be used to install Red Hat Ceph Storage. The **ceph-ansible** utility contains a set of Ansible playbooks that allow users to add or update monitor, OSD, Ceph Object Gateway, and Ceph Metadata Server nodes.

For details, see the [Installation Guide for Red Hat Enterprise Linux](#) and [Installation Guide for Ubuntu](#).

Red Hat Storage Console

Red Hat Storage Console is a new Red Hat product that provides a graphical management platform for Red Hat Ceph Storage 2. Red Hat Storage Console allows users to install, monitor, and manage a Red Hat Ceph Storage cluster.

For details, see the [Red Hat Ceph Storage Console documentation set](#).

SELinux is now in enforcing mode by default

The SELinux policy for Red Hat Ceph Storage was added in Red Hat Ceph Storage 1.3.2 but was not enabled by default. With this release, SELinux is enabled by default for all nodes, except for the Red Hat Storage Console node, where it runs in permissive mode.

To learn more about SELinux, see the [SELinux User's and Administrator's Guide](#) for Red Hat Enterprise Linux 7.

Changed behavior of `calamari-ctl initialize`

The **calamari-ctl initialize** command no longer prompts for creating the administrator's account. With this release, **calamari-ctl initialize** requires to specify the administrator's account by using the following command-line arguments:

```
calamari-ctl initialize --admin-username <username> --admin-password  
<password> --admin-email <email>
```

The Ceph API contains more details about OSD nodes

The Ceph API now includes locations of the data that is stored on an OSD and the path to the OSD journal.

The API is accessible from `/api/v2/cluster/<fsid>/osd?format=json`. Also, the Red Hat Storage Console uses this information to infer what devices the OSD daemons use.

Support for per tenant namespaces has been added

The Ceph Object Gateway now supports multiple tenants with associated namespaces. Previously, the Ceph Object Gateway had only a single global namespace for buckets and containers. As a consequence, names of buckets and containers had to be unique, even across distinct accounts.

With this new feature, the names have to be unique only within a tenant.

For details, see the [Ceph Object Gateway Guide for Red Hat Enterprise Linux](#) and [Ceph Object Gateway Guide for Ubuntu](#).

The AWS4 authentication is now supported

Red Hat Ceph Storage now supports the new Amazon Web Services (AWS) Signature Version 4 authentication.

crushtool now checks for overlapping CRUSH rules

The **crushtool** now checks if the CRUSH rules overlap. To check if the CRUSH rules overlap, use the following command. If **crushtool** finds any overlapping rules, it outputs them:

```
$ crushtool --check -i <crush_file>
overlapped rules in ruleset 0: rule-r0, rule-r1, rule-r2
```

Image metadata is now supported

With this release, it is possible to tag images with custom key-value pairs. Also, metadata can be used to override the RADOS Block Device (RBD) image configuration settings for particular images.

For details, see the [Working with Image Metadata](#) section in the [Block Device](#) guide for Red Hat Ceph Storage 2.

Calamari now supports passing the rbd commands

With this release, the Calamari `/cli` endpoint supports passing the **rbd** commands.

Support for randomized scrub scheduling has been added

The new Ceph **osd_scrub_interval_randomize_ratio** configuration option randomizes scrub scheduling by adding a random delay to the value specified by the **osd_scrub_min_interval** option. As a result, scrubbing of newly created pool or placement groups does not happen at the same time, and I/O impact is reduced.

Scrub operations are now in the client queue

The scrub and trimming operations have been moved to the client operations queue to prioritize client I/O more accurately.

The deep flatten feature has been added

Without the **deep flatten** feature, it is not possible to dissociate a clone from its parent if the clone has any snapshots. As a consequence, the parent image or snapshot cannot be deleted. With **deep flatten** enabled, it is possible to dissociate a clone from its parent and to delete no longer needed snapshots.

The **deep flatten** feature must be enabled when creating an image. It is not possible to enable it on already existing images.

Support for Swift Static Large Object (SLO) has been added

Ceph now supports the Swift Static Large Object (SLO) feature. This feature allows uploading large objects. To do so, SLO splits a large object into smaller objects, uploads the smaller objects, and then treats the result as a single large object.

For details, see the [Static Large Object \(SLO\) support](#) chapter in the [Configuration Reference](#) guide for Red Hat OpenStack Platform 8.

Support for renaming snapshots

With this release, support for renaming snapshots has been added to the Red Hat Ceph Storage 2. To rename a snapshot:

```

rbd snap rename <pool-name>/<image-name>@<original-snapshot-name>
<pool-name>/<image-name>@<new-snapshot-name>

```

For example:

```

$ rbd snap rename data/dataset@snap1 data/dataset@snap2

```

For details, see the [Renaming Snapshots](#) section in the [Block Device](#) guide for Red Hat Ceph Storage 2.

Keystone 3 authentication is now supported

Red Hat Ceph Storage now supports OpenStack Keystone 3 authentication. As a result, users can use Keystone 3 to authenticate to the Ceph Object Gateway.

For details, see the [Using Keystone to Authenticate Ceph Object Gateway Users](#) guide.

Block Device mirroring is now supported

The RADOS Block Device (RBD) mirroring feature has been added to the Red Hat Ceph Storage 2. RBD mirroring is a process of replicating Ceph Block Device images between two peer Ceph Storage Clusters. The mirroring is asynchronous, crash-consistent, and serves primarily for disaster recovery.

To learn more about RBD mirroring, see the [Block Device Mirroring](#) chapter in the [Block Device](#) guide for Red Hat Ceph Storage 2.

systemd now restarts failed Ceph services

When a Ceph service, such as **ceph-mon** or **ceph-osd**, fails to start, the **systemd** daemon now attempts to restart the service. Prior to this update, Ceph services remained in the failed state.

Image features can be now enabled or disabled on already existing images

Image features, such as **fast diff**, **exclusive-lock**, **object map**, or **journaling**, can now be enabled or disabled on already existing images. The **deep-flatten** feature can be disabled but not enabled on existing images.

For details, see the [Enabling and Disabling Image Features](#) section in the [Block Device](#) guide for Red Hat Ceph Storage 2.

The scrub feature has been enhanced

The OSD daemon now persists scrub results. In addition, a new **librados** API is provided to query the scrub results in detail.

Changes in the rbd command

This release introduces the following changes in the **rbd** command:

- ✦ The **rbd help** command can now output help for a particular subcommand instead of for all options and commands. To do so, use:

```
rbd help <subcommand>
```



Note

The **-h** option still displays help for all available commands.

- ✦ Size can now be specified in units, for example in megabytes, gigabytes, or terabytes.
- ✦ Object size can now be specified directly rather than as a power of 2.

New multi-site configurations for the Ceph Object Gateway

The Red Hat Ceph Storage now supports an active-active zone configuration of Ceph Object Gateways. Previously, only active-passive configuration was supported, meaning the users could write only to one (master) zone. With this update, users can write also to non-master zones.

In addition, this update introduces the following changes:

- ✦ A region is now called **zone-group**.
- ✦ Support for configuring realms has been added. A realm is a single zone group or multiple zone groups with a globally unique namespace.
- ✦ A configuration period has been added. The period is a representation of the multi-site configuration at a point in time. The current period contains an epoch that advances as changes are made to the configuration. A new period is generated when the location of the multi-site master zone changes.
- ✦ The **ceph-radosgw** daemon handles the synchronization, eliminating the need for a separate synchronization agent.

For details see the [Multi-site](#) chapter in the [Object Gateway Guide for Red Hat Enterprise Linux](#) and the [Multi-site](#) chapter in the [Object Gateway Guide for Ubuntu](#).

The fast-diff feature has been added

The **fast-diff** feature adds the ability to quickly determine if an object has changed since the last snapshot was made. As a result, generating the difference between two images as well as using the **rbd disk-usage** command is faster.

For details, see the [Enabling and Disabling Image Features](#) section in the [Block Device](#) guide for Red Hat Ceph Storage 2.

Support for Swift bulk-delete has been added

Red Hat Ceph Storage now supports the Swift **bulk-delete** feature. With this new feature, users can delete multiple objects from an account with a single request.

For details see the [Bulk delete](#) section in the [Configuration Reference](#) guide for Red Hat OpenStack Platform 8.

Swift Object Expiration is now supported

Ceph now supports expiration of Swift objects.

For details see the [Expiring Object Support](#) section of the OpenStack documentation.

Ceph daemons now run as the ceph user and group

With this release, Ceph daemons, such as **ceph-osd** or **ceph-mon**, no longer run as **root** but run as the **ceph** user that belongs to the **ceph** group. This change improves security of the Ceph cluster.

Active Directory/LDAP authentication is now supported

Red Hat Ceph Storage now supports LDAP authentication. As a result, users can use LDAP accounts to access buckets in the Ceph Object Gateway. In addition, they can use LDAP authentication to authenticate against Microsoft Active Directory.

For details see the [Ceph Object Gateway with LDAP/AD](#) guide.

CHAPTER 4. 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/>.

OSD BlueStore

BlueStore is a new back end for the OSD daemons that allows for storing objects directly on the Ceph block devices. Because BlueStore does not need any file system interface, it improves performance of Ceph storage clusters.

To learn more about the BlueStore OSD back end, see the [OSD BlueStore \(Technology Preview\)](#) chapter in the [Administration Guide](#).

The NFS interface for the Ceph Object Gateway

The Ceph Object Gateway now includes a new NFS interface that presents buckets and objects as directories and files. The main purpose of this new functionality is to facilitate migration from POSIX applications and workloads to the Ceph Object Gateway.

The Ceph File System (CephFS) and Metadata Server (MDS)

The Ceph File System (CephFS) and Ceph Metadata Server (MDS) have been added as Technology Preview features.

The Ceph File System is a POSIX-compliant file system that uses the Ceph Storage Cluster to store its data. CephFS can be mounted with a kernel driver or as a file system in user space (FUSE).

The Ceph Metadata Server manages metadata related to files stored on the Ceph File System. The MDS daemon (**ceph-mds**) also coordinates access to the shared cluster.

In addition, the CephFS repair utilities are now supported.

For details, see the [Ceph File System Guide](#).

CHAPTER 5. DEPRECATED FUNCTIONALITY

This section provides an overview of functionality that has been deprecated in all minor releases up to this release of Red Hat Ceph Storage.

The `ceph-deploy` and `ice_setup` utilities are now deprecated

With this release, the `ceph-deploy` and `ice_setup` utilities are deprecated in favor of the Red Hat Storage Console. All the administration tasks related to `ceph-deploy` or `ice_setup`, such as installation and monitoring, are now handled by Red Hat Storage Console.

For detailed information about Red Hat Storage Console, see the [Red Hat Ceph Storage Console documentation set](#).

In addition, the Ansible automation application can be used to install the Ceph Storage Cluster. For details, see the [Installation Guide for Red Hat Enterprise Linux](#) and [Installation Guide for Ubuntu](#).

Changes in the Calamari Architecture

The Calamari architecture has been simplified and made more lightweight:

- ✦ The Calamari UI has been deprecated in favor of the Red Hat Storage Console.
- ✦ The Calamari API is now co-located with monitor nodes by default and it is present as a default component of any cluster.

Cache tiering is now deprecated

The RADOS-level `cache tiering` feature has been deprecated. The feature does not provide any performance improvements for most Ceph workloads and introduces stability issues to the cluster when activated.

Alternative cache options will be provided by using the `dm-cache` feature in future versions of Red Hat Ceph Storage.

CHAPTER 6. KNOWN ISSUES

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

The Calamari API sometimes incorrectly prints standard output as an error

When using the Calamari `/cli` endpoint, the Calamari API response swaps the standard output (`stdout`) and standard error (`stderr`). As a consequence, the `err` field contains information from the `stdout` field. This behavior does not have any impact on the Calamari API commands. ([BZ#1287904](#), [BZ#1287905](#))

Renaming snapshots returns errors on overloaded clusters

When a Ceph storage cluster is overloaded and an image is currently in-use for I/O, the cluster might not be able to service rename requests in a timely fashion. Consequently, the RADOS Block Device (RBD) CLI keeps sending the rename request every 5 seconds because it receives a message that the request timed out. This behavior causes error log messages to be returned within the logs of the process performing I/O on the image. To avoid this problem, do not rename snapshots when the cluster is overloaded. ([BZ#1340772](#))

Multi-site configuration of the Ceph Object Gateway sometimes fails when options are changed at runtime

When the `rgw md log max shards` and `rgw data log num shards` options are changed at runtime in multi-site configuration of the Ceph Object Gateway, the `radosgw` process terminates unexpectedly with a segmentation fault.

To avoid this issue, do not change the aforementioned options at runtime, but set them during the initial configuration of the Ceph Object Gateway. ([BZ#1330952](#))

Calamari hangs when interactive commands are issued

The Calamari REST API incorrectly includes the following interactive commands:

- ✦ `rbid import`
- ✦ `rbid import-diff`
- ✦ `rbid journal import`
- ✦ `rbid merge-diff`
- ✦ `rbid watch`

When these commands are executed from Calamari, Calamari becomes unresponsive because it waits for an action from the user.

To work around this issue, do not execute the aforementioned commands from the Calamari REST API and use the command line instead. ([BZ#1354459](#))

Bucket creation fails after upgrading Red Hat Ceph Storage 1.3 to 2.0

After upgrading an Object Gateway node from Red Hat Ceph Storage 1.3 to 2.0, an attempt to create a bucket fails. To work around this issue, restart the **radosgw** service and create the bucket again. ([BZ#1352888](#))

Unable to write data on a promoted image

In RBD mirroring configuration, an image can be demoted to non-primary on the local cluster and promoted to primary on the remote cluster. If this happens and the **rbd-mirror** daemon is not restarted on the remote cluster, it is not possible to write data on the promoted image because **rbd-daemon** considers the demoted image on the local cluster to be the primary one.

To avoid this issue, restart the **rbd-mirror** daemon to gain the read/write access to the promoted image. ([BZ#1365648](#))

Mirroring image metadata is not supported

Image metadata are not currently replicated to a peer cluster. ([BZ#1344212](#))

The radosgw-admin command sometimes returns unnecessary error messages

The log level of certain common error messages that are returned by the **radosgw-admin** command is set to **0**. Consequently, these messages appear even though **radosgw-admin** command succeeds. These messages can be safely ignored. ([BZ#1364353](#))

Ansible fails to install OSDs if they point to directories

Ansible does not support installation of OSDs that point to directories and not to partitions. As a consequence, an attempt to install such OSDs fails. ([BZ#1361228](#))

The default value of journal_size is 0

By default, the **journal_size** option in the **ceph-ansible** utility is set to **0**. This option is mandatory and Red Hat strongly recommends to reserve enough disk space for the journal, at least 5 GB depending on the size of your cluster. Using a journal smaller than 5 GB leads to critical performance issues and production stability problems.

For more details, see the [Journal Settings](#) section in the [Configuration Guide](#) for Red Hat Ceph Storage 2. ([BZ#1359889](#))

Certain features available in ceph-ansible are not supported

The **ceph-ansible** utility contains the following features that are not supported with Red Hat Ceph Storage 2.0:

- ✦ the **purge-cluster** feature
- ✦ the **restapis** role ([BZ#1366394](#))

Sync point snapshots are not properly removed after a failover or failback

In a RADOS Block Device (RBD) mirroring configuration, it is possible that sync point snapshots created by the **rbd-mirror** daemon incorrectly remain on the clusters after a failover or failback. These snapshots can be safely removed if an image synchronization is not in progress.

(BZ#1350003)

Buckets sometimes have incorrect time stamps

Buckets created by the Simple Storage Service (S3) API on the Ceph Object Gateway before mounting the Ganesh NFS interface have incorrect time stamps. ([BZ#1359404](#))

The UID and GID of NFS directories is 4294967294 when ID mapping is not properly configured

When ID mapping between a client node and the node with the NFS interface for the Ceph Object Gateway is not properly configured, the UID and GID of newly created directories are set to **4294967294**, which is the default. This behavior is equivalent to the **all_squash** option of the **knfsd** NFS server. ([BZ#1359407](#))

Ansible fails to add a monitor to an upgraded cluster

An attempt to add a monitor to a cluster by using the Ansible automation application after upgrading the cluster from Red Hat Ceph Storage 1.3 to 2.0 fails on the following task:

```
TASK: [ceph-mon | collect admin and bootstrap keys]
```

This happens because the original monitor keyring was created with the **mds "allow"** capability while the newly added monitor requires a keyring with the **mds "allow *"** capability.

To work around this issue, after installing the **ceph-mon** package, manually copy the administration keyring from an already existing monitor node to the new monitor node:

```
scp /etc/ceph/<cluster_name>.client.admin.keyring  
<target_host_name>:/etc/ceph
```

For example:

```
# scp /etc/ceph/ceph.client.admin.keyring node4:/etc/ceph
```

Then use Ansible to add the monitor as described in the [Adding a Monitor with Ansible](#) section of the [Administration Guide](#) for Red Hat Ceph Storage 2. ([BZ#1357292](#))

Certain maintenance image operations are incorrectly allowed on non-primary images

With RADOS Block Device (RBD) mirroring enabled, non-primary images are expected to be read-only. Under certain conditions, the **rbd** command does not properly restrict **rbd** maintenance operations against non-primary images. The affected operations include:

- ✦ updating snapshots
- ✦ resizing images
- ✦ renaming and creating clones using the non-primary image as the parent
- ✦ disabling image features

These operations should be disallowed on non-primary images. To avoid this issue, make sure to perform the maintenance operations only against the primary image instance. ([BZ#1348928](#), [BZ#1352878](#), [BZ#1353877](#), [BZ#1349332](#))

The NFS interface for the Ceph Object Gateway does not show bucket size or number of blocks

The NFS interface of the Ceph Object Gateway lists buckets as directories. However, the interface always shows that the directory size and the number of blocks is **0**, even if some data is written to the buckets. ([BZ#1359408](#))

Certain image features are not supported with the RBD kernel module

The following image features are not supported with the current version of the RADOS Block Device (RBD) kernel module (**krbd**) that is shipped with the Red Hat Enterprise Linux 7.2:

- ✦ **object-map**
- ✦ **deep-flatten**
- ✦ **journaling**
- ✦ **exclusive-lock**
- ✦ **fast-diff**

However, by default the **ceph-installer** utility creates RBDs with the aforementioned features enabled. As a consequence, an attempt to map the kernel RBDs by running the **rbd map** command fails.

To work around this issue, disable the unsupported features by setting the **rbd_default_features** option in the Ceph configuration file for kernel RBDs or dynamically disable them by running the following command:

```
rbd feature disable <image> <feature>
```

This issue is a limitation only in kernel RBDs, and the features work as expected with user-space RBDs. ([BZ#1340080](#))

The **rbd bench write** command fails when **--io-size** is greater than or equal to the image size

The **rbd bench-write --io-size <size> <image>** command fails with a segmentation fault when the size specified by the **--io-size** option is greater than or equal to the image size.

To avoid this issue, make sure that the value of **--io-size** is smaller than the image size. ([BZ#1362014](#))

FUSE clients cannot be mounted permanently on Red Hat Enterprise Linux 7.2

The **util-linux** package shipped with Red Hat Enterprise Linux 7.2 does not support mounting CephFS Filesystem in Userspace (FUSE) clients in the **/etc/fstab** file. Red Hat Enterprise Linux 7.3 will include a new version of **util-linux** that will support mounting CephFS FUSE clients permanently. ([BZ#1360849](#))

Image synchronization starts from the beginning after restarting `rbd-mirror`

When the `rbd-mirror` daemon is restarted during image synchronization, the synchronization starts from the beginning.

To avoid this issue, do not restart `rbd-mirror` during image synchronization. ([BZ#1348940](#))

Results from deep scrubbing are overwritten by shallow scrubbing

When performing shallow scrubbing after deep scrubbing, results from deep scrubbing are overwritten by results from shallow scrubbing. As a consequence, the deep scrubbing results are lost. There is no workaround for this issue yet. ([BZ#1330023](#))

Calamari sometimes does not respond when sending a PATCH Request

The Calamari API does not respond when making PATCH requests to `/api/v2/cluster/FSID/osd/OSD_ID` if the requests does not change any fields on the OSD from their present values. ([BZ#1338688](#))

RBD mirroring is not disabled if peer clusters are not removed first

Disabling of RADOS Block Device (RBD) mirroring in pool mode consists of two steps:

1. Removing peer clusters by using the `rbd mirror pool remove` command.
2. Disabling pool mirroring by using the `rbd mirror pool disable` command.

If pool mirroring is disabled before removing the peer clusters, all the existing images on the mirrored pool stop being replicated but the pool remains mirrored. As a consequence, newly created images on the pool will be still replicated.

In addition, an error message similar to the following one is returned:

```
2016-08-02 10:30:31.823726 7f96fd7e8d80 -1 librbd: Failed to set mirror mode: (16) Device or resource busy
```

To avoid this issue, first remove peer clusters and then disable pool mirroring. ([BZ#1362647](#))

An LDAP user can access buckets created by a local RGW user with the same name

The RADOS Object Gateway (RGW) does not differentiate between a local RGW user and an LDAP user with the same name. As a consequence, the LDAP user can access the buckets created by the local RGW user.

To work around this issue, use different names for RGW and LDAP users. ([BZ#1361754](#))

POSIX ACL support is disabled by default in CephFS FUSE clients

Support for Access Control Lists (ACL) is disabled by default for Ceph File Systems (CephFS) mounted as FUSE clients. To use the ACL feature with FUSE clients, enable it manually. For details, see the [Limitations](#) section in the [Ceph File System Guide](#) for Red Hat Ceph Storage 2.

In addition, ACL in CephFS kernel clients is supported on Red Hat Enterprise Linux with kernel version `kernel-3.10.0-327.18.2.el7` or later. ([BZ#1342751](#))

Failed instances of the ceph-disk service

After installation of a new Ceph storage cluster, failed instances of the **ceph-disk** service appear. This is because the service is started twice - once to activate the data partition and once to activate the journal partition. After the disk activation, one of these instances fails because of limited resources. This is an expected behavior and the failed instance does not have any impact on the disk activation. ([BZ#1326740](#))

The Calamari API returns the "server error (500)" error when changing the take step

When changing a CRUSH rule, modifying the **take** step type to any other value than **take** causes the Calamari API to return the "server error (500)" error.

To avoid this issue, do not change the **take** step to any other value. ([BZ#1329216](#))

Client installation by using Ansible is not supported

The current version of the **ceph-ansible** utility does not support installation of the Ceph command-line interface (CLI). To use the Ceph CLI, install it manually. For details, see the [Client Installation](#) chapter in in the [Installation Guide for Red Hat Enterprise Linux](#) or the [Client Installation](#) chapter in in the [Installation Guide for Ubuntu](#). ([BZ#1335308](#))

S3 versioning cannot be set on buckets when accessing a non-master zone

Amazon Simple Storage Service (S3) object versioning cannot be set on buckets when accessing any zone other than the metadata master zone. ([BZ#1350522](#))

An error message is returned when downloading an S3 multipart file by using Swift

When uploading a multiple part file by using the Simple Storage Service (S3) and then downloading it by using the Swift service, the following error message is returned:

```
Error downloading object: md5sum != etag
```

Despite the message, the upload and download operations succeed, and the message can be safely ignored. ([BZ#1361044](#))

The rados list-inconsistent-obj command does not highlight inconsistent shards when it could have

The output of the **rados list-inconsistent-obj** command does not explicitly show which shard is inconsistent when it could have. ([BZ#1363949](#))

Restart of the radosgw service on clients is needed after rebooting the cluster

After rebooting the Ceph cluster, the **radosgw** service must be restarted on the Ceph Object Gateway clients to restore the connection with the cluster. ([BZ#1363689](#))

Ansible does not support adding encrypted OSDs

The current version of the **ceph-ansible** utility does not support adding encrypted OSD nodes. As a consequence, an attempt to perform asynchronous updates between releases by using the **rolling_update** playbook fails to upgrade encrypted OSD nodes. In addition, Ansible returns the following error message during the disk activation task:

```
mount: unknown filesystem type 'crypto_LUKS'
```

To work around this issue, open the **rolling_update.yml** file located in the Ansible working directory and find all instances of the **roles:** list. Then remove or comment out all roles (**ceph-mon**, **ceph-rgw**, **ceph-osd** or **ceph-mds**) except the **ceph-common** role from the lists, for example:

```
roles:
  - ceph-common
  #- ceph-mon
```



Note

Make sure to edit all instances of the **roles:** list in the **rolling_update.yml** file.

Then run the **rolling_update** playbook to update the nodes. ([BZ#1366808](#))

Dynamic feature updates are not replicated

When a feature is disabled or enabled on an already existing image and the image is mirrored to a peer cluster, the feature is not disabled or enabled on the replicated image. ([BZ#1344262](#))

Users created by using the Calamari API do not have permissions to run the API commands

When a user is created by using the Calamari REST API (**api/v2/user**), the user does not have permissions to run most of the Calamari API commands. Consequently, an attempt to run the commands fails with the following error message:

```
"You do not have permission to perform this action"
```

To work around this issue, use the **calamari-ctl add_user** command from the command line when creating new users. ([BZ#1356872](#))

Ansible does not support removing monitor or OSD nodes

The current version of the **ceph-ansible** utility does not support removing monitor or OSD nodes. To remove monitor or OSD nodes from a cluster, use the manual procedure. For more information, see the [Administration Guide](#) for Red Hat Ceph Storage 2. ([BZ#1335569](#))

Error messages are returned after disabling journaling

When the **journaling** feature is disabled on an image that was previously mirrored by the **rbd-mirror** daemon, an error message similar to the following is returned on the primary cluster:

```
2016-06-15 22:10:40.462481 7fed3d10b700 -1 rbd::mirror::ImageReplayer:
0x7fec8003b80 [1/29d86f79-7bba-4316-9ab9-c8a3f600e0f2] operator():
start failed: (2) No such file or directory
```

These harmless error messages indicates that the image journal was deleted because **journaling** was disabled and can be safely ignored. ([BZ#1346946](#))

Calamari sometimes incorrectly outputs "null" as a value

When the Calamari REST-based API is used to get details of a CRUSH rule in the Ceph cluster, the output contains "null" as a value for certain fields in the **steps** section of the CRUSH rule. The fields containing null values can be safely ignored for the respective steps in the CRUSH rule. However, do not use "null" as a value for any field when doing a PATCH operation. Using null values in such a case causes the operation to fail. ([BZ#1342504](#))

The Ceph Object Gateway is restarted after switching the zone from master to non-master

When a non-master zone is promoted to the master zone, all I/O requests become unresponsive until the **radosgw** process is restarted on both zones. Consequently, the I/O requests time out. ([BZ#1359712](#))

Images are synchronizing from the beginning after their demotion and promotion

With RADOS Block Device (RBD) mirroring enabled, an image can be demoted to non-primary on one cluster and promoted to primary on a peer cluster. When this happens, the **rbd-mirror** daemon starts to synchronize the newly demoted image with the newly promoted image even though the image has been already successfully synchronized before the demotion and promotion of the image. This behavior is not optimal and will be fixed in a future release. ([BZ#1349955](#))

Ansible does not properly handle unresponsive tasks

Certain tasks, for example adding monitors with the same host name, cause the **ceph-ansible** utility to become unresponsive. Currently, there is no timeout set after which the unresponsive tasks is marked as failed. ([BZ#1313935](#))

The Calamari REST-based API fails to edit user details

An attempt to use the Calamari REST-based API to edit user details fails with an error. To change user details, use the **calamari-ctl** command-line utility. ([BZ#1338649](#))

Object sync requests are sometimes skipped

In multi-site configurations of the Ceph Object Gateway, a non-master zone can be promoted to the master zone. In most cases, the master zone's gateway or gateways are still running when this happens. However, if the gateways are down, it can take up to 30 seconds after their restart until the gateways notice that another zone was promoted. During this time, the gateways can miss changes to buckets that occur on other zones. Consequently, object sync requests are skipped.

To work around this issue, pull the new master's period to the old master zone before restarting the old master zone:

```
$ radosgw-admin period pull --remote=<new-master-zone-id>
```

For details on pulling the period, see the [Ceph Object Gateway Guide for Red Hat Enterprise Linux](#) or the [Ceph Object Gateway Guide for Ubuntu](#). (BZ#1362639)

CHAPTER 7. NOTABLE BUG FIXES

This section describes bugs fixed in this release of Red Hat Ceph Storage that have significant impact on users.

OSD now deletes old OSD maps as expected

When new OSD maps are received, the OSD daemon marks the unused OSD maps as **stale** and deletes them to keep up with the changes. Previously, an attempt to delete stale OSD maps could fail for various reasons. As a consequence, certain OSD nodes were sometimes marked as **down** if it took too long to clean their OSD map caches when booting. With this update, the OSD daemon deletes old OSD maps as expected, thus fixing this bug. (BZ#1339061)

OSDs no longer crash when running `mark_lost_revert` or `mark_lost_delete`

Previously, issuing the `mark_lost_revert` and `mark_lost_delete` commands during active client write operations could cause the OSD daemons to terminate unexpectedly. The underlying code has been modified, and OSDs no longer crash in the described scenario. (BZ#1305152)

CHAPTER 8. SOURCES

The updated Red Hat Ceph Storage packages are available at the following locations:

- ✦ for Red Hat Enterprise Linux:
<ftp://ftp.redhat.com/redhat/linux/enterprise/7Server/en/RHCEPH/SRPMS/>
- ✦ for Ubuntu: <https://rhcs.download.redhat.com/ubuntu/>