



Red Hat Ceph Storage 2.1 Release Notes

Release notes for Red Hat Ceph Storage 2.1

Red Hat Ceph Storage Documentation
Team

Red Hat Ceph Storage 2.1 Release Notes

Release notes for Red Hat Ceph Storage 2.1

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	6
CHAPTER 5. KNOWN ISSUES	7
CHAPTER 6. NOTABLE BUG FIXES	15
CHAPTER 7. SOURCES	19

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.

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:

- ✧ Intel
- ✧ Fujitsu
- ✧ UnitedStack
- ✧ Yahoo
- ✧ UbuntuKylin
- ✧ Mellanox
- ✧ CERN
- ✧ Deutsche Telekom
- ✧ Mirantis
- ✧ SanDisk

CHAPTER 3. MAJOR UPDATES

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

Support for rebuilding Monitor store from OSD nodes

The **ceph-objectstore-tool** and **ceph-monstore-tool** utilities now enables you to rebuild the Monitor database and keyring files from OSD nodes. This ability is especially useful when all Monitors fail at the same time to boot due to an underlying **leveldb** corruption. For details, see the [All Monitors Failed to Start Because of a Corrupted Store](#) section in the Administration Guide for Red Hat Ceph Storage 2.

Collection splitting can be now done offline

This update adds the **apply-layout-settings** command to the **ceph-objectstore-tool** utility. The new command enables users to perform collection splitting offline on OSD nodes in cases when doing it online would cause too large an impact.

Ceph Object Gateway now supports static website hosting

With this feature, users can host static websites in Simple Storage Service (S3) buckets. For details, see the [Configuring Gateways for Static Web Hosting](#) chapter in the Red Hat Ceph Storage Ceph Object Gateway Guide for Red Hat Enterprise Linux.

Ceph Object Gateway now supports indexless buckets

Indexless buckets do not maintain an index and therefore require much less resources on Ceph OSDs. Consequently, they have higher performance than indexed buckets. Indexless buckets cannot be listed or replicated, therefore use them only when these capabilities are not required. For details, see the [Storage Policies](#) chapter in the Ceph Object Gateway Guide for Red Hat Enterprise Linux or the [Storage Policies](#) chapter in the Ceph Object Gateway Guide for Ubuntu.

Ceph Object Gateway now supports the Amazon "requestPayment" API call

The Amazon **requestPayment** API call enables users of a public storage object to provide data publicly without incurring payment for the retrieval of that data by other users.

"rados bench" now shows how much time it took to clean up objects

With this update, the **rados bench** command output includes a line that shows how much time it took to clean up objects:

```
Clean up completed and total clean up time :8.492848
```

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/>.

iSCSI support for the Ceph Block Device

The iSCSI gateway integrates Red Hat Ceph Storage with the iSCSI standard to provide a Highly Available (HA) iSCSI target that exports RADOS Block Device (RBD) images as SCSI disks.

For details, see the [Using an iSCSI Gateway](#) chapter in the Red Hat Ceph Storage 2 Block Device Guide.

CHAPTER 5. KNOWN ISSUES

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

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](#))

Mirroring image metadata is not supported

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

Disabling image features is incorrectly allowed on non-primary images

With RADOS Block Device (RBD) mirroring enabled, non-primary images are expected to be read-only. An attempt to disable image features on non-primary images could cause an indefinite wait. This operation should be disallowed on non-primary images.

To avoid this issue, make sure to disable image features only on the primary image. ([BZ#1353877](#))

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](#))

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. ([BZ#1330023](#))

Buckets sometimes have incorrect time stamps

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

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](#))

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](#))

The **rd** **bench write** command fails when **--io-size** is equal to the image size

The **rd bench-write --io-size <size> <image>** command fails with a segmentation fault if the size specified by the **--io-size** option is equal to the image size.

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

Setting file permissions and ownership attributes fails on existing files and directories

The NFS Ganesha file system fails to serialize and store UNIX attributes on existing files and directories. Consequently, file permissions and ownership attributes that are set after file or directory creation are not correctly stored. To avoid this problem, set file permissions and ownership attributes during file or directory creation. ([BZ#1358020](#))

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](#))

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](#))

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](#))

Ceph OSD daemons fail to initialize and DM-Multipath disks are not automatically mounted on iSCSI nodes

The **ceph-iscsi-gw.yml** Ansible playbook enables device mapper multipathing (DM-Multipath) and disables the **kpartx** utility. This behavior causes the multipath layer to claim a device before Ceph disables automatic partition setup for other system disks that use DM-Multipath. Consequently, after a reboot, Ceph OSD daemons fail to initialize, and system disks that use DM-Multipath with partitions are not automatically mounted. Because of that the system can fail to boot.

To work around this problem:

1. After executing the **ceph-iscsi-gw.yml**, log into each node that runs an iSCSI target and display the current multipath configuration:

```
$ multipath -ll
```

2. If you see any devices that you did not intend to be used by DM-Multipath, for example OSD disks, remove them from the DM-Multipath configuration.

- a. Remove the devices World Wide Identifiers (WWIDs) from the WWIDs file:

```
$ multipath -w <device_name>
```

- b. Flush the devices multipath device maps:

```
$ multipath -f device_name
```

3. Edit the `/etc/multipath.conf` file on each node that runs an iSCSI target.

- a. Comment out the `skip-partx` variable.

- b. Set the `user_friendly_names` variable to `yes`:

```
defaults {
    user_friendly_names yes
    find_multipaths no
}
```

- c. Blacklist all devices:

```
blacklist {
    devnode ".*"
}
```

- d. DM-Multipath is used with Ceph Block Devices, therefore you must add an exception for them. Edit `^rbd[0-9]` as needed:

```
blacklist_exceptions {
    devnode "^rbd[0-9]"
}
```

- e. Add the following entry for the Ceph Block Devices:

```
devices {
    device {
        vendor "Ceph"
        product "RBD"
        skip_kpartx yes
        user_friendly_names no
    }
}
```

4. Reboot the nodes. The OSD and iSCSI gateway services will initialize automatically after the reboot. ([BZ#1389484](#))

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 Ceph CLI installation

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](#))

Listing bucket info data can cause the OSD daemon to terminate unexpectedly

Due to invalid memory access in an object class operation, the **radosgw-admin bi list --max-entries=1** command in some cases causes the Ceph OSD daemon to terminate unexpectedly with a segmentation fault.

To avoid this problem, do not use the **--max-entries** option, or set its value to 2 or higher when listing bucket info data. ([BZ#1390716](#))

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](#))

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 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)

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 serial parameter must be set to 1

The **rolling_update.yml** Ansible playbook contains a comment about changing the value for the **serial** parameter to adjust the number of servers to be updated. However, upgrading many nodes in parallel can cause disruption to I/O operations. To avoid this problem, ensure that **serial** is set to **1**. (BZ#1396742)

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 update to a newer minor version or 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 problem, do not use **ceph-ansible** to update clusters with encrypted OSDs but update them by using the **Yum** utility as described in the [Upgrading from Red Hat Ceph Storage 1.3 to 2](#) section of the Red Hat Ceph Storage 2 Installation Guide for Red Hat Enterprise Linux. (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](#))

The Ceph Object Gateway fails certain Tempest tests

Currently, the Ceph Object Gateway fails the following tests of the Tempest test utility for OpenStack:

- ✦ `tempest.api.object_storage.test_object_version.ContainerTest.test_versioned_container [id-a151e158-dcbf-4a1f-a1e7-46cd65895a6f]`
- ✦ `tempest.api.object_storage.test_object_services.ObjectTest.test_delete_object [id-17738d45-03bd-4d45-9e0b-7b2f58f98687]`
- ✦ `tempest.api.object_storage.test_object_temp_url.ObjectTempUrlTest.test_put_object_using_temp_url [id-9b08dade-3571-4152-8a4f-a4f2a873a735]`
- ✦ `tempest.api.object_storage.test_object_temp_url.ObjectTempUrlTest.test_get_object_using_temp_url [id-f91c96d4-1230-4bba-8eb9-84476d18d991]` ([BZ#1252600](#))

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 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](#))

An error is returned when removing a Ceph Monitor

When removing a Ceph Monitor by using the **ceph mon remove** command, the Monitor is successfully removed but an error message similar to the following is returned:


```
Error EINVAL: removing mon.magna072 at 10.8.128.72:6789/0, there will
be 3 monitors
```

You can safely ignore this error message. ([BZ#1394495](#))

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](#))

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](#))

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](#))

iSCSI gateway setup fails if the cluster name is different than "ceph" (Technology Preview)

The **device-mapper-multipath rbd** path checker currently only supports the default cluster name, which is "ceph". As a consequence, an attempt to set up an iSCSI gateway fails during Logical Unit Number (LUN) creation if the cluster name is different than "ceph". In addition, the **ansible-playbook ceph-iscsi-gw.yml** command returns the following error:

```
Could not find dm multipath device for <image_name>.
```

To work around this problem:

1. In the **ceph-iscsi-gw** Ansible configuration file:
 - a. Set the **cluster_name** variable to **ceph**.

b. Set the **gateway_keyring** variable to **ceph.client.admin.keyring**.

2. On the **seed_monitor** host, create the following symbolic links:

✦ from **/etc/ceph/<your_cluster_name>.conf** to **/etc/ceph/ceph.conf**

```
ln -s /etc/ceph/<your_cluster_name>.conf /etc/ceph/ceph.conf
```

✦ from **/etc/ceph/<your_cluster_name>.client.admin.keyring** to **/etc/ceph/ceph.client.admin.keyring**

```
ln -s /etc/ceph/<your_cluster_name>.client.admin.keyring \
/etc/ceph/ceph.client.admin.keyring
```

([BZ#1386617](#))

Data exported though multiple iSCSI targets can be overwritten (Technology Preview)

When exporting a RADOS Block Device (RBD) image through multiple iSCSI targets, the RBD kernel module takes an exclusive lock before executing I/O requests. This behavior can prevent the module from holding the lock before the iSCSI initiator times out the request and the **multipath** layer retries the request on another target. As a consequence, I/O requests that wait for the RBD kernel module to hold an exclusive lock could be executed at a later time and overwrite newer data.

([BZ#1392124](#))

CHAPTER 6. NOTABLE BUG FIXES

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

Sync point snapshots are now properly removed after a failover or failback

In a RADOS Block Device (RBD) mirroring configuration, sync point snapshots created by the **rbd-mirror** daemon incorrectly remained on the clusters after a failover or failback. This update fixes this behavior, and sync point snapshots are now properly removed as expected. (BZ#1350003)

The ceph-disk service no longer fails because of limited resources

After installation of a new Ceph storage cluster, failed instances of the **ceph-disk** service appeared because the service was started twice: once to activate the data partition, and once to activate the journal partition. After the disk activation, one of these instances failed because of limited resources. With this update, the instance without the resources terminates with the **0** exit status and an informative message is returned. (BZ#1326740)

Bucket creation no longer fails after upgrading Red Hat Ceph Storage 1.3 to 2.0

Previously, after upgrading an Ceph Object Gateway node from Red Hat Ceph Storage 1.3 to 2.0, an attempt to create a bucket failed. This bug has been fixed, and bucket creation no longer fails in this case. (BZ#1352888)

Ceph Object Gateway now deletes large buckets in parallel

Previously, the Ceph Object Gateway was unable to delete multiple large buckets at the same time. As a consequence, the process of deleting large buckets containing millions of objects was slow. A patch has been applied, and the Ceph Object Gateway now deletes large buckets in parallel, which makes the whole process significantly faster. (BZ#1318409)

RBD mirroring is now properly disabled even if peer clusters are not removed first

Disabling 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.

Previously, if pool mirroring was disabled before removing the peer clusters, all the existing images on the mirrored pool stopped being replicated but the pool remained mirrored. As a consequence, newly created images on the pool were still replicated.

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

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

This update fixes this inconsistent state if disabling RBD mirroring is attempted before peer clusters are removed. (BZ#1362647)

Ceph Object Gateway now supports new token formats with the Keystone integration

Users can now use the following token formats for Keystone authentication:

- ✦ PKIZ - compressed Public Key Infrastructure (PKI)
- ✦ Fernet ([BZ#1300855](#))

The `radosgw-admin` command no longer returns unnecessary error messages

The log level of certain common error messages that are returned by the `radosgw-admin` command was set to `0`. Consequently, these messages appeared even though `radosgw-admin` command succeeded. The log level of these messages has been increased, and the messages no longer appear if `radosgw-admin` succeeds. ([BZ#1364353](#))

Ceph log messages are now `LOG_INFO` priority

Previously, Ceph log messages had the `LOG_DEBUG` priority. However, with the default `syslog` configuration, `LOG_DEBUG` messages do not appear in the logs. With this update, Ceph log messages have `LOG_INFO` priority, ensuring that they appear in the logs. ([BZ#1378675](#))

A restart of the `radosgw` process is no longer required after switching the zone from master to non-master

When a non-master zone was promoted to the master zone, all I/O requests became unresponsive until the `radosgw` process was restarted on both zones. Consequently, the I/O requests timed out. The underlying source code has been modified, and restarting `radosgw` is no longer required in the described situation. ([BZ#1359712](#))

Image synchronization no longer starts from the beginning after restarting `rbd-mirror`

When the `rbd-mirror` daemon was restarted during image synchronization, the synchronization started from the beginning. With this update, the sync point object number is updated periodically during the synchronization. As a result, the image synchronization no longer starts from the beginning after restarting `rbd-mirror`. ([BZ#1348940](#))

Certain maintenance image operations are no longer 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 did not properly restrict `rbd` maintenance operations against non-primary images. The affected operations included:

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

With this update, these operations are disallowed on non-primary images as expected. ([BZ#1348928](#), [BZ#1352878](#), [BZ#1349332](#))

Calamari no longer hangs when interactive commands are issued

Previously, the Calamari REST API incorrectly attached a standard input to interactive commands, such as **rbd import**. When these commands were executed from Calamari, the Calamari user interface became unresponsive because it waited for an action from the user. The underlying source code has been modified, and Calamari no longer hangs when interactive commands are issued. ([BZ#1354459](#))

FUSE clients can now be mounted permanently on Red Hat Enterprise Linux 7.3

Previously, the **util-linux** package shipped with Red Hat Enterprise Linux 7.2 did not support mounting the CephFS File System in User Space (FUSE) clients in the **/etc/fstab** file. Red Hat Enterprise Linux 7.3 now includes a new version of **util-linux** that supports mounting CephFS FUSE clients permanently in **/etc/fstab**. ([BZ#1360849](#))

Enabling write-back cache no longer causes potential data corruption

Previously, cache read requests were not properly handled if a subsection of an in-progress read operation was interrupted with an unrelated I/O operation. As a consequence, when **librbd** and **ceph-fuse** write-back cache was enabled, read requests serviced through the cache were in some cases corrupted. With this update, the incoming read result from the OSD nodes is properly reconstructed if a split is encountered. As a result, the cache return the expected data in the described situation. ([BZ#1392136](#))

S3 versioning can now be set on buckets when accessing a non-master zone

Previously, it was not possible to set Amazon Simple Storage Service (S3) object versioning on buckets when accessing any zone other than the metadata master zone. The underlying source code has been modified, and S3 versioning can now be set on buckets when accessing a non-master zone as expected. ([BZ#1350522](#))

Renaming snapshots no longer returns errors on overloaded clusters

When a Ceph Storage Cluster was overloaded and an image was in-use for I/O operations, performing rename requests sometimes took unexpectedly long. Consequently, the RADOS Block Device (RBD) CLI kept sending the rename request every 5 seconds because it received a message that the request had timed out. This caused error messages were returned in the logs of the process performing the I/O operations on the image. This update fixes this bug, and the error log messages are no longer returned in the described scenario. ([BZ#1340772](#))

Error messages are no longer returned after disabling journaling

When the **journaling** feature was disabled on an image that was previously mirrored by the **rbd-mirror** daemon, an error message similar to the following was 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 indicated that the image journal was deleted because **journaling** was disabled. The underlying source code has been modified, and the error messages are no longer returned in the described situation. ([BZ#1346946](#))

"rbd-mirror" no longer synchronizes images 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. Previously, when this happened, the **rbd-mirror** daemon started to synchronize the newly demoted image with the newly promoted image even though the image was already successfully synchronized. This behavior has been fixed, and **rbd-mirror** no longer synchronizes images from the beginning after their demotion and promotion. ([BZ#1349955](#))

CHAPTER 7. 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/>