Red Hat Ceph Storage 4.1 Release Notes

Release notes for Red Hat Ceph Storage 4.1
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. NEW FEATURES ............................................................... 5
   3.1. THE CEPH-ANSIBLE UTILITY ................................. 5
   3.2. OBJECT GATEWAY .................................................. 5
   3.3. RADOS ................................................................. 6

## CHAPTER 4. BUG FIXES ................................................................. 7
   4.1. THE CEPH-ANSIBLE UTILITY .................................. 7
   4.2. CONTAINERS .......................................................... 9
   4.3. OBJECT GATEWAY .................................................. 9
   4.4. RADOS ................................................................. 9
   4.5. BLOCK DEVICES (RBD) ......................................... 10

## CHAPTER 5. KNOWN ISSUES ............................................................ 11
   5.1. CEPH MANAGEMENT DASHBOARD ............................. 11
   5.2. THE CEPH-VOLUME UTILITY .................................. 11
   5.3. OBJECT GATEWAY MULTISITE ............................... 11
   5.4. PACKAGES ........................................................... 11

## CHAPTER 6. DEPRECATED FUNCTIONALITY ....................................... 13

## CHAPTER 7. SOURCES ................................................................. 14
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/red-hat-ceph-storage/.
CHAPTER 2. ACKNOWLEDGMENTS

Red Hat Ceph Storage version 4.1 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
This section lists all major updates, enhancements, and new features introduced in this release of Red Hat Ceph Storage.

### 3.1. THE CEPH-ANSIBLE UTILITY

The dedicated journal devices retain their configuration when migrating from Filestore OSD to Bluestore

Previously, dedicated journal devices for Filestore OSD could not be reused when migrating to Bluestore OSD DB. An example of a dedicated device configuration is using a HDD for data and an SSD for journaling.

With this update, dedicated journal devices retain their configuration during the migration, so that they can be reused with the Bluestore OSD DB.

*ceph-ansible* now supports multisite deployments with multiple realms

Previously, *ceph-ansible* multisite deployments supported a single RGW realm. With this update, *ceph-ansible* now supports multiple realms with their associated zones, zonegroups, and endpoints.

For more information, see *Configuring multisite Ceph Object Gateways* in the *Red Hat Ceph Storage Installation Guide*.

OpenStack users can deploy Ceph Dashboard with a default admin account with read-only privileges

Previously, changes made from Ceph Dashboard by OpenStack users with full admin privileges could override cluster settings or status. With this feature, Ceph Dashboard admin account can only monitor Ceph cluster status and retrieve information and settings.

*purge-container-cluster.yml* playbook now supports clusters with three-digit IDs

Previously, *purge-container-cluster* only supported Red Hat Ceph Storage clusters with up to 99 OSDs. This is because the playbook supported ceph-osd services with two-digit indices. With this update, you can properly purge clusters with three-digit IDs.

### 3.2. OBJECT GATEWAY

Support for Amazon S3 resources in Ceph Object Gateway

AWS provides the Secure Token Service (STS) to allow secure federation with existing OpenID Connect/ OAuth2.0 compliant identity services such as Keycloak. STS is a standalone REST service that provides temporary tokens for an application or user to access a Simple Storage Service (S3) endpoint after the user authenticates against an identity provider (IDP).

Previously, users without permanent Amazon Web Services (AWS) credentials could not access S3 resources through Ceph Object Gateway. With this update, Ceph Object Gateway supports STS AssumeRoleWithWebIdentity. This service allows web application users who have been authenticated with an OpenID Connect/OAuth 2.0 compliant IDP to access S3 resources through Ceph Object Gateway.

AWS S3 ListObjects v2 operation provides an improved mechanism to list the objects in the S3 bucket without additional configuration
Previously, S3 protocol clients, like S3A and the awscli command-line tool, had to be configured with the older ListObjects method. With this feature, AWS S3 ListObjects v2 operation is implemented, that provides an improved mechanism to list objects in an S3 bucket.

### 3.3. RADOS

**BlueStore compression stats added to the dashboard**

With this release, compression related performance metrics for BlueStore OSDs will now be visible in the dashboard.
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 CEPH-ANSIBLE UTILITY

Ceph installations with custom software repositories fail

Previously, using custom repositories to install Ceph were not allowed. This occurred because the `redhat_custom_repository.yml` file was removed. With this update, the `redhat_custom_repository.yml` file is included, and custom repositories can be used to install Red Hat Ceph Storage.

NOTE

Only Red Hat-signed packages can use custom software repositories to install Ceph. Custom third-party software repositories are not supported.

(BZ#1808362)

Red Hat Ceph Storage installation on Red Hat OpenStack Platform fails

Previously, the `ceph-ansible` utility became unresponsive when attempting to install Red Hat Ceph Storage with the Red Hat OpenStack Platform 16, and it returns an error similar to the following:

```
'Error: unable to exec into ceph-mon-dcn1-computehci1-2: no container with name or ID ceph-mon-dcn1-computehci1-2 found: no such container'
```

This occurred because `ceph-ansible` reads the value of the fact `container_exec_cmd` from the wrong node in handler_osds.yml.

With this update, `ceph-ansible` reads the value of `container_exec_cmd` from the correct node, and the installation proceeds successfully.

(BZ#1792320)

Ansible upgrades a multisite Ceph Object Gateway when the Dashboard is enabled

Previously, when Red Hat Ceph Storage Dashboard is enabled and an attempt to use Ansible to upgrade to a later version of Red Hat Ceph Storage is made, the upgrade to the secondary Ceph Object Gateway site in a multisite setup failed. With this update to Red Hat Ceph Storage, upgrade of the secondary site works as expected.

(BZ#1794351)

Ceph Ansible supports multiple grafana instances during a Ceph dashboard deployment

Previously, in a multi-node environment, `ceph-ansible` was not able to configure multiple `grafana` instances as only one node was supported, leaving the remaining nodes unconfigured. With this update, `ceph-ansible` supports multiple instances and injects Ceph-specific layouts on all the Ceph Monitor nodes during the deployment of the Ceph Dashboard.

(BZ#1784011)
Using a standalone nfs-ganesha daemon with an external Ceph storage cluster fails to copy the keyring during deployment

Previously, in configurations consisting of a standalone nfs-ganesha daemon and an external Ceph storage cluster, the Ceph keyring was not copied to /etc/ceph during a Ceph Ansible deployment.

With this update, the Ceph keyring is copied to /etc/ceph/ directory.

(BZ#1814942)

Ceph Ansible works with Ansible 2.9

Previously, ceph-ansible versions 4.0 and above did not work with Ansible version 2.9. This occurred because the ceph-validate role did not allow ceph-ansible to be run against Ansible 2.9. With this update, ceph-ansible works with Ansible 2.9.

(BZ#1807085)

Running the Ansible purge-cluster.yml fails with an error message when the dashboard feature is disabled

Previously, using the purge-cluster.yml playbook to purge clusters failed when the dashboard feature was disabled with the following error message:

```
registry.redhat.io/openshift4/ose-prometheus-node-exporter:v4.1
msg: '[Errno 2] No such file or directory'
```

This occurred because the dashboard_enabled variable was ignored. With this update, the dashboard_enabled variable is correctly handled, and purge-cluster.yml runs successfully.

(BZ#1785736)

Ansible unsets the norebalance flag after it completes

Previously, Ansible did not unset the norebalance flag and it had to be unset manually. With this update, the rolling-update.yml Ansible playbook unsets the norebalance flag automatically after it completes and a manual unset is not required.

(BZ#1793564)

The size of the replication pool can now be modified after the Ceph cluster deployment

Previously, increasing the size of the replication pool failed after the Ceph cluster was deployed using director. This occurred because an issue with the task in charge of customizing the pool size prevented it from executing when the playbook was rerun. With this update, you can now modify pool size after cluster deployment.

(BZ#1743242)

ceph-ansible purge playbook does not fail if dashboard related resources is not part of the deployment

Previously, when the dashboard was not deployed, purge playbook failed when purging the cluster because it tried to remove dashboard related resources that did not exist. Consequently, purge playbook stated that the dashboard is deployed. With this update, ceph-ansible does not purge dashboard related resources, if not part of the deployment.

(BZ#1808988)
4.2. CONTAINERS

The **nfs-ganesha** daemon would not start due to a missing library

Previously, a configuration using **nfs-ganesha** with the RADOS backend would not start because the **nfs-ganesha-rados-urls** library was missing. This occurred because the **nfs-ganesha** library package for the RADOS backend was moved to a dedicated package. With this update, the **nfs-ganesha-rados-urls** package is added to the Ceph container image, so the **nfs-ganesha** daemon starts successfully.

(BZ#1797075)

4.3. OBJECT GATEWAY

The **radosgw-admin bucket check** command displays incomplete multipart uploads

Previously, running the **radosgw-admin bucket check** command displayed incomplete multipart uploads. This could cause confusion for a site admin because the output might have appeared as though the bucket index were damaged. With this update, the command displays only errors and orphaned objects, and the incomplete uploads are filtered out.

(BZ#1687971)

Ceph Object Gateway properly applies AWS request signing

Previously, the Ceph Object Gateway did not properly apply an AWS request for signing headers, and was generating the following error message:

```
SignatureDoesNotMatch
```

With this release, the Ceph Object Gateway code was fixed to properly sign headers. This results in the signing request to succeed when requested.

(BZ#1665683)

Increase in overall throughput of Object Gateway lifecycle processing performance

Previously, Object Gateway lifecycle processing performance was constrained by the lack of parallelism due to the increasing workload of objects or buckets with many buckets or containers in the given environment. With this update, parallelism is in two dimensions, a single object gateway instance can have several lifecycle processing threads, and each thread has multiple work-pool threads executing the lifecycle work. Additionally, this update improved the allocation of **shards** to workers, thereby increasing overall throughput.

(BZ#1794715)

4.4. RADOS

A health warning status is reported when no Ceph Managers or OSDs are in the storage cluster

In previous Red Hat Ceph Storage releases, the storage cluster health status was **HEALTH_OK** even though there were no Ceph Managers or OSDs in the storage cluster. With this release, this health status has changed, and will report a health warning if a storage cluster is not set up with Ceph Managers, or if all the Ceph Managers go down. Because Red Hat Ceph Storage heavily relies on the Ceph Manager to deliver key features, it is not advisable to run a Ceph storage cluster without Ceph Managers or OSDs.
Small objects and files in RADOS use more space than required

The Ceph Object Gateway and the Ceph file system (CephFS) stores small objects and files as individual objects in RADOS. Previously, objects smaller than BlueStore’s default minimum allocation size (min_alloc_size) of 16 KB used more space than required. This happened because the earlier default value of BlueStore’s min_alloc_size was 16 KB for solid state devices (SSDs). Currently, the default value of min_alloc_size for SSDs is 4 KB. This enables better use of space with no impact on performance.

(BZ#1788347)

Slow ops not being logged in cluster logs

Previously, slow ops were not being logged in cluster logs. They were logged in the osd or mon logs, but lacked the expected level of detail. With this release, slow ops are being logged in cluster logs, at a level of detail that makes the logs useful for debugging.

(BZ#1807184)

4.5. BLOCK DEVICES (RBD)

If more than one `rbd unmap` command is issued concurrently, the commands might fail and the corresponding RBD block devices might remain mapped.

Previously, issuing concurrent rbd unmap commands could result in udev-related event race conditions. The commands would sporadically fail, and the corresponding RBD block devices might remain mapped to their node. With this update, the udev-related event race conditions have been fixed, and the commands no longer fail.

(BZ#1784895)
CHAPTER 5. KNOWN ISSUES

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

5.1. CEPH MANAGEMENT DASHBOARD

The Dashboard does not provide correct Ceph iSCSI error messages

If the Ceph iSCSI returns an error, for example the HTTP "400" code when trying to delete an iSCSI target while a user is logged in, the Red Hat Ceph Storage Dashboard does not forward that error code and message to the Dashboard user using the pop-up notifications, but displays a generic "500 Internal Server Error". Consequently, the message that the Dashboard provides is not informative and even misleading; an expected behavior ("users cannot delete a busy resource") is perceived as an operational failure ("internal server error"). To work around this issue, see the Dashboard logs.

(BZ#1786457)

5.2. THE CEPH-VOLUME UTILITY

Ceph OSD fails to start because udev resets the permissions for BlueStore DB and WAL devices

When specifying the BlueStore DB and WAL partitions for an OSD using the ceph-volume lvm create command or specifying the partitions, using the lvm_volume option with Ceph Ansible can cause those devices to fail on startup. The udev subsystem resets the partition permissions back to root:disk.

To work around this issue, manually start the systemd ceph-volume service. For example, to start the OSD with an ID of 8, run the following: systemctl start 'ceph-volume@lvm-8-*'. You can also use the service command, for example: service ceph-volume@lvm-8-4c6ddc44-9037-477d-903c-63b5a789ade5 start. Manually starting the OSD results in the partition having the correct permission, ceph:ceph.

(BZ#1822134)

5.3. OBJECT GATEWAY MULTISITE

The radosgw-admin commands that create and modify users are not allowed in secondary zones for multi-site Ceph Object Gateway environments

Using the radosgw-admin commands to create or modify users and subusers on the secondary zone does not propagate those changes to the master zone, even if the --yes-i-really-mean-it option was used.

To workaround this issue, use the REST APIs instead of the radosgw-admin commands. The REST APIs enable you to create and modify users in secondary zone, and then propagate those changes to the master zone.

(BZ#1553202)

5.4. PACKAGES

Current version of Grafana causes certain bugs in the Dashboard


Red Hat Ceph Storage 4 uses the Grafana version 5.2.4. This version causes the following bugs in the Red Hat Ceph Storage Dashboard:

- When navigating to **Pools > Overall Performance**, Grafana returns the following error:

  ```
  TypeError: l.c[t.type] is undefined
  true
  ```

- When viewing a pool’s performance details (**Pools > select a pool from the list > Performance Details**) the Grafana bar is displayed along with other graphs and values, but it should not be there.

These bugs will be fixed after rebasing to a newer Grafana version in a future release of Red Hat Ceph Storage.

(BZ#1786107)
CHAPTER 6. 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.

Ubuntu is no longer supported

Installing a Red Hat Ceph Storage 4 cluster on Ubuntu is no longer supported. Use Red Hat Enterprise Linux as the underlying operating system.
CHAPTER 7. SOURCES

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

- For Red Hat Enterprise Linux 7:

- For Red Hat Enterprise Linux 8: