



# Red Hat Ceph Storage 3.2

## Release Notes

Release notes for Red Hat Ceph Storage 3.2



# Red Hat Ceph Storage 3.2 Release Notes

---

Release notes for Red Hat Ceph Storage 3.2

## Legal Notice

Copyright © 2019 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, the Red Hat logo, JBoss, OpenShift, Fedora, the Infinity logo, and RHCE are trademarks of Red Hat, Inc., registered in the United States and other countries.

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

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

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

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

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

The OpenStack<sup>®</sup> 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

<b>CHAPTER 1. INTRODUCTION</b> .....	<b>3</b>
<b>CHAPTER 2. ACKNOWLEDGMENTS</b> .....	<b>4</b>
<b>CHAPTER 3. NEW FEATURES</b> .....	<b>5</b>
3.1. THE CEPH-ANSIBLE UTILITY	5
3.2. CEPH FILE SYSTEM	6
3.3. THE CEPH-VOLUME UTILITY	7
3.4. CONTAINERS	7
3.5. DISTRIBUTION	7
3.6. ISCSI GATEWAY	7
3.7. OBJECT GATEWAY	8
3.8. OBJECT GATEWAY MULTISITE	8
3.9. PACKAGES	8
3.10. RADOS	8
3.11. BLOCK DEVICES (RBD)	9
<b>CHAPTER 4. BUG FIXES</b> .....	<b>10</b>
4.1. THE CEPH-ANSIBLE UTILITY	10
4.2. CEPH MANAGEMENT DASHBOARD	11
4.3. CEPH FILE SYSTEM	12
4.4. CEPH MANAGER PLUGINS	13
4.5. THE CEPH-VOLUME UTILITY	13
4.6. CONTAINERS	13
4.7. OBJECT GATEWAY	13
4.8. OBJECT GATEWAY MULTISITE	17
4.9. RADOS	18
4.10. BLOCK DEVICES (RBD)	20
<b>CHAPTER 5. TECHNOLOGY PREVIEWS</b> .....	<b>21</b>
5.1. BLOCK DEVICES (RBD)	21
5.2. CEPH FILE SYSTEM	21
5.3. OBJECT GATEWAY	21
<b>CHAPTER 6. KNOWN ISSUES</b> .....	<b>23</b>
6.1. THE CEPH-ANSIBLE UTILITY	23
6.2. CEPH MANAGEMENT DASHBOARD	24
6.3. CEPH FILE SYSTEM	26
6.4. THE CEPH-VOLUME UTILITY	26
6.5. ISCSI GATEWAY	26
6.6. OBJECT GATEWAY	27
6.7. RADOS	27
<b>CHAPTER 7. DEPRECATED FUNCTIONALITY</b> .....	<b>29</b>
7.1. THE CEPH-ANSIBLE UTILITY	29
<b>CHAPTER 8. SOURCES</b> .....	<b>30</b>



## 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 3.2 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
- SUSE



## CHAPTER 3. NEW FEATURES

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

The main features added by this release are:

- [General availability of the BlueStore OSD back end](#)
- [Technology preview of erasure coding for Ceph File System](#)

### 3.1. THE CEPH-ANSIBLE UTILITY

#### Ansible now configures `firewalld` by default

The **ceph-ansible** utility now configures the **firewalld** service by default when creating a new cluster. Previously, it only checked if required ports were opened or closed, but it did not configure any firewall rules.

#### Pool size can now be customized when deploying clusters with **ceph-ansible**

Previously, the **ceph-ansible** utility set the pool size to 3 by default and did not allow the user to change it. However, in Red Hat OpenStack deployments, setting the size of each pool is sometimes required. With this update, the pool size can be customized. To do so, change the **size** setting in the **all.yml** file. Each time, the value of **size** is changed, a new size is applied.

#### Ansible now validates CHAP settings before running playbooks

Previously, when the Challenge Handshake Authentication Protocol (CHAP) settings were set incorrectly, the **ceph-ansible** utility returned an unclear error message during deploying Ceph iSCSI gateway. With this update, **ceph-ansible** validates the CHAP settings before deploying Ceph iSCSI gateways.

#### The **noup** flag is now set before creating OSDs to distribute PGs properly

The **ceph-ansible** utility now sets the **noup** flag before creating OSDs to prevent them from changing their status to **up** before all OSDs are created. Previously, if the flag was not set, placement groups (PGs) were created on only one OSD and got stuck in creation or activation. With this update, the **noup** flag is set before creating OSDs and unset after the creation is complete. As a result, PGs are distributed properly among all OSDs.

#### Variables are now validated at the beginning of an invocation of **ceph-ansible** playbooks

The **ceph-ansible** utility now validates variables specified in configuration files located in the **group\_vars** or **host\_vars** directories at the beginning of playbooks invocation. This change makes it easier to discover misconfigured variables.

#### Ceph Ansible supports a multi-site Ceph Object Gateway configuration

With previous versions of **ceph-ansible**, only one Object Gateway endpoint was configurable. With this release, **ceph-ansible** supports a multi-site Ceph Object Gateway for multiple endpoints. Zones can be configured with multiple Object Gateways and can be added to a zone automatically by appending their endpoint information to a list. With the **rgw\_multisite\_proto** option, users can set it to **http** or **https** depending on whether the endpoint is configured to use SSL or not.

When more than one Ceph Object Gateway is in the master zone or in the secondary zone, then the **rgw\_multisite\_endpoints** option needs to be set. The **rgw\_multisite\_endpoints** option is a comma separated list, with no spaces. For example:

```
rgw_multisite_endpoints:  
http://foo.example.com:8080,http://bar.example.com:8080,http://baz.example.com:8080
```

When adding a new Object Gateway, append it to the end of the **rgw\_multisite\_endpoints** list with the endpoint URL of the new Object Gateway before running the Ansible playbook.

### Ansible now has the ability to start OSD containers using numactl

With this update, the **ceph-ansible** utility has the ability to start OSD containers using the **numactl** utility. **numactl** allows use of the **--preferred** option, which means the program can allocate memory outside of the NUMA socket and running out of memory causes less problems.

## 3.2. CEPH FILE SYSTEM

### A new subcommand: drop\_cache

The **ceph tell** command now supports the **drop\_cache** subcommand. Use this subcommand to drop Metadata Server (MDS) cache without restarting, trim its journal, and ask clients to drop all capabilities that are not in use.

### New option: mds\_cap\_revoke\_eviction\_timeout

This update adds a new configurable timeout for evicting clients that have not responded to capability revoke request by the Metadata Server (MDS). MDS can request clients to release its capabilities under certain conditions, such as another client requesting a capability that is currently held by a client. The client then releases its capabilities and acknowledges the MDS which can handover the capability to other clients. However, a misbehaving client might not acknowledge or could totally ignore the capability revoke request by the MDS, causing other clients to wait and thereby stalling requested I/O operations. Now, MDS can evict clients that have not responded to capability revoke requests for a configured timeout. This is disabled by default and can be enabled by setting the **mds\_cap\_revoke\_eviction\_timeout** configuration parameter.

### SELinux support for CephFS

This update adds the SELinux policy for the Metadata Server (MDS) and **ceph-fuse** daemons so that users can use Ceph File System (CephFS) with SELinux in enforcing mode.

### MDS now records the IP address and source port for evicted clients

The Red Hat Ceph Storage Metadata Server (MDS) now logs the IP address and source port for evicted clients. If you want to correlate client evictions with machines, review the cluster log for this information.

### Improved logging for Ceph MDS

Now, the Ceph MetaData Server (MDS) outputs more metrics concerning client sessions by default to the debug log. This includes the creation of the client session and other metadata. This information is useful for storage administrators to see when a new client session is created and how long it took to establish a connection.

### *session\_timeout* and *session\_autoclose* are now configurable by *ceph fs set*

You can now configure the **session\_timeout** and **session\_autoclose** options by using the **ceph fs set** command instead of setting them in the Ceph configuration file.

## 3.3. THE CEPH-VOLUME UTILITY

### Specifying more than one OSD per device is now possible

With this version, a new **batch** subcommand has been added. The **batch** subcommand includes the **--osds-per-device** option that allows specifying multiple OSD per device. This is especially useful when using high-speed devices, such as Non-volatile Memory Express (NVMe).

### New subcommand: `ceph-volume lvm batch`

This update adds the **ceph-volume lvm batch** subcommand that allows creation of volume groups and logical volumes for OSD provisioning from raw disks. The **batch** subcommand makes creating logical volumes easier for users who are not familiar with the Logical Volume Manager (LVM). With **batch**, one or many OSDs can be created by passing an array of devices and an OSD count per device to the **ceph-volume lvm batch** command.

## 3.4. CONTAINERS

### Support the iSCSI gateway in containers

Previously, the iSCSI gateway could not be run in a container. With this update to Red Hat Ceph Storage, a containerized version of the Ceph iSCSI gateway can be deployed with a containerized Ceph cluster.

## 3.5. DISTRIBUTION

### **nfs-ganesha** rebased to 2.7

The **nfs-ganesha** package has been upgraded to upstream version 2.7, which provides a number of bug fixes and enhancements over the previous version.

## 3.6. ISCSI GATEWAY

### Target-level control parameters can be now overridden

Only if instructed to by Red Hat Support, the following configuration settings can now be overridden by using the **gwcli reconfigure** subcommand:

- `cmds_n_depth`
- `immediate_data`
- `initial_r2t`
- `max_outstanding_r2t`
- `first_burst_length`
- `max_burst_length`
- `max_recv_data_segment_length`
- `max_xmit_data_segment_length`

Tuning these variables might be useful for high IOPS/throughput environments. **Only** use these variables if instructed to by Red Hat Support

## Automatic rotation of iSCSI logs

This update implements automatic log rotation for the **rbd-target-gw**, **rbd-target-api**, and **tcmu-runner** daemons that are used by Ceph iSCSI gateways.

## 3.7. OBJECT GATEWAY

### Changed the `reshard_status` output

Previously, the **radogw-admin reshard status --bucket <bucket\_name>** command displayed a numerical value for the **reshard\_status** output. These numerical values corresponded with an actual status, as follows:

```
CLS_RGW_RESHARD_NONE      = 0
CLS_RGW_RESHARD_IN_PROGRESS = 1
CLS_RGW_RESHARD_DONE      = 2
```

In this release, these numerical values were replaced by the actual status.

## 3.8. OBJECT GATEWAY MULTISITE

### New performance counters added

This update adds the following performance counters to multi-site configuration of the Ceph Object Gateway to measure data sync:

- **poll\_latency** measures the latency of requests for remote replication logs.
- **fetch\_bytes** measures the number of objects and bytes fetched by data sync.

## 3.9. PACKAGES

### ceph rebased to 12.2.8

The **ceph** package has been upgraded to upstream version 12.2.8, which provides a number of bug fixes and enhancements over the previous version.

## 3.10. RADOS

### OSD BlueStore is now fully supported

BlueStore is a new back end for the OSD daemons that allows for storing objects directly on the 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](#) chapter in the Administration Guide for Red Hat Ceph Storage 3.

### New option: `osd_scrub_max_preemptions`

With this release a new **osd\_scrub\_max\_preemptions** option has been added. This option sets the maximum number of times Ceph preempts a deep scrub due to a client operation before blocking the client I/O to complete the scrubbing process. The option is set to 5 by default.

### Offline splitting FileStore directories to a target hash level is now supported

The **ceph-objectstore-tool** utility now supports splitting FileStore directories to a target hash level.

### New option: **osd\_memory\_target**

A new option, **osd\_memory\_target**, has been added with the release. This option sets a target memory size for OSDs. The BlueStore back end adjusts its cache size and attempts to stay close to this target. The **ceph-ansible** utility automatically adjusts **osd\_memory\_target** based on host memory. The default value is 4 GiB. The **osd\_memory\_target** option is set differently for Hyper-converged infrastructure (HCI) and non-HCI setups. To differentiate between them, use the **is\_hci** configuration parameter. This parameter is set to **false** by default. To change the default values of **osd\_memory\_target** and **is\_hci**, set them in the **all.yml** file.

### New options: **osd\_delete\_sleep**, **osd\_remove\_threads**, and **osd\_recovery\_threads**

This update adds a new configuration option, **osd\_delete\_sleep** to throttle object delete operations. In addition, the **osd\_disk\_threads** option has been replaced with the **osd\_remove\_threads** and **osd\_recovery\_threads** options so that users can separately configure the threads for these tasks. These changes help to throttle the rate of object delete operations to reduce the impact on client operations. This is especially important when migrating placement groups (PGs). When using these options, every removal thread sleeps for the number of seconds specified between small batches of removal operations.

### Upgrading to the latest version no longer causes cluster data movement

Previously, upgrading a Red Hat Ceph Storage cluster to the latest version when CRUSH device classes were enabled, the **crushtool** utility rebalanced data in the cluster because of changes in the CRUSH map. This data movement should not have occurred. With this update, a reclassify functionality is available to help transition from older CRUSH maps that maintains parallel hierarchies for OSDs of different types to a modern CRUSHmap that makes use of the device class feature without triggering data movement.

## 3.11. BLOCK DEVICES (RBD)

### Support for RBD mirroring to multiple secondary clusters

Mirroring RADOS Block Devices (RBD) from one primary cluster to multiple secondary clusters is now fully supported.

### **rbd ls** now uses IEC units

The **rbd ls** command now uses International Electrotechnical Commission (IEC) units to display image sizes.

## 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

#### **osd\_scenario: lvm now works when deploying Ceph in containers**

Previously, the **lvm** installation scenario did not work when deploying a Ceph cluster in containers. With this update, the **osd\_scenario: lvm** installation method is supported as expected in this situation.

([BZ#1509230](#))

#### **The --limit mdss option now creates CephFS pools as expected**

Previously, when deploying the Metadata Server (MDS) nodes by using the Ansible and the **--limit mdss** option, Ansible did not create the Ceph File System (CephFS) pools. This bug has been fixed, and Ansible creates the CephFS pools as expected.

([BZ#1518696](#))

#### **Ceph Ansible no longer fails if network interface names include dashes**

When **ceph-ansible** makes an inventory of network interfaces if they have a dash ( - ) in the name the inventory must convert the dashes to underscores ( \_ ) in order to use them. In some cases conversion did not occur and Ceph installation failed. With this update to Red Hat Ceph Storage, all dashes in the names of network interfaces are converted in the facts and installation completes successfully.

([BZ#1540881](#))

#### **Ansible now sets container and service names that correspond with OSD numbers**

When containerized Ceph OSDs were deployed with the **ceph-ansible** utility, the resulting container names and service names of the OSDs did not correspond in any way to the OSD number and were thus difficult to find and use. With this update, **ceph-ansible** has been improved to set container and service names that correspond with OSD numbers. Note that this change does not affect existing deployed OSDs.

([BZ#1544836](#))

#### **Expanding clusters deployed with osd\_scenario: lvm works**

Previously, the **ceph-ansible** utility could not expand a cluster that was deployed by using the **osd\_scenario: lvm** option. The underlying source code has been modified, and clusters deployed with **osd\_scenario: lvm** can be expanded as expected.

([BZ#1564214](#))

#### **Ansible now stops and disables the iSCSI gateway services when purging the Ceph iSCSI gateway**

Previously, the **ceph-ansible** utility did not stop and disable the Ceph iSCSI gateway services when using the **purge-iscsi-gateways.yml** playbook. Consequently, the services had to be stopped manually. The playbook has been improved, and the iSCSI services are now stopped and disabled as expected when purging the iSCSI gateway.

([BZ#1621255](#))

### The values passed into devices in `osds.yml` are now validated

Previously in the `osds.yml` of the Ansible playbook, the values passed into the `devices` parameter were not validated. This caused errors when `ceph-disk`, `parted`, or other device preparation tools failed to operate on devices that did not exist. It also caused errors if the number of values passed into the `dedicated_devices` parameter was not equal to the number of values passed into `devices`. With this update, the values are validated as expected, and none of the above mentioned errors occur.

([BZ#1648168](#))

### Purging clusters using `ceph-ansible` deletes logical volumes as expected

When using the `ceph-ansible` utility to purge a cluster that deployed OSDs with the `ceph-volume` utility, the logical volumes were not deleted. This behavior caused logical volumes to remain in the system after the purge process completed. This bug has been fixed, and purging clusters using `ceph-ansible` deletes logical volumes as expected.

([BZ#1653307](#))

### The `--limit osds` option now works as expected

Previously, an attempt to add OSDs by using the `--limit osds` option failed on container setup. The underlying source code has been modified, and adding OSDs with `--limit osds` works as expected.

([BZ#1670663](#))

### Increased CPU CGroup limit for containerized Ceph Object Gateway

The default CPU CGroup limit for containerized Ceph Object Gateway (RGW) was very low and has been increased with this update to be more reasonable for typical Hard Disk Drive (HDD) production environments. However, consider evaluating what limit to set for the site's configuration and workload. To customize the limit, adjust the `ceph_rgw_docker_cpu_limit` parameter in the Ansible `group_vars/rgws.yml` file.

([BZ#1680171](#))

### SSL works as expected with containerized Ceph Object Gateways

Previously, the SSL configuration in containerized Ceph Object Gateways did not work because the Certificate Authority (CA) certificate was only added to the TLS bundle on the hypervisor and was not propagated to the Ceph Object Gateway container due to missing container bind mounts on the `/etc/pki/ca-trusted/` directory. This bug has been fixed, and SSL works as expected with containerized Ceph Object Gateways.

([BZ#1684283](#))

### The `rolling-upgrade.yml` playbook now restarts all OSDs as expected

Due to a bug in a regular expression, the `rolling-upgrade.yml` playbook did not restart OSDs that used Non-volatile Memory Express devices. The regular expression has been fixed, and `rolling-upgrade.yml` now restarts all OSDs as expected.

([BZ#1687828](#))

## 4.2. CEPH MANAGEMENT DASHBOARD

### **The OSD node details are now displayed in the Host OSD Breakdown panel as expected**

Previously, in the Red Hat Ceph Storage Dashboard, the **Host OSD Breakdown** information was not displayed on the **OSD Node Detail** panel under the **All OSD Overview** section. With this update, the underlying issue has been fixed, and the OSD node details are displayed as expected.

([BZ#1610876](#))

## **4.3. CEPH FILE SYSTEM**

### **The Ceph Metadata Server no longer allows recursive stat rctime to go backwards**

Previously, the Ceph Metadata Server used the client's time to update rctime. But because client time may not be synchronized with the MDS, the inode rctime could go backwards. The underlying source code has been modified, and the Ceph Metadata Server no longer allows recursive stat rctime to go backwards.

([BZ#1632506](#))

### **The ceph-fuse client no longer indicates incorrect recursive change time**

Previously, the **ceph-fuse** client did not update change time when file content was modified. Consequently, incorrect recursive change time was indicated. With this update, the bug has been fixed, and the client now indicates the correct change time.

([BZ#1632509](#))

### **The Ceph MDS no longer allows dumping of cache larger than 1 GB**

Previously, if you attempted to dump a Ceph Metadata Server (MDS) cache with a size of around 1 GB or larger, the MDS could terminate unexpectedly. With this update, MDS no longer allows dumping of cache that size so the MDS no longer terminates in the described situation.

([BZ#1636037](#))

### **When Monitors cannot reach an MDS, they no longer incorrectly mark its rank as damaged**

Previously, Monitors were evicting and fencing an unreachable Metadata Server (MDS), then MDS was signaling that its rank was damaged due to improper handling of blacklist errors. Consequently, Monitors were incorrectly marking the rank as damaged, and the file system became unavailable because of one or more damaged ranks. In this release, the Monitors are setting the correct rank.

([BZ#1652464](#))

### **The reconnect timeout for MDS clients has been extended**

When the Metadata Server (MDS) daemon was handling a large number of reconnecting clients with a huge number of capabilities to aggregate, the reconnect timeout was reached. Consequently, the MDS rejected clients that attempted to reconnect. With this update, the reconnect timeout has been extended, and MDS now handles reconnecting clients as expected in the described situation.

([BZ#1656969](#))

### **Shrinking large MDS cache no longer causes the MDS daemon to appear to hang**

Previously, an attempt to shrink a large Metadata Server (MDS) cache caused the primary MDS daemon to become unresponsive. Consequently, Monitors removed the unresponsive MDS and a standby MDS became the primary MDS. With this update, shrinking large MDS cache no longer causes the primary MDS daemon to hang.



([BZ#1664468](#))

## 4.4. CEPH MANAGER PLUGINS

### HDD and SSD devices can now be mixed when accessing the `/osd` endpoint

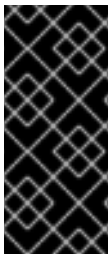
Previously, the Red Hat Ceph Storage RESTful API did not handle when HDD and SSD devices were mixed when accessing the `/osd` endpoint and returned an error. With this update, the OSD traversal algorithm has been improved to handle this scenario as expected.

([BZ#1594746](#))

## 4.5. THE CEPH-VOLUME UTILITY

### `ceph-volume` does not break custom named clusters

When using a custom storage cluster name other than `ceph`, the OSDs could not start after a reboot. With this update, `ceph-volume` provisions OSDs in a way that allows them to boot properly when a custom name is used.



#### IMPORTANT

Despite this fix, Red Hat does not support clusters with custom names. This is because the upstream Ceph project removed support for custom names in the Ceph OSD, Monitor, Manager, and Metadata server daemons. The Ceph project removed this support because it added complexities to systemd unit files. This fix was created before the decision to remove support for custom cluster names was made.

([BZ#1621901](#))

## 4.6. CONTAINERS

### Deploying encrypted OSDs in containers by using `ceph-disk` works as expected

When attempting to deploy a containerized OSD by using the `ceph-disk` and `dmccrypt` utilities, the container process failed to start because the OSD ID could not be found by the mounts table. With this update, the OSD ID is correctly determined, and the container process no longer fails.

([BZ#1695852](#))

## 4.7. OBJECT GATEWAY

### CivetWeb was rebased to upstream version 1.10 and the `enable_keep_alive` CivetWeb option works as expected

When using the Ceph Object Gateway with the CivetWeb front end, the CivetWeb connections timed out despite the `enable_keep_alive` option enabled. Consequently, S3 clients that did not reconnect or retry were not reliable. With this update, CivetWeb has been updated, and the `enable_keep_alive` option works as expected. As a result, CivetWeb connections no longer time out in this case.

In addition, the new CivetWeb version introduces more strict header checks. This new behavior can cause certain return codes to change because invalid requests are detected sooner. For example, in previous version CivetWeb returned the **403 Forbidden** error on an invalid HTTP request, but in the new version it returns the **400 Bad Request** error instead.

[\(BZ#1670321\)](#)

### Red Hat Ceph Storage passes the Swift Tempest test in the RefStack 15.0 toolset

Various improvements have been made to the Ceph Object Gateway Swift service. As a result, when configured correctly, Red Hat Ceph Storage 3.2, which includes the **ceph-12.2.8** package, passes the Swift Tempest **tempest.api.object\_storage** test suite with the exception of the **test\_container\_synchronization** test case. Red Hat Ceph Storage includes a different synchronization model, multisite operations, for users who require that feature.

[\(BZ#1436386\)](#)

### Mounting the NFS Ganesha file server in a containerized IPv6 cluster no longer fails

When a containerized IPv6 Red Hat Ceph Storage cluster with an **nfs-ganesha-rgw** daemon was deployed by using the **ceph-ansible** utility, an attempt to mount the NFS Ganesha file server on a client failed with the **Connection Refused** error. Consequently, I/O requests were unable to run. This update fixes the default configuration IPv6 connections, and mounting the NFS Ganesha server works as expected in this case.

[\(BZ#1566082\)](#)

### Stale lifecycle configuration data of deleted buckets no longer persists in OMAP consuming space

Previously, in the Ceph Object Gateway (RGW), incorrect key formatting in the **RGWDeleteLC::execute()** function caused bucket lifecycle configuration metadata to persist after the deletion of the corresponding bucket. This caused stale lifecycle configuration data to persist in **OMAP** consuming space. With this update, the correct name for the lifecycle object is now used in **RGWDeleteLC::execute()**, and the lifecycle configuration is removed as expected on removal of the corresponding bucket.

[\(BZ#1588731\)](#)

### The Keystone credentials were moved to an external file

When using the Keystone identity service to authenticate a Ceph Object Gateway user, the Keystone credentials were set as plain text in the Ceph configuration file. With this update, the Keystone credentials are configured in an external file that only the Ceph user can read.

[\(BZ#1637529\)](#)

### Wildcard policies match objects with colons in the name

Previously, using colons in the name caused an error in a matching function not allowing wildcards to match beyond colons. In this release, colons can be used to match objects.

[\(BZ#1650674\)](#)

### Lifecycle rules with multiple tag filters are no longer rejected

Due to a bug in lifecycle rule processing, an attempt to install the lifecycle rules with multiple tag filters was rejected and the **InvalidRequest** error message was returned. With this update, other rule forms are used, and lifecycle rules with multiple tag filters are no longer rejected.

[\(BZ#1654588\)](#)

### An object can no longer be deleted even if a bucket or user policy with DENY s3:DeleteObject exists

Previously, this issue was caused by an incorrect value being returned by a method which evaluates policies. In this release, the correct value is being returned.

([BZ#1654694](#))

### The Ubuntu `nfs_ganesha` package did not install the `systemd` unit file properly

When running `systemctl enable nfs-ganesha` the following error would be printed: **Failed to execute operation: No such file or directory**. This was because the `nfs-ganesha-lock.service` file was not created properly. With this release, the file is created properly and the `nfs-ganehsa` service can be enabled successfully.

([BZ#1660063](#))

### The Ceph Object Gateway supports a string as a delimiter

Invalid logic was used to find and project a delimiter sequence greater than one character. This was causing the Ceph Object Gateway to fail any request with a string as the delimiter, returning an **invalid utf-8 character** message. The logic handling the delimiter has been replaced by an 8-bit shift-carry equivalent. As a result, a string delimiter will work correctly. Red Hat has only tested this against the **US-ascii** character set.

([BZ#1660962](#))

### Mapping NFS exports to Object Gateway tenant user IDs works as expected

Previously, the NFS server for the Ceph Object Gateway (`nfs-ganesha`) did not correctly map Object Gateway tenants into their correct namespace. As a consequence, an attempt to map an NFS export onto Ceph Object Gateway with a tenanted user ID silently failed; the account could authenticate and NFS mounts could succeed, but the namespace did not contain buckets and objects. This bug has been fixed, and tenanted mappings are now set correctly. As a result, NFS exports can now be mapped to Object Gateway tenant user IDs and buckets and objects are visible as expected in the described situation.

([BZ#1661882](#))

### An attempt to get bucket ACL for non-existing bucket returns an error as expected

Previously, an attempt to get bucket Access Control Lists (ACL) for a non-existent bucket by calling the `GetBucketAcl()` function returned a result instead of returning a **NoSuchBucket** error. This bug has been fixed, and the **NoSuchBucket** error is returned in the aforementioned situation.

([BZ#1667142](#))

### The log level for `gc_iterate_entries` has been changed to 10

Previously, the log level for the `gc_iterate_entries` log message was set to 0. As a consequence, OSD log files included unnecessary information and could grow significantly. With this update, the log level for `gc_iterate_entries` has been changed to 10.

([BZ#1671169](#))

### Garbage collection no longer consumes bandwidth without making forward progress

Previously, some underlying bugs prevented garbage collection (GC) from making forward progress. Specifically, the marker was not always being advanced, GC was unable to process entries with zero-length chains, and the truncated flag was not always being set correctly. This caused GC to consume bandwidth without making any forward progress, thereby not freeing up disk space, slowing down other

cluster work, and allowing OMAP entries related to GC to continue to increase. With this update, the underlying bugs have been fixed, and GC is able to make progress as expected freeing up disk space and OMAP entries.

([BZ#1674436](#))

### The `radosgw-admin` utility no longer gets stuck and creates high read operations when creating greater than 999 buckets per user

An issue with a limit check caused the `radosgw-admin` utility to never finish when creating 1,000 or more buckets per user. This problem has been fixed and `radosgw-admin` no longer gets stuck or creates high read operations.

([BZ#1679263](#))

### LDAP authentication is available again

Previously, a logic error caused LDAP authentication checks to be skipped. Consequently, the LDAP authentication was not available. With this update, the checks for a valid LDAP authentication setup and credentials have been fixed, and LDAP authentication is available again.

([BZ#1687800](#))

### NFS Ganesha no longer aborts when an S3 object name contains a `//` sequence

Previously, the NFS server for the Ceph Object Gateway (**RGW NFS**) would abort when an S3 object name contained a `//` sequence. With this update, **RGW NFS** ignores such sequence as expected and no longer aborts.

([BZ#1687970](#))

### Expiration time is calculated the same as S3

Previously, a Ceph Object Gateway computed relative object's life cycle expiration rules from the time of creation, rather than rounded to midnight UTC as in AWS. This could cause the following error: **botocore.exceptions.ClientError: An error occurred (InvalidArgument) when calling the PutBucketLifecycleConfiguration operation: 'Date' must be at midnight GMT**. Expiration is now rounded to midnight UTC for greater AWS compatibility.

([BZ#1688330](#))

### Operations waiting for resharding to complete are able to complete after resharding

Previously, when using dynamic resharding, some operations that were waiting to complete after resharding failed to complete. This was due to code changes to the Ceph Object Gateway when automatically cleaning up no longer used bucket index shards. While this reduced storage demands and eliminated the need for manual clean up, the process removed one source of an identifier needed for operations to complete after resharding. The code has been updated so that identifier is retrieved from a different source after resharding and operations requiring it can now complete.

([BZ#1688378](#))

### `radosgw-admin bi put` now sets the correct `mtime` time stamp

Previously, the `radosgw-admin bi put` command did not set the `mtime` time stamp correctly. This bug has been fixed.

([BZ#1688541](#))

### Ceph Object Gateway lifecycle works properly after a bucket is resharded

Previously, after a bucket was resharded using the dynamic resharding feature, if a lifecycle policy was applied to the bucket, it did not complete and the policy failed to update the bucket. With this update to Red Hat Ceph Storage, a lifecycle policy is properly applied after resharding of a bucket.

([BZ#1688869](#))

### The RGW server no longer returns an incorrect S3 error code **NoSuchKey** when asked to return non-existent CORS rules

Previously, the Ceph Object Gateway (RGW) sever would return an incorrect S3 error code **NoSuchKey** when asked to return non-existent **CORS** rules. This caused the **s3cmd** tool and other programs to misbehave. With this update, the RGW server now returns **NoSuchCORSConfiguration** for this case, and the **s3cmd** tool and other programs that expect this error behave correctly.

([BZ#1689410](#))

### Decrypting multipart uploads was corrupting data

When doing multipart uploads with SSE-C, the part size was not a multiple of the 4k encryption block size. While the multipart uploads were encrypted correctly, the decryption process failed to account for the part boundaries and was returning corrupted data. With this release, the decryption process correctly handles the part boundaries when using SSE-C. As a result, all encrypted multipart uploads can be successfully decrypted.

([BZ#1690941](#))

## 4.8. OBJECT GATEWAY MULTISITE

### Redundant multi-site replication sync errors were moved to debug level 10

A few multi-site replication sync errors were logged multiple times at log level 0 and consumed extra space in logs. This update moves the redundant messages to debug level 10 to hide them from the log.

([BZ#1635381](#))

### Buckets with false entries can now be deleted as expected

Previously, bucket indices could include "false entries" that did not represent actual objects and that resulted from a prior bug. Consequently, during the process of deleting such buckets, encountering a false entry caused the process to stop and return an error code. With this update, when a false entry is encountered, Ceph ignores it, and deleting buckets with false entries works as expected.

([BZ#1658308](#))

### Datalogs are now trimmed regularly as expected

Due to a regression in decoding of the JSON format of data sync status objects, automated datalog trimming logic was unable to query the sync status of its peer zones. Consequently, the datalog trimming process did not progress. This update fixes the JSON decoding and adds more regression test coverage for log trimming. As a result, datalogs are now trimmed regularly as expected.

([BZ#1662353](#))

### Objects are now synced correctly in versioning-suspended buckets

Due to a bug in multi-site sync of versioning-suspended buckets, certain object versioning attributes

were overwritten with incorrect values. Consequently, the objects failed to sync and attempted to retry endlessly, blocking further sync progress. With this update, the sync process no longer overwrites versioning attributes. In addition, any broken attributes are now detected and repaired. As a result, objects are synced correctly in versioning-suspended buckets.

([BZ#1663570](#))

### Objects are now synced correctly in versioning-suspended buckets

Due to a bug in multi-site sync of versioning-suspended buckets, certain object versioning attributes were overwritten with incorrect values. Consequently, the objects failed to sync and attempted to retry endlessly, blocking further sync progress. With this update, the sync process no longer overwrites versioning attributes. In addition, any broken attributes are now detected and repaired. As a result, objects are synced correctly in versioning-suspended buckets.

([BZ#1690927](#))

### Buckets with false entries can now be deleted as expected

Previously, bucket indices could include "false entries" that did not represent actual objects and that resulted from a prior bug. Consequently, during the process of deleting such buckets, encountering a false entry caused the process to stop and return an error code. With this update, when a false entry is encountered, Ceph ignores it, and deleting buckets with false entries works as expected.

([BZ#1690930](#))

### radosgw-admin sync status now shows timestamps for master zone

Previously in Ceph Object Gateway multisite, running **radosgw-admin sync status** on the master zone did not show timestamps, which made it difficult to tell if data sync was making progress. This bug has been fixed, and timestamps are shown as expected.

([BZ#1692555](#))

### Synchronizing a multi-site Ceph Object Gateway was getting stuck

When recovering versioned objects, other operations were unable to finish. These stuck operations were caused by the removing of expired **user.rgw.olh.pending** extended attributes (xattrs) all at once on those versioned objects. Another [bug](#) was causing too many of the **user.rgw.olh.pending** xattrs to be written to those recovering versioned objects. With this release, batches of expired xattrs are removed instead of all at once. This results in versioned objects recovering correctly so other operations can proceed normally.

([BZ#1693445](#))

### A multi-site Ceph Object Gateway is not trimming the data and bucket index logs

Configuring zones for a multi-site Ceph Object Gateway without setting the **sync\_from\_all** option, was causing the data and bucket index logs not to be trimmed. With this release, the automated trimming process only consults the synchronization status of peer zones that are configured to synchronize. As result, this allows the data and bucket index logs to be trimmed properly.

([BZ#1699478](#))

## 4.9. RADOS

### A PG repair no longer sets the storage cluster to a warning state

When doing a repair of a placement group (PG) it was considered a damaged PG. This was placing the storage cluster into a warning state. With this release, repairing a PG does not place the storage cluster into a warning state.

([BZ#1506782](#))

### The **ceph-mgr** daemon no longer crashes after starting balancer module in automatic mode

Previously, due to a CRUSH bug, invalid mappings were created. When an invalid mapping was encountered in the **\_apply\_upmap** function, the code caused a segmentation fault. With this release, the code has been updated to check that the values are within an expected range. If not, the invalid values are ignored.

([BZ#1593110](#))

### RocksDB compaction no longer exhausts free space of BlueFS

Previously, the balancing of free space between main storage and storage for RocksDB, managed by BlueFS, happened only when write operations were underway. This caused an **ENOSPC** error for BlueFS to be returned when RocksDB compaction was triggered right before long interval without write operations. With this update, the code has been modified to periodically check free space balance even if no write operations are ongoing so that compaction no longer exhausts free space of BlueFS.

([BZ#1600138](#))

### PGs per OSD limits have been increased

In some situations, such as widely varying disk sizes, the default limit on placement groups (PGs) per OSD could prevent PGs from going active. These limits have been increased by default to make this situation less likely.

([BZ#1633426](#))

### Ceph installation no longer fails when FIPS mode is enabled

Previously, installing Red Hat Ceph Storage using the **ceph-ansible** utility failed at **TASK [ceph-mon : create monitor initial keyring]** when FIPS mode was enabled. To resolve this bug, the symmetric cipher cryptographic key is now wrapped with a one-shot wrapping key before it is used to instantiate the cipher. This allows Red Hat Ceph Storage to install normally when FIPS mode is enabled.

([BZ#1636251](#))

### Slow request messages have been re-added to the OSD logs

Previously, slow request messages were removed from the OSD logs, which made debugging harder. This update re-adds these warnings to the OSD logs.

([BZ#1659156](#))

### Force backfill and recovery preempt a lower priority backfill or recovery

Previously, force backfill or force recovery did not preempt an already running recovery or backfill process. As a consequence, although force backfill or recovery set priority to the max value, recovery process for placement groups (PGs) already running at a lower priority was finished first. With this update, force backfill and recovery preempt a lower priority backfill or recovery processes.

([BZ#1668362](#))



### **Ceph Manager no longer crashes when two or more Ceph Object Gateway daemons use the same name**

Previously, when two or more Ceph Object Gateway daemons used the same name in a cluster, Ceph Manager terminated unexpectedly. The underlying source code has been modified, and Ceph Manager no longer crashes in the described scenario.

([BZ#1670781](#), [BZ#1634964](#))

### **A race condition was causing threads to deadlock with the standby ceph-mgr daemon**

Some threads can cause a race condition when acquiring a local lock and the Python global interpreter lock, which is causing a deadlock issue for each thread. As the thread holds on to one of the locks, it wants to acquire the other lock, but cannot. In this release, the code was fixed to close the window of opportunity for the race condition to occur. This is done by changing the location of the lock acquisition and releasing the appropriate locks. Doing this results in the threads not causing a deadlock, which allows progress to be made.

([BZ#1674549](#))

### **An OSD daemon no longer crashes when a block device has read errors**

Previously, an OSD daemon would crash when a block device had read errors, because the daemon expected only a general EIO error code, not the more specific errors the kernel generates. With this release, low-level errors are mapped to EIO, resulting in an OSD daemon not crashing because of an unrecognized error code.

([BZ#1678470](#))

### **Read retries no longer cause the client to hang after a failed sync read**

Previously, when an OSD daemon failed to sync read an object, the length of the object to be read was set to 0. This caused the read retry to incorrectly read the entire object. The underlying code has been fixed, and the read retry uses the correct length and does not cause the client to hang.

([BZ#1682966](#))

## **4.10. BLOCK DEVICES (RBD)**

### **The *python-rbd list\_snaps()* method no longer segfaults after an error**

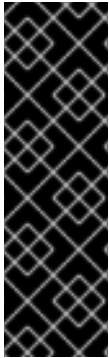
This issue was discovered with OpenStack Cinder Backup when **rados\_connect\_timeout** was set. Normally the timeout is not enabled. If the cluster was highly loaded the timeout could be reached, causing the segfault. With this update to Red Hat Ceph Storage, if the timeout is reached a segfault no longer occurs.

([BZ#1655681](#))



## 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/>.

## 5.1. BLOCK DEVICES (RBD)

### Erasure Coding for Ceph Block Devices

Erasure coding for Ceph Block Devices is supported as a Technology Preview. For details, see the [Erasure Coding with Overwrites \(Technology Preview\)](#) section in the Storage Strategies Guide for Red Hat Ceph Storage 3.

## 5.2. CEPH FILE SYSTEM

### Erasure Coding for Ceph File System

Erasure coding for Ceph File System is now supported as a Technology Preview. For details, see the [Creating Ceph File Systems with erasure coding](#) section in the Ceph File System Guide for Red Hat Ceph Storage 3.

## 5.3. OBJECT GATEWAY

### Improved interoperability with S3 and Swift by using a unified tenant namespace

This enhancement allows buckets to be moved between tenants. It also allows buckets to be renamed.

In Red Hat Ceph Storage 2 the **rgw\_keystone\_implicit\_tenants** option only applied to Swift. As of Red Hat Ceph Storage 3 this option applies to s3 also. Sites that used this feature with Red Hat Ceph Storage 2 now have outstanding data that depends on the old behavior. To accommodate that issue this enhancement also expands **rgw\_keystone\_implicit\_tenants** so it can be set to any of "none", "all", "s3", or "swift".

For more information, see Bucket management in the [Object Gateway Guide for Red Hat Enterprise Linux](#) or [Object Gateway Guide for Ubuntu](#) depending on your distribution. The **rgw\_keystone\_implicit\_tenants** setting is documented in the [Using Keystone to Authenticate Ceph Object Gateway Users](#) guide.

### AWS4 signature support in S3 authentication for Ceph Object Gateway when using Keystone

With this update, S3 user authentication using the new AWS4 signatures as a part of the Keystone service is supported as a Technology Preview.

The Ceph Object Gateway supports a subset of the Amazon Secure Token Service (STS) REST APIs. STS Lite is one supported API. It provides access to a set of temporary credentials for identity and access management. For more information, see [Authentication using the STS Lite API \(Technology Preview\)](#) in the [Developer Guide](#).

### **The Beast HTTP front end**

This update adds a new Ceph Object Gateway HTTP front end called Beast as a Technology Preview. The Beast front end uses the Boost.Beast library for HTTP parsing and the Boost.Asio library for asynchronous I/O.

### **Experimental support for delegated authorization using the Open Policy Agent (OPA)**

The Open Policy Agent is a distributed policy-based authorization framework being incubated in the Cloud-Native Computing Foundation (CNCF). This feature is in development and is not to be used in a production environment.

## CHAPTER 6. KNOWN ISSUES

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

### 6.1. THE CEPH-ANSIBLE UTILITY

The **shrink-osd.yml** playbook currently has no support for removing OSDs created by **ceph-volume**

The **shrink-osd.yml** playbook assumes all OSDs are created by the **ceph-disk** utility. Consequently, OSDs deployed by using the **ceph-volume** utility cannot be shrunk.

To work around this issue, remove OSDs deployed by using **ceph-volume** manually.

([BZ#1569413](#))

**Partitions are not removed from NVMe devices by shrink-osd.yml in certain situations**

The Ansible playbook **infrastructure-playbooks/shrink-osd.yml** does not properly remove partitions on NVMe devices when used with **osd\_scenario: non-collocated** in containerized environments.

To work around this issue, manually remove the partitions.

([BZ#1572933](#))

**When putting a dedicated journal on an NVMe device installation can fail**

When the **dedicated\_devices** setting contains an NVMe device and it has partitions or signatures on it Ansible installation might fail with an error like the following:

```
journal check: ondisk fsid 00000000-0000-0000-0000-000000000000 doesn't match expected
c325f439-6849-47ef-ac43-439d9909d391, invalid (someone else's?) journal
```

To work around this issue, ensure there are no partitions or signatures on the NVMe device.

([BZ#1619090](#))

**When deploying Ceph NFS Ganesha gateways on Ubuntu IPv6 systems ceph-ansible may fail to start the nfs-ganesha services**

This issue causes Ceph NFS Ganesha gateways to fail to deploy.

To work around this issue, rerun ceph-ansible playbook **site.yml** to deploy only the Ceph NFS Ganesha gateways:

```
[root@ansible ~]# ansible-playbook /usr/share/ceph-ansible/site.yml --limit nfss
```

([BZ#1656908](#))

**When using dedicated devices for BlueStore the default sizes for *block.db* and *block.wal* might be too small**

By default **ceph-ansible** does not override the default values **bluestore block db size** and **bluestore block wal size**. The default sizes are 1 GB and 576 MB respectively. These sizes might be too small when using dedicated devices with BlueStore.

To work around this issue, set **bluestore\_block\_db\_size** or **bluestore\_block\_wal\_size**, or both, using **ceph\_conf\_overrides** in **ceph.conf** to override the default values.

([BZ#1657883](#))

## 6.2. CEPH MANAGEMENT DASHBOARD

### Ceph OSD encryption summary is not displayed in the Red Hat Ceph Storage Dashboard

On the *Ceph OSD Information* dashboard, under the *OSD Summary* panel, the *OSD Encryption Summary* information is not displayed.

There is no workaround at this time.

([BZ#1605241](#))

### The Prometheus **node-exporter** service is not removed after purging the Dashboard

When purging the Red Hat Ceph Storage Dashboard, the **node-exporter** service is not removed, and is still running.

To work around this issue, manually stop and remove the **node-exporter** service.

Perform the following commands as **root**:

```
# systemctl stop prometheus-node-exporter
# systemctl disable prometheus-node-exporter
# rpm -e prometheus-node-exporter
# reboot
```

For Ceph Monitor, OSD, Object Gateway, MDS, and Dashboard, nodes, reboot these one at a time.

([BZ#1609713](#))

### The **OSD down** tab shows an incorrect value

When rebooting OSDs, the **OSD down** tab in the **CEPH Backend storage** dashboard shows the correct number of OSDs that are **down**. However, when all OSDs are **up** again after the reboot, the tab continues showing the number of **down** OSDs.

There is no workaround at this time.

([BZ#1652233](#))

### The **Top 5 pools by Throughput** graph lists all pools

The *Top 5 pools by Throughput* graph in the *Ceph Pools* tab lists all pools in the cluster instead of listing only the top five pools with the highest throughput.

There is no workaround at this time.

([BZ#1652807](#))

The **MDS Performance** dashboard displays the wrong value for **Clients** after increasing and decreasing the number of active MDS servers and clients multiple times.

This issue causes the Red Hat Ceph Storage dashboard to display the wrong number of CephFS clients. This can be verified by comparing the value in the Red Hat Ceph Storage dashboard with the value printed by the **ceph fs status \$FILESYSTEM\_NAME** command.

There is no workaround at this time.

([BZ#1652896](#))

### **Request Queue Length displays an incorrect value**

In the **Ceph RGW Workload** dashboard, the **Request Queue Length** parameter always displays **0** even when running Ceph Object Gateways I/Os from different clients.

There is no workaround at this time.

([BZ#1653725](#))

### **Capacity Utilization in Ceph - At Glance dashboard shows the wrong value when an OSD is down**

This issue causes the Red Hat Ceph Dashboard to show capacity utilization which is less than what **ceph df** shows.

There is no workarond at this time.

([BZ#1655589](#))

### **Some links on the Ceph - At Glance page do not work after installing ceph-metrics**

After installing **ceph-metrics**, some of the panel links on the *Ceph - At Glance* page in the Ceph Dashboard do not work.

To work around this issue, clear the browser cache and reload the Ceph Dashboard site.

([BZ#1655630](#))

### **The iSCSI Overview dashboard does not display graphs if the [iscsigws] role is included in the Ansible inventory file.**

When deploying the Red Hat Ceph Storage Dashboard, the *iSCSI Overview* dashboard does not display any graphs or values if the Ansible inventory file has the [iscsigws] role included for iSCSI gateways.

To work around this issue, add **[iscsis]** as a role in the Ansible inventory file and run the Ansible playbook for **cephmetrics-ansible**. The *iSCSI Overview* dashboard then displays the graphs and values.

([BZ#1656053](#))

### **In the Ceph Cluster dashboard the Pool Capacity graphs display values higher than actual capacity**

This issue causes the *Pool Capacity* graph to display values around one percent higher than what **df --cluster** shows.

There is no workaround at this time.

([BZ#1656820](#))

### **Graphs on the OSD Node Detail dashboard might appear incorrect when used with All**

Graphs generated under *OSD Node Detail > OSD Host Name > All* do not show all OSDs in the cluster. A graph with data for hundreds or thousands of OSDs would not be usable. The ability to set *All* is intended to show cluster-wide values. For some dashboards it does not make sense and should not be used.

There is no workaround at this time.

([BZ#1659036](#))

## 6.3. CEPH FILE SYSTEM

### The Ceph Metadata Server might crash during scrub with multiple MDS

This issue is triggered when the **scrub\_path** command is run in an environment with multiple Ceph Metadata Servers.

There is no workaround at this time.

([BZ#1642015](#))

## 6.4. THE CEPH-VOLUME UTILITY

### Deploying an OSD on devices with GPT headers fails

Drives with GPT headers will cause an error to be returned by LVM when deploying an OSD on them. The error says the device has been excluded by a filter.

To work around this issue ensure there is no GPT header present on devices to be used by OSDs.

([BZ#1644321](#))

## 6.5. ISCSI GATEWAY

### Using **ceph-ansible** to deploy the iSCSI gateway does not allow the user to adjust the **max\_data\_area\_mb** option

Using the **max\_data\_area\_mb** option with the **ceph-ansible** utility sets a default value of 8 MB. To adjust this value, set it manually using the **gwcli** command. See the Red Hat Ceph Storage [Block Device Guide](#) for details on setting the **max\_data\_area\_mb** option.

([BZ#1613826](#))

### Ansible fails to purge RBD images with snapshots

The **purge-iscsi-gateways.yml** Ansible playbook does not purge RBD images with snapshots. To purge the images and their snapshots, use the **rbd** command-line utility:

- To purge a snapshot:

```
rbd snap purge pool-name/image-name
```

For example:

```
# rbd snap purge data/image1
```

- To delete an image:

```
rdm rm image-name
```

For example:

```
# rdm rm image1
```

([BZ#1654346](#))

## 6.6. OBJECT GATEWAY

### Ceph Object Gateway garbage collection decreases client performance by up to 50% during mixed workload

In testing during a mixed workload of 60% read operations, 16% write operations, 14% delete operations, and 10% list operations, at 18 hours into the testing run, client throughput and bandwidth drop to half their earlier levels.

([BZ#1596401](#))

### Pushing a docker image to the Ceph Object Gateway over s3 does not complete

In certain situations when configuring **docker-distribution** to use Ceph Object Gateway with the s3 interface the **docker push** command does not complete. Instead the command fails with an HTTP 500 error.

There is no workaround at this time.

([BZ#1604979](#))

### Delete markers are not removed with a lifecycle configuration

In certain situations after deleting a file and a lifecycle triggers, delete markers are not removed.

There is no workaround at this time.

([BZ#1654820](#))

### The Ceph Object Gateway's S3 does not always work in FIPS mode

If a secret key of a Ceph Object Gateway user or sub-user is less than 112 bits in length, it can cause the **radosgw** daemon to exit unexpectedly when a user attempts to authenticate using S3.

This is because the FIPS mode Red Hat Enterprise Linux security policy forbids construction of a cryptographic HMAC based on a key of less than 112 bits, and violation of this constraint yields an exception that is not correctly handled in Ceph Object Gateway.

To work around this issue, ensure that the secret keys of Ceph Object Gateway users and sub-users are at least 112 bits in length.

([BZ#1687567](#))

## 6.7. RADOS

### Performing I/O in CephFS erasure-coded pools can cause a failure on assertion

This issue is being investigated as a possible latent bug in the messenger layer which could be causing out of order operations on the OSD.

The issue causes the following error:

```
FAILED assert(repop_queue.front() == repop)
```

There is no workaround at this time. CephFS with erasure-coded pools are a Technology Preview. For more information see [Creating Ceph File Systems with erasure coding](#) in the [Ceph File System Guide](#)

([BZ#1637948](#))



## CHAPTER 7. 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.

### 7.1. THE CEPH-ANSIBLE UTILITY

#### The `rgw_dns_name` parameter

The `rgw_dns_name` parameter is deprecated. Instead, configure the RADOS Gateway (RGW) zonegroup with the RGW DNS name. For more information, see: [Ceph - How to add hostnames in RGW zonegroup](#) in the Red Hat Customer Portal.

## CHAPTER 8. SOURCES

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

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