Red Hat Ceph Storage 7.1

Release Notes

Release notes for Red Hat Ceph Storage 7.1

Last Updated: 2024-06-13

Red Hat Ceph Storage 7.1 Release Notes

Release notes for Red Hat Ceph Storage 7.1

Legal Notice

Abstract

The release notes describes the major features, enhancements, known issues, and bug fixes implemented for the Red Hat Ceph Storage 7.1 product release.

Table of Contents

MAKING OPEN SOURCE MORE INCLUSIVE

Red Hat is committed to replacing problematic language in our code, documentation, and web properties. We are beginning with these four terms: master, slave, blacklist, and whitelist. Because of the enormity of this endeavor, these changes will be implemented gradually over several upcoming releases. For more details, see our CTO Chris Wright’s message .

PROVIDING FEEDBACK ON RED HAT CEPH STORAGE DOCUMENTATION

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

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

CHAPTER 1. INTRODUCTION

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

The Red Hat Ceph Storage documentation is available at https://access.redhat.com/documentation/en-us/red_hat_ceph_storage/7.

CHAPTER 2. ACKNOWLEDGMENTS

Red Hat Ceph Storage version 7.1 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 ®

CHAPTER 3. NEW FEATURES

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

3.1. THE CEPHADM UTILITY

Users can now configure various NFS options in **idmap.conf**

With this enhancement, the ability to configure NFS options, such as "Domain", "Nobody-User", "Nobody-Group" and the like, in **idmap.conf** is introduced.

Bugzilla:2068026

Client IP restriction is now possible over the new haproxy protocol mode for NFS

Previously, client IP restriction did not work in setups using haproxy over NFS.

With this enhancement, Cephadm deployed NFS supports the haproxy protocol. If users add **enable_haproxy_protocol: True** to both their ingress and haproxy specification or pass **--ingress-mode haproxy-protocol** to the **ceph nfs cluster create** command, the NFS daemon will make use of the haproxy protocol.

Bugzilla:2068030

Users must now enter a username and password to access the Grafana API URL

Previously, anyone who could connect to the Grafana API URL would have access to it without needing any credentials.

With this enhancement, Cephadm deployed Grafana is set up with a username and password for users to access the Grafana API URL.

Bugzilla:2079815

Ingress service with NFS backend can now be set up to use only **keepalived** to create a virtual IP (VIP) for the NFS daemon to bind to, without the HAProxy layer involved

With this enhancement, ingress service with an NFS backend can be set up to only use **keepalived** to create a virtual IP for the NFS daemon to bind to, without the HAProxy layer involved. This is useful in cases where the NFS daemon is moved around and clients need not use a different IP to connect to it.

Cephadm deploys **keepalived** to set up a VIP and then have the NFS daemon bind to that VIP. This can also be setup using the NFS module via the **ceph nfs cluster create** command, using the flags **--ingress --ingress-mode keepalive-only --virtual-ip <VIP>**.

The specification file looks as follows:

```
service_type: ingress
service_id: nfs.nfsganesha
service_name: ingress.nfs.nfsganesha
placement:
  count: 1
  label: foo
spec:
  backend_service: nfs.nfsganesha
  frontend_port: 12049
  monitor_port: 9049
  virtual_ip: 10.8.128.234/24
  virtual_interface_networks: 10.8.128.0/24
  keepalive_only: true
```

that includes the **keepalive_ony: true** setting.

An NFS specification looks as below:

```
networks:
    - 10.8.128.0/21
service_type: nfs
service_id: nfsganesha
placement:
  count: 1
  label: foo
spec:
  virtual_ip: 10.8.128.234
  port: 2049
```

that includes the **virtual_ip** field that should match the VIP in the ingress specification.

Bugzilla:2089167

The HAProxy daemon binds to its front-end port only on the VIP created by the accompanying keepalived

With this enhancement, the HAProxy daemon will bind to its front-end port only on the VIP created by the accompanying keepalived, rather than on 0.0.0.0. Cephadm deployed HAProxy will bind its front-end port to the VIP, allowing other services, such as an NFS daemon, to potentially bind to port 2049 on other IPs on the same node.

Bugzilla:2176297

Haproxy health check interval for ingress service is now customizable

Previously, in some cases, the two second default health check interval was too frequent and it caused unnecessary traffic.

With this enhancement, HAProxy health check interval for ingress service is customizable. By applying an ingress specification that includes the **health_check_interval** field, the HAProxy configuration generated by Cephadm for each HAProxy daemon for the service will include that value for the health check interval.

Ingress specification file:

```
service_type: ingress
service_id: rgw.my-rgw
placement:
  hosts: ['ceph-mobisht-7-1-07lum9-node2', 'ceph-mobisht-7-1-07lum9-node3']
spec:
  backend_service: rgw.my-rgw
  virtual_ip: 10.0.208.0/22
  frontend_port: 8000
  monitor_port: 1967
  health_check_interval: 3m
```

Valid units for the interval are: **us** : microseconds **ms** : milliseconds **s** : seconds **m** : minutes **h** : hours **d** : days

Bugzilla:2199129

Grafana now binds to an IP within a specific network on a host, rather that always binding to 0.0.0.0

With this enhancement, using a Grafana specification file that includes both the networks' section with the network that Grafana binds to an IP on, and **only_bind_port_on_networks: true** included in the "spec" section of the specification, Cephadm configures the Grafana daemon to bind to an IP within that network rather than 0.0.0.0. This enables users to use the same port that Grafana uses for another service but on a different IP on the host. If it is a specification update that does not cause them all to be moved, **ceph orch redeploy grafana** can be run to pick up the changes to the settings.

Grafana specification file:

```
service_type: grafana
service_name: grafana
placement:
  count: 1
networks:
  192.168.122.0/24
spec:
  anonymous_access: true
  protocol: https
  only_bind_port_on_networks: true
```

Bugzilla:2233659

All bootstrap CLI parameters are now made available for usage in the **cephadm-ansible** module

Previously, only a subset of the bootstrap CLI parameters were available and it was limiting the module usage.

With this enhancement, all bootstrap CLI parameters are made available for usage in the **cephadm-ansible** module.

Bugzilla:2246266

Prometheus scrape configuration is added to the nfs-ganesha exporter

With this enhancement, the Prometheus scrape configuration is added to the nfs-ganesha exporter. This is done to scrape the metrics exposed by nfs-ganesha prometheus exporter into the Prometheus instance running in Ceph, which would be further consumed by Grafana Dashboards.

Bugzilla:2263898

Prometheus now binds to an IP within a specific network on a host, rather that always binding to 0.0.0.0

With this enhancement, using a Prometheus specification file that includes both the networks section with the network that Prometheus binds to an IP on, and **only_bind_port_on_networks: true** included in the "spec" section of the specification, Cephadm configures the Prometheus daemon to bind to an IP within that network rather than 0.0.0.0. This enables users to use the same port that Prometheus uses for another service but on a different IP on the host. If it is a specification update that does not cause them all to be moved, **ceph orch redeploy prometheus** can be run to pick up the changes to the settings.

Prometheus specification file:

```
service_type: prometheus
service_name: prometheus
placement:
  count: 1
networks:
- 10.0.208.0/22
spec:
  only_bind_port_on_networks: true
```

Bugzilla:2264812

Users can now mount snapshots (exports within .snap directory)

With this enhancement, users can mount snapshots (exports within **.snap** directory) to look at in a RO mode. NFS exports created with the NFS MGR module now include the **cmount_path** setting (this cannot be configured and should be left as "/") which allows snapshots to be mounted.

Bugzilla:2245261

Zonegroup hostnames can now be set using the specification file provided in the **ceph rgw realm bootstrap…** command

With this release, in continuation to the automation of Ceph Object gateway multi-site setup, users can now set zonegroup hostnames through the initial specification file passed in the bootstrap command **ceph rgw realm bootstrap…** instead of requiring additional steps.

For example,

```
zonegroup_hostnames:
- host1
- host2
```

If users add the above section to the "specification" section of the Ceph Object gateway specification file passed in the realm bootstrap command, Cephadm will automatically add those hostnames to the zonegroup defined in the specification after the Ceph Object gateway module finishes creation of the realm/zonegroup/zone. Note that this may take a few minutes to occur depending on what other activity the Cephadm module is currently completing.

3.2. CEPH DASHBOARD

CephFS snapshot schedules management on the Ceph dashboard

Previously, CephFS snapshot schedules could only be managed through the command-line interface.

With this enhancement, CephFS snapshot schedules can be listed, created, edited, activated, deactivated, and removed from the Ceph dashboard.

Bugzilla:2264145

Ceph dashboard now supports NFSv3-based exports in Ceph dashboard

With this enhancement, support is enabled for NFSv3-based export management in the Ceph dashboard.

Bugzilla:2267763

Ability to manage Ceph users for CephFS is added

With this enhancement, the ability to manage the Ceph users for CephFS is added. This provides the ability to manage the users' permissions for volumes, subvolume groups, and subvolumes from the File System view.

Bugzilla:2271110

A new API endpoint for multi-site sync status is added

Previously, multi-site sync status was available only via the CLI command.

With this enhancement, multi-site status is added via an API in the Ceph dashboard. The new API endpoint for multi-site sync status is **api/rgw/multisite/sync_status**.

Bugzilla:2258951

Improved monitoring of NVMe-oF gateway

With this enhancement, to improve monitoring of NVMe-oF gateway, alerts of NVMe-oF gateway are added based on the metrics emitted and also, metrics from the embedded prometheus exporter are scraped in the NVMe-oF gateway.

Bugzilla:2276038

CephFS clone management in Ceph dashboard

With this enhancement, CephFS clone management functionality is provided in the Ceph dashboard. Users can create and delete subvolume clone through the Ceph dashboard.

Bugzilla:2264142

CephFS snapshot management in Ceph dashboard

With this enhancement, CephFS snapshot management functionality is provided in the Ceph dashboard. Users can create and delete subvolume snapshot through the Ceph dashboard.

Bugzilla:2264141

Labeled Performance Counters per user/bucket

With this enhancement, users can not only obtain information on the operations happening per Ceph Object Gateway node, but can also view the Ceph Object Gateway performance counters per-user and per-bucket in the Ceph dashboard.

Labeled Sync Performance Counters into Prometheus

With this enhancement, users can gather real-time information from Prometheus about the replication health between zones for increased observability of the Ceph Object Gateway multi-site sync operations.

Add and edit bucket in Ceph dashboard

With this enhancement, as part of the Ceph Object Gateway improvements to the Ceph dashboard, the capability to apply, list and edit Buckets from the Ceph dashboard is added.

- ACL(Public, Private)
- Tags(adding/removing)

Add, List, Delete, and Apply bucket policies in Ceph dashboard

With this enhancement, as part of the Ceph Object Gateway improvements to the Ceph dashboard, the capability to add, list, delete, and apply bucket policies from the Ceph dashboard is added.

3.3. CEPH FILE SYSTEM

MDS dynamic metadata balancer is off by default

Previously, poor balancer behavior would fragment trees in undesirable ways by increasing the **max_mds** file system setting.

With this enhancement, MDS dynamic metadata balancer is off, by default. Operators must turn on the balancer explicitly to use it.

Bugzilla:2227309

CephFS supports quiescing of subvolumes or directory trees

Previously, multiple clients would interleave reads and writes across a consistent snapshot barrier where out-of-band communication existed between clients. This communication led to clients wrongly believing they have reached a checkpoint that is mutually recoverable via a snapshot.

With this enhancement, CephFS supports quiescing of subvolumes or directory trees to enable the execution of crash-consistent snapshots. Clients are now forced to quiesce all I/O before the MDS executes the snapshot. This enforces a checkpoint across all clients of the subtree.

Bugzilla:2235753

MDS Resident Segment Size (RSS) performance counter is tracked with a higher priority

With this enhancement, the MDS Resident Segment Size performance counter is tracked with a higher priority to allow callers to consume its value to generate useful warnings. This allows Rook to identify the MDS RSS size and act accordingly.

Bugzilla:2256560

Laggy clients are now evicted only if there are no laggy OSDs

Previously, monitoring performance dumps from the MDS would sometimes show that the OSDs were laggy, **objecter.op_laggy** and **objecter.osd_laggy**, causing laggy clients (dirty data could not be flushed for cap revokes).

With this enhancement, if the **defer_client_eviction_on_laggy_osds** option is set to true and a client gets laggy because of a laggy OSD then client eviction will not take place until OSDs are no longer laggy.

Bugzilla:2260003

cephfs-mirror daemon exports snapshot synchronization performance counters via **perf dump** command

ceph-mds daemon export per-client performance counters included in the already existing **perf dump** command.

Bugzilla:2264177

A new **dump dir** command is introduced to dump the directory information

With this enhancement, the **dump dir** command is introduced to dump the directory information and print the output.

Bugzilla:2269687

Snapshot scheduling support for subvolumes

With this enhancement, snapshot scheduling support is provided for subvolumes. All snapshot scheduling commands accept **--subvol** and **--group** arguments to refer to appropriate subvolumes and subvolume groups. If a subvolume is specified without a subvolume group argument, then the default subvolume group is considered. Also, a valid path need not be specified when referring to subvolumes and just a placeholder string is sufficient due to the nature of argument parsing employed.

Example

```
# ceph fs snap-schedule add - 15m --subvol sv1 --group g1
# ceph fs snap-schedule status - --subvol sv1 --group g1
```

Bugzilla:2238537

Ceph commands that add or modify MDS caps give an explanation about why the MDS caps passed by user was rejected

Previously, Ceph commands that add or modify MDS caps printed "Error EINVAL: mds capability parse failed, stopped at 'allow w' of 'allow w'".

With this enhancement, the commands give an explanation about why the MDS caps passed by user were rejected and print Error EINVAL: Permission flags in MDS caps must start with 'r' or 'rw' or be '*' or 'all'.

Bugzilla:2247586

3.4. CEPH OBJECT GATEWAY

Admin interface is now added to manage bucket notification

Previously, the S3 REST APIs were used to manage bucket notifications. However, if an admin wanted to override them, there was no easy way to do that over the radosgw-admin tool.

With this enhancement, an admin interface with the following commands is added to manage bucket notifications:

```
radosgw-admin notification get --bucket <bucket name> --notification-id <notification id>

radosgw-admin notification list --bucket <bucket name>

radosgw-admin notification rm --bucket <bucket name> [--notification-id <notification id>]
```

Bugzilla:2130292

RGW labeled user and bucket operation counters are now in different sections when the **ceph counter dump** is run

Previously, all RGW labeled operation counters were in the **rgw_op`** section of the output of the **ceph counter dump** command but would either have a user label or a bucket label.

With this enhancement, RGW labeled user and bucket operation counters are in **rgw_op_per_user** or **rgw_op_per_bucket** sections respectively when the **ceph counter dump** command is executed.

Bugzilla:2265574

Users can now place temporary files into a directory using the -t command-line option

Previously, the **/usr/bin/rgw-restore-bucket-index** tool just used **/tmp** and that directory sometimes did not have enough free space to hold all the temporary files.

With this enhancement, the user can specify a directory into which the temporary files can be placed using the **-t** command-line option and will be notified if they run out of space, thereby knowing what adjustments to make to re-run the tool. Also, users can periodically check if the tool's temporary files have exhausted the available space on the file system where the temporary files are present.

Bugzilla:2267715

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

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

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

Bugzilla:2149450

A new Ceph Object Gateway admin-ops capability is added to allow reading user metadata but not their associated authorization keys

With this enhancement, a new Ceph Object Gateway admin-ops capability is added to allow reading Ceph Object gateway user metadata but not their associated authorization keys. This is to reduce the privileges of automation and reporting tools and to avoid impersonating users or view their keys.

Bugzilla:2112325

Cloud Transition: add new supported S3-compatible platforms

With this release, to be able to move object storage to the cloud or other on-premise S3 endpoints, the current lifecycle transition and storage class model is extended. S3-compatible platforms, such as IBM Cloud Object Store (COS) and IBM Storage Ceph are now supported for the cloud archival feature.

NFS with RGW backend

With this release, NFS with Ceph Object Gateway backend is re-GAed with the existing functionalities.

3.5. MULTI-SITE CEPH OBJECT GATEWAY

A retry mechanism is introduced in the radosgw-admin sync status command

Previously, when the multisite sync sent requests to a remote zone, it used a round robin strategy to choose one of its zone endpoints. If that endpoint was not available, the http client logic used by the **radosgw-admin sync status** command would not provide a retry mechanism, and thus report input/output error.

With this enhancement, a retry mechanism is introduced in the sync status command by virtue of which, if the chosen endpoint is unavailable, a different endpoint is selected to serve the request.

Bugzilla:1995152

NewerNoncurrentVersions, **ObjectSizeGreaterThan**, and **ObjectSizeLessThan** filters are added to the lifecycle

With this enhancement, support for **NewerNoncurrentVersions**, **ObjectSizeGreaterThan**, and **ObjectSizeLessThan** filters are added to the lifecycle.

Bugzilla:2172162

User S3 replication APIs are now supported

With this enhancement, user S3 replication APIs are now supported. With these APIs, users can set replication policies at bucket-level. The API is extended to include additional parameters to specify source and destination zone names.

Bugzilla:2279461

Bucket Granular Sync Replication GA (Part 3)

With this release, the ability to replicate a bucket or a group of buckets to a different Red Hat Ceph Storage cluster is added with bucket granular support. The usability requirements are as Ceph Object Gateway multi-site.

3.6. RADOS

Setting the **noautoscale** flag on/off retains each pool's original autoscale mode configuration

Previously, the **pg_autoscaler** did not persist in each pool's **autoscale mode** configuration when the **noautoscale** flag was set. Due to this, whenever the **noautoscale** flag was set, the **autoscale** mode had to be set for each pool repeatedly.

With this enhancement, the **pg_autoscaler** module persists individual pool configuration for the autoscaler mode after the **noautoscale flag** is set. Setting the **noautoscale** flag on/off still retains each pool's original autoscale mode configuration.

Bugzilla:2136766

reset_purged_snaps_last OSD command is introduced

With this enhancement, **reset_purged_snaps_last** OSD command is introduced to resolve cases in which the **purged_snaps** keys (PSN) are missing in the OSD and exist in the monitor. The **purged_snaps_last** command will be zeroed and as a result, the monitor will share all its **purged_snaps** information with the OSD on the next boot.

Bugzilla:2251188

BlueStore's RocksDB compression enabled

With this enhancement, to ensure that the metadata (especially OMAP) takes less space, RocksDB configuration is modified to enable internal compression of its data.

As a result, * database size is smaller * write amplification during compaction is smaller * average I/O is higher * CPU usage is higher

Bugzilla:2253313

OSD is now more resilient to fatal corruption

Previously, special OSD layer object "superblock" would be overwritten due to being located at the beginning of the disk, resulting in a fatal corruption.

With this enhancement, OSD "superblock" is redundant and is migrating on disk. Its copy is stored in the database. OSD is now more resilient to fatal corruption.

Bugzilla:2079897

3.7. RADOS BLOCK DEVICES (RBD)

Improved rbd_diff_iterate2() API performance

Previously, RBD diff-iterate was not guaranteed to execute locally if exclusive lock was available when diffing against the beginning of time (**fromsnapname == NULL**) in fast-diff mode (**whole_object == true** with **fast-diff** image feature enabled and valid).

With this enhancement, **rbd_diff_iterate2()** API performance is improved, thereby increasing the performance for QEMU live disk synchronization and backup use cases, where the **fast-diff** image feature is enabled.

Bugzilla:2258997

CHAPTER 4. BUG FIXES

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

4.1. THE CEPHADM UTILITY

Using the --name NODE flag with the cephadm shell to start a stopped OSD no longer returns the wrong image container

Previously, in some cases, when using the **cephadm shell --name NODE** command, the command would start the container with the wrong version of the tools. This would occur when a user has a newer ceph container image on the host than the one that their OSDs are using.

With this fix, Cephadm determines the container image for stopped daemons when using the cephadm shell command with the **--name** flag. Users no longer have any issues with the **--name** flag, and the command works as expected.

Bugzilla:2258542

4.2. THE CEPH ANSIBLE UTILITY

Playbooks now remove the RHCS version repositories matching the running RHEL version

Previously, playbooks would try to remove Red Hat Ceph Storage 4 repositories from RHEL 9 even though they do not exist on RHEL 9. This would cause the playbooks to fail.

With this fix, playbooks remove existing Red Hat Ceph Storage version repositories matching the running RHEL version and the correct repositories are removed.

Bugzilla:2258940

4.3. NFS GANESHA

All memory consumed by the configuration reload process is now released

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

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

Bugzilla:2265322

4.4. CEPH DASHBOARD

Users can create volumes with multiple hosts in the Ceph dashboard

With this fix, users can now create volumes with multiple hosts in the Ceph dashboard.

Bugzilla:2241056

Unset subvolume size is no longer set as 'infinite'

Previously, the unset subvolume size was set to 'infinite', resulting in the failure of the update.

With this fix, the code that sets the size to 'infinite' is removed and the update works as expected.

Bugzilla:2251192

Missing options are added in the kernel mount command

Previously, a few options were missing in the kernel mount command for attaching the filesystem causing the command to not work as intended.

With this fix, the missing options are added and the kernel mount command works as expected.

Bugzilla:2266256

Ceph dashboard now supports both NFS v3 and v4-enabled export management

Previously, the Ceph dashboard only supported the NFSv4-enabled exports management and not the NFSv3-enabled exports. Due to this, any management done for exports via CLI for NFSv3 was corrupted.

With this fix, support for NFSv3-based exports management is enabled by having an additional checkbox. The Ceph dashboard now supports both v3 and v4-enabled export management.

Bugzilla:2267814

Access/secret keys are now not compulsory while creating a zone

Previously, access/secret keys were compulsory when creating a zone in Ceph Object Gateway multi-site. Due to this, users had to first set the non-system user’s keys in the zone and later update with the system user’s keys.

With this fix, access/secret keys are not compulsory while creating a zone.

Bugzilla:2275463

Importing multi-site configuration no longer throws an error on submitting the form

Previously, the multi-site period information did not contain the 'realm' name. Due to this, importing the multi-site configuration threw an error on submitting the form.

With this fix, the check for fetching 'realm' name from period information is removed and the token import works as expected.

Bugzilla:2275861

The Ceph Object Gateway metrics label names are aligned with the Prometheus label naming format and they are now visible in Prometheus

Previously, the metrics label names were not aligned with the Prometheus label naming format, causing the Ceph Object Gateway metrics to not be visible in Prometheus.

With this fix, the hyphen (-) is replaced with an underscore (_) in Ceph Object Gateway metrics label names, wherever applicable and all Ceph Object Gateway metrics are now visible in Prometheus.

Bugzilla:2276340

Full names can now include dot in Ceph dashboard

Previously, in the Ceph dashboard, it was not possible to create or modify a full name with a dot in it due to incorrect validation.

With this fix, validation is properly adapted to include a dot in full names in Ceph dashboard.

Bugzilla:2249812

4.5. CEPH FILE SYSTEM

MDS metadata with FSMap changes are now added in batches to ensure consistency

Previously, monitors would sometimes lose track of MDS metadata during upgrades and cancelled PAXOS transactions resulting in MDS metadata being no longer available.

With this fix, MDS metadata with FSMap changes are added in batches to ensure consistency. The **ceph mds metadata** command now functions as intended across upgrades.

Bugzilla:2144472

The **ENOTEMPTY** output is detected and the message is displayed correctly

Previously, when running the **subvolume group rm** command, the **ENOTEMPTY** output was not detected in the volume's plugin causing a generalized error message instead of a specific message.

With this fix, the **ENOTEMPTY** output is detected for the **subvolume group rm** command when there is subvolume present inside the subvolumegroup and the message is displayed correctly.

Bugzilla:2240138

MDS now queues the next client replay request automatically as part of request cleanup

Previously, sometimes, MDS would not queue the next client request for replay in the **up:client-replay** state causing the MDS to hang.

With this fix, the next client replay request is queued automatically as part of request cleanup and MDS proceeds with failover recovery normally.

Bugzilla:2243105

cephfs-mirroring overall performance is improved

With this fix, the incremental snapshot sync is corrected, which improves the overall performance of cephfs-mirroring.

Bugzilla:2248639

The loner member is set to **true**

Previously, for a file lock in the LOCK_EXCL_XSYN state, the non-loner clients would be issued empty caps. However, since the loner of this state is set to **false**, it could make the locker to issue the Fcb caps to them, which is incorrect. This would cause some client requests to incorrectly revoke some caps and infinitely wait and cause slow requests.

With this fix, the loner member is set to **true** and as a result the corresponding request is not blocked.

Bugzilla:2251258

snap-schedule repeat and retention specification for monthly snapshots is changed from **m** to **M**

Previously, the snap-schedule repeat specification and retention specification for monthly snapshots was not consistent with other Ceph components.

With this fix, the specifications are changed from **m** to **M** and it is now consistent with other Ceph components. For example, to retain 5 monthly snapshots, you need to issue the following command:

```
# ceph fs snap-schedule retention add /some/path M 5 --fs cephfs
```

Bugzilla:2264348

ceph-mds no longer crashes when some inodes are replicated in multi-mds cluster

Previously, due to incorrect lock assertion in ceph-mds, ceph-mds would crash when some inodes were replicated in a multi-mds cluster.

With this fix, the lock state in the assertion is validated and no crash is observed.

Bugzilla:2265415

Missing fields, such as **date**, **client_count**, **filters** are added to the **--dump** output

With this fix, missing fields, such as **date**, **client_count**, **filters** are added to the **--dump** output.

Bugzilla:2272468

MDS no longer fails with the assert function during recovery

Previously, MDS would sometimes report metadata damage incorrectly when recovering a failed rank and thus, fail with an assert function.

With this fix, the startup procedure is corrected and the MDS does not fail with the assert function during recovery.

Bugzilla:2272979

The target mon_host details are removed from the peer List and mirror daemon status

Previously, the snapshot mirror peer-list showed more information than just the peer list. This output caused confusion if there should be only one MON IP or all the MON host IP’s should be displayed.

With this fix, mon_host is removed from the fs snapshot mirror peer_list command and the target mon_host details are removed from the peer List and mirror daemon status.

Bugzilla:2277143

The target mon_host details are removed from the peer List and mirror daemon status

Previously, a regression was introduced by the quiesce protocol code. When killing the client requests, it would just skip choosing the new batch head for the batch operations. This caused the stale batch head requests to stay in the MDS cache forever and then be treated as slow requests.

With this fix, choose a new batch head when killing requests and no slow requests are caused by the batch operations.

Bugzilla:2277944

File system upgrade happens even when no MDS is up

Previously, monitors would not allow an MDS to upgrade a file system when all MDS were down. Due to this, upgrades would fail when the **fail_fs** setting was set to 'true'.

With this fix, monitors allow the upgrades to happen when no MDS is up.

Bugzilla:2244417

4.6. CEPH OBJECT GATEWAY

Auto-generated internal topics are no longer shown in the admin topic list command

Previously, auto-generated internal topics were exposed to the user via the topic list command due to which the users could see a lot more topics than what they had created.

With this fix, internal, auto-generated topics are not shown in the admin topic list command and users now see only the expected list of topics.

Bugzilla:1954461

The deprecated bucket name field is no longer shown in the topic list command

Previously, in case of pull mode notifications (**pubsub**), the notifications were stored in a bucket. However, despite this mode being deprecated, an empty bucket name field is still shown in the topic list command.

With this fix, the empty bucket name field is removed.

Bugzilla:1954463

Notifications are now sent on lifecycle transition

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

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

Bugzilla:2166576

RGWCopyObjRequest is fixed and rename operations work as expected

Previously, incorrect initialization of **RGWCopyObjRequest**, after zipper conversion, broke the rename operation. Due to this, many **rgw_rename()** scenarios failed to copy the source object, and due to a secondary issue, also deleted the source even though the copy had failed.

With this fix, **RGWCopyObjRequest** is corrected and several unit test cases are added for different renaming operations.

Bugzilla:2217499

Ceph Object Gateway can no longer be illegally accessed

Previously, a variable representing a Ceph Object Gateway role was being accessed before it was initialized, resulting in a segfault.

With this fix, operations are reordered and there is no illegal access. The roles are enforced as required.

Bugzilla:2252048

An error message is now shown per wrong CSV object structure

Previously, a CSV file with unclosed double-quotes would cause an assert, followed by a crash.

With this fix, an error message is introduced which pops up per wrong CSV object structure.

Bugzilla:2252396

Users no longer encounter 'user not found' error when querying user-related information in the Ceph dashboard

Previously, in the Ceph dashboard, end users could not retrieve the user-related information from the Ceph Object Gateway due to the presence of a namespace in the full **user_id** which the dashboard would not identify, resulting in encountering the “user not found” error.

With this fix, a fully constructed user ID, which includes **tenant**, **namespace**, and **user_id** is returned as well as each field is returned individually when a GET request is sent to admin ops for fetching user information. End users can now retrieve the correct **user_id**, which can be used to further fetch other user-related information from Ceph Object Gateway.

Bugzilla:2255255

Ceph Object gateway now passes requests with well-formed payloads of the new stream encoding forms

Previously, Ceph Object gateway would not recognize **STREAMING-AWS4-HMAC-SHA256-PAYLOAD** and **STREAMING-UNSIGNED-PAYLOAD-TRAILER** encoding forms resulting in request failures.

With this fix, the logic to recognize, parse, and wherever applicable, verify new trailing request signatures provided for the new encoding forms is implemented. The Ceph Object gateway now passes requests with well-formed payloads of the new stream encoding forms.

Bugzilla:2256967

The check stat calculation for radosgw admin bucket and bucket reshard stat calculation are now correct

Previously, due to a code change, radosgw-admin bucket check stat calculation and bucket reshard stat calculation were incorrect when there were objects that transitioned from unversioned to versioned.

With this fix, the calculations are corrected and incorrect bucket stat outputs are no longer generated.

Bugzilla:2257978

Tail objects are no longer lost during a multipart upload failure

Previously, during a multipart upload, if an upload of a part failed due to scenarios, such as a time-out, and the upload was restarted, the cleaning up of the first attempt would remove tail objects from the subsequent attempt. Due to this, the resulting Ceph Object Gateway multipart object would be damaged as some tail objects would be missing. It would respond to a HEAD request but fail during a GET request.

With this fix, the code cleans up the first attempt correctly. The resulting Ceph Object Gateway multipart object is no longer damaged and can be read by clients.

Bugzilla:2262650

ETag values in the **CompleteMultipartUpload** and its notifications are now present

Previously, changes related to notifications caused the object handle corresponding to the completing multipart upload to not contain the resulting ETag. Due to this, ETags were not present for completing multipart uploads as the result of **CompleteMultipartUpload** and its notifications. (The correct ETag was computed and stored, so subsequent operations contained a correct ETag result.)

With this fix, **CompleteMultipartUpload** refreshes the object and also prints it as expected. ETag values in the **CompleteMultipartUpload** and its notifications are present.

Bugzilla:2266579

Listing a container (bucket) via swift no longer causes a Ceph Object Gateway crash

Previously, a **swift-object-storage** call path was missing a call to update an object handle with its corresponding bucket (zipper backport issue). Due to this, listing a container (bucket) via swift would cause a Ceph Object Gateway crash when an S3 website was configured for the same bucket.

With this fix, the required zipper logic is added and the crash no longer occurs.

Bugzilla:2269038

Processing a lifecycle on a bucket with no lifecycle policy does not crash now

Previously, attempting to manually process a lifecycle on a bucket with no lifecycle policy induced a null pointer reference causing the radosgw-admin program to crash.

With this fix, a check for a null bucket handle is made before operating on the handle to avoid the crash.

Bugzilla:2270402

Zone details for a datapool can now be modified

The **rgw::zone_create()** function initializes the default placement target and pool name on zone creation. This function was also previously used for radosgw-admin zone set with **exclusive=false**. But, **zone set** does not allow the STANDARD storage class’s data_pool to be modified.

With this fix, the default-placement target should not be overwritten if it already exists and the zone details for a datapool can be modified as expected.

Bugzilla:2254480

Modulo operation on float numbers now return correct results

Previously, modulo operation on float numbers returned wrong results.

With this fix, the SQL engine is enhanced to handle modulo operations on floats and return correct results.

Bugzilla:2254125

SQL statements correctly return results for case-insensitive boolean expressions

Previously, SQL statements contained a boolean expression with capital letters in parts of the statement resulting in wrong interpretation and wrong results.

With this fix, the interpretation of a statement is case-insensitive and hence, the correct results are returned for any case.

Bugzilla:2254122

SQL engine returns the correct NULL value

Previously, SQL statements contained cast into type from NULL, as a result of which, the wrong result was returned instead of returning NULL.

With this fix, the SQL engine identifies cast from NULL and returns NULL.

Bugzilla:2254121

ETags values are now present in **CompleteMultipartUpload** and its notifications

Previously, the changes related to notifications caused the object handle, corresponding to the completing multipart upload, to not contain the resulting ETag. As a result, ETags were not present for **CompleteMultipartUpload** and its notifications. (The correct ETag was computed and stored, so subsequent operations contained a correct ETag result.)

With this fix, **CompleteMultipartUpload** refreshes the object and also prints it as expected. ETag values are now present in the **CompleteMultipartUpload** and its notifications.

Bugzilla:2249744

Sending workloads with embedded backslash (/) in object names to cloud-sync no longer causes sync failures

Previously, incorrect URL-escaping of object paths during cloud sync caused sync failures when workloads contained objects with an embedded backslash (/) in the names, that is, when virtual directory paths were used.

With this fix, incorrect escaping is corrected and workloads with embedded backslash (/) in object names can be sent to cloud-sync as expected.

Bugzilla:2249068

SQL statements containing boolean expression return boolean types

Previously, SQL statements containing boolean expression (a projection) would return a string type instead of boolean type.

With this fix, the engine identifies a string as a boolean expression, according to the statement syntax, and the engine successfully returns a boolean type (true/false).

Bugzilla:2254582

The work scheduler now takes the next date into account in the **should_work** function

Previously, the logic used in the **should_work** function, that decides whether the lifecycle should start running at the current time, would not take the next date notion into account. As a result, any custom work time "XY:TW-AB:CD" would break the lifecycle processing when AB < XY.

With this fix, the work scheduler now takes the next date into account and the various custom lifecycle work schedules now function as expected.

Bugzilla:2255938

merge_and_store_attrs() method no longer causes attribute update operations to fail

Previously, a bug in the **merge_and_store_attrs()** method, which deals with reconciling changed and the unchanged bucket instance attributes, caused some attribute update operations to fail silently. Due to this, some metadata operations on a subset of buckets would fail. For example, a bucket owner change would fail on a bucket with a rate limit set.

With this fix, the **merge_and_store_attrs()** method is fixed and all affected scenarios now work correctly.

Bugzilla:2262919

Checksum and malformed trailers can no longer induce a crash

Previously, an exception from **AWSv4ComplMulti** during **java AWS4Test.testMultipartUploadWithPauseAWS4** led to a crash induced by some client input, specifically, by those which use checksum trailers.

With this fix, an exception handler is implemented in **do_aws4_auth_completion()**. Checksum and malformed trailers can no longer induce a crash.

Bugzilla:2266092

Implementation of improved trailing chunk boundary detection

Previously, one valid-form of 0-length trailing chunk boundary formatting was not handled. Due to this, the Ceph Object Gateway failed to correctly recognize the start of the trailing chunk, leading to the 403 error.

With this fix, improved trailing chunk boundary detection is implemented and the unexpected 403 error in the anonymous access case no longer occurs.

Bugzilla:2266411

Default values for Kafka message and idle timeouts no longer cause hangs

Previously, the default values for Kafka message and idle timeouts caused infrequent hangs while waiting for the Kafka broker.

With this fix, the timeouts are adjusted and it no longer hangs.

Bugzilla:2269381

Delete bucket tagging no longer fails

Previously, an incorrect logic in RADOS SAL **merge_and_store_attrs()** caused deleted attributes to not materialize. This also affected **DeleteLifecycle**. As a result, a pure attribute delete did not take effect in some code paths.

With this fix, the logic to store bucket tags uses RADOS SAL **put_info()** instead of **merge_and_store_attrs()**. Delete bucket tagging now succeeds as expected.

Bugzilla:2271806

Object mtime now advances on S3 PutACL and ACL changes replicate properly

Previously, **S3 PutACL** operations would not update object **mtime**. Due to this, the ACL changes once applied would not replicate as the timestamp-based object-change check incorrectly returned false.

With this fix, the object **mtime** always advances on **S3 PutACL** and ACL changes properly replicate.

Bugzilla:2271938

All transition cases can now dispatch notifications

Previously, the logic to dispatch notifications on transition was mistakenly scoped to the cloud-transition case due to which notifications on pool transition were not sent.

With this fix, notification dispatch is added to the pool transition scope and all transition cases can dispatch notifications.

Bugzilla:2279607

RetainUntilDate after the year 2106 no longer truncates and works as expected for new **PutObjectRetention** requests

Previously, **PutObjectRetention** requests specifying a **RetainUntilDate** after the year 2106 would truncate, resulting in an earlier date used for object lock enforcement. This did not affect ` PutBucketObjectLockConfiguration` requests, where the duration is specified in days.

With this fix, the **RetainUntilDate** now saves and works as expected for new **PutObjectRetention** requests. Requests previously existing are not automatically repaired. To fix existing requests, identify the requests by using the **HeadObject** request based on the **x-amz-object-lock-retain-until-date** and save again with the **RetainUntilDate**.

For more information, see *S3 put object retention*

Bugzilla:2265890

Bucket lifecycle processing rules are no longer stalled

Previously, enumeration of per-shard bucket-lifecycle rules contained a logical error related to concurrent removal of lifecycle rules for a bucket. Due to this, a shard could enter a state which would stall processing of that shard, causing some bucket lifecycle rules to not be processed.

With this fix, enumeration can now skip past a removed entry and the lifecycle processing stalls related to this issue are resolved.

Bugzilla:2270334

Deleting objects in versioned buckets causes statistics mismatch

Due to versioned buckets having a mix of current and non-current objects, deleting objects might cause bucket and user statistics discrepancies on local and remote sites. This does not cause object leaks on either site, just statistics mismatch.

Bugzilla:1871333

4.7. MULTI-SITE CEPH OBJECT GATEWAY

Ceph Object Gateway no longer deadlocks during object deletion

Previously, during object deletion, the Ceph Object Gateway S3 **DeleteObjects** would run together with a multi-site deployment, causing the Ceph Object Gateway to deadlock and stop accepting new requests. This was caused by the **DeleteObjects** requests processing several object deletions at a time.

With this fix, the replication logs are serialized and the deadlock is prevented.

Bugzilla:2249651

CURL path normalization is now disabled at startup

Previously, due to "path normalization" performed by CURL, by default (part of the Ceph Object Gateway replication stack), object names were illegally reformatted during replication. Due to this, objects whose names contained embedded **.** and **..** were not replicated.

With this fix, the CURL path normalization is disabled at startup and the affected objects replicate as expected.

Bugzilla:2265148

The authentication of the forwarded request on the primary site no longer fails

Previously, an S3 request issued to secondary failed if temporary credentials returned by STS were used to sign the request. The failure occured because the request would be forwarded to the primary and signed using a system user's credentials which do not match the temporary credentials in the session token of the forwarded request. As a result of unmatched credentials, the authentication of the forwarded request on the primary site fails, which results in the failure of the S3 operation.

With this fix, the authentication is by-passed by using temporary credentials in the session token in case a request is forwarded from secondary to primary. The system user's credentials are used to complete the authentication successfully.

Bugzilla:2271399

4.8. RADOS

Ceph reports a POOL_APP_NOT_ENABLED warning if the pool has zero objects stored in it

Previously, Ceph status failed to report pool application warning if the pool was empty resulting in RGW bucket creation failure if the application tag was enabled for RGW pools.

With this fix, Ceph reports a **POOL_APP_NOT_ENABLED** warning even if the pool has zero objects stored in it.

Bugzilla:2029585

Checks are added for uneven OSD weights between two sites in a stretch cluster

Previously, there were no checks for equal OSD weights after stretch cluster deployment. Due to this, users could make OSD weights unequal.

With this fix, checks are added for uneven OSD weights between two sites in a stretch cluster. The cluster now gives a warning about uneven OSD weight between two sites.

Bugzilla:2125107

Autoscaler no longer runs while the norecover flag is set

Previously, the autoscaler would run while the **norecover** flag was set leading to creation of new PGs and these PGs requiring to be backfilled. Running of autoscaler while the **norecover** flag is set allowed in cases where I/O is blocked on missing or degraded objects in order to avoid client I/O hanging indefinitely.

With this fix, the autoscaler does not run while the **norecover** flag is set.

Bugzilla:2134786

The ceph config dump command output is now consistent

Previously, the **ceph config dump** command without the pretty print formatted output showed the localized option name and its value. An example of a normalized vs localized option is shown below:

> Normalized: mgr/dashboard/ssl_server_port
>
> Localized: mgr/dashboard/x/ssl_server_port

However, the pretty-printed (for example, JSON) version of the command only showed the normalized option name as shown in the example above. The **ceph config dump** command result was inconsistent between with and without the pretty-print option.

With this fix, the output is consistent and always shows the localized option name when using the **ceph config dump --format TYPE** command, with **TYPE** as the pretty-print type.

Bugzilla:2213766

MGR module no longer takes up one CPU core every minute and CPU usage is normal

Previously, expensive calls from the placement group auto-scaler module to get OSDMap from the Monitor resulted in the MGR module taking up one CPU core every minute. Due to this, the CPU usage was high in the MGR daemon.

With this fix, the number of OSD map calls made from the placement group auto-scaler module is reduced. The CPU usage is now normal.

Bugzilla:2241030

The correct CRUSH location of the OSDs parent (host) is determined

Previously, when the **osd_memory_target_autotune** option was enabled, the memory target was applied at the host level. This was done by using a host mask when auto-tuning the memory. But the code that applied to the memory target would not determine the correct CRUSH location of the parent host for the change to be propagated to the OSD(s) of the host. As a result, none of the OSDs hosted by the machine got notified by the config observer and the **osd_memory_target** remained unchanged for those set of OSDs.

With this fix, the correct CRUSH location of the OSDs parent (host) is determined based on the host mask. This allows the change to propagate to the OSDs on the host. All the OSDs hosted by the machine are notified whenever the auto-tuner applies a new **osd_memory_target** and the change is reflected.

Bugzilla:2244604

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

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

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

Bugzilla:2248939

'Error getting attr on' message no longer occurs

Previously, **ceph-objectstore-tool** listed pgmeta objects when using **--op list** resulting in "Error getting attr on" message.

With this fix, pgmeta objects are skipped and the error message no longer appears.

Bugzilla:2251004

LBA alignment in the allocators are no longer used and the OSD daemon does not assert due to allocation failure

Previously, OSD daemons would assert and fail to restart which could sometimes lead to data unavailability or data loss. This would happen as the OSD daemon would not assert if the allocator got to 4000 requests and configured with a different allocation unit.

With this fix, the LBA alignment in the allocators are not used and the OSD daemon does not assert due to allocation failure.

Bugzilla:2260306

A sqlite database using the "libcephsqlite" library no longer may be corrupted due to short reads failing to correctly zero memory pages.

Previously, “libcephsqlite” would not handle short reads correctly which may cause corruption of sqlite databases.

With this fix, “libcephsqlite” zeros pages correctly for short reads to avoid potential corruption.

Bugzilla:2240139

4.9. RBD MIRRORING

The image status description now shows "orphan (force promoting)" when a peer site is down during force promotion

Previously, upon a force promotion, when a peer site went down, the image status description showed "local image linked to unknown peer", which is not a clear description.

With this fix, the mirror daemon is improved to show image status description as "orphan (force promoting)".

Bugzilla:2190366

rbd_support module no longer fails to recover from repeated block-listing of its client

Previously, it was observed that the **rbd_support** module failed to recover from repeated block-listing of its client due to a recursive deadlock in the rbd_support module, a race condition in the **rbd_support** module’s librbd client, and a bug in the librbd cython bindings that sometimes crashed the ceph-mgr.

With this release, all these 3 issues are fixed and rbd_support` module no longer fails to recover from repeated block-listing of its client

Bugzilla:2247531

CHAPTER 5. KNOWN ISSUES

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

5.1. THE CEPHADM UTILITY

Cephadm does not maintain the previous OSD weight when draining an OSD

Cephadm does not maintain the previous OSD weight when draining an OSD. Due to this, if the **ceph orch osd rm <osd-id>** command is run and later, the OSD removal is stopped, Cephadm will not set the crush weight of the OSD back to its original value. The crush weight will remain at 0.

As a workaround, users have to manually adjust the crush weight of the OSD to its original value, or complete removal of the OSD and deploy a new one. Users should be careful when cancelling a **ceph orch osd rm** operation, as the crush weight of the OSD will not be returned to its original value before the removal process begins.

Bugzilla:2247211

Repeated use of the Ceph Object Gateway realm bootstrap command causes setting the zonegroup hostname to fail

Setting the zonegroup hostnames using the Ceph Object Gateway realm bootstrap command fails when being done multiple times. Due to this, the repeated use of the Ceph Object Gateway realm bootstrap command to recreate a realm/zonegroup/zone does not work properly with **zonegroup_hostnames** field and the hostnames will not be set in the zonegroup.

As a workaround, set the zonegroup hostnames manually using the radosgw-admin tool. Bugzilla:2241321

5.2. CEPH OBJECT GATEWAY

Processing a query on a large Parquet object causes Ceph Object gateway processes to stop

Previously, in some cases, upon processing a query on a Parquet object, that object would be read chunk after chunk and these chunks could be quite big. This would cause the Ceph Object Gateway to load a large buffer into memory that is too big for a low-end machine; especially, when Ceph Object Gateway is co-located with OSD processes, which consumes a large amount of memory. This situation would trigger the OS to kill the Ceph Object Gateway process.

As a workaround, place the Ceph Object Gateway on a separate node and as a result, more memory is left for Ceph Object gateway, enabling it to complete processing successfully.

Bugzilla:2275323

Current RGW STS implementation does not support encryption keys larger than 1024 bytes

The current RGW STS implementation does not support encryption keys larger than 1024 bytes.

As a workaround, in **Keycloak: realm settings - keys**, edit the 'rsa-enc-generated' provider to have priority 90 rather than 100 and **keySize** as 1024 instead of 2048.

Bugzilla:2276931

Intel QAT Acceleration for Object Compression & Encryption

Intel QuickAssist Technology (QAT) is implemented to help reduce node CPU usage and improve the performance of Ceph Object Gateway when enabling compression and encryption. In this release, QAT can only be configured on new setups (Greenfield), which is a limitation of this feature. QAT Ceph Object Gateway daemons cannot be configured in the same cluster as non-QAT (regular) Ceph Object Gateway daemons.

Bugzilla:2284394

CHAPTER 6. SOURCES

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

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