Red Hat Ceph Storage 7.1

发行注记

Red Hat Ceph Storage 7.1 发行注记

Last Updated: 2024-07-03

Red Hat Ceph Storage 7.1 发行注记

Red Hat Ceph Storage 7.1 发行注记

法律通告

摘要

本发行注记介绍了为 Red Hat Ceph Storage 7.1 产品发行版本实施的主要功能、增强功能、已知问题和程序错误修复。

目录

使开源包含更多

红帽致力于替换我们的代码、文档和 Web 属性中存在问题的语言。我们从这四个术语开始：master、slave、黑名单和白名单。由于此项工作十分艰巨，这些更改将在即将推出的几个发行版本中逐步实施。有关更多详情，请参阅我们的首席技术官 Chris Wright 提供的消息。

提供有关 RED HAT CEPH STORAGE 文档的反馈

我们感谢您对文档提供反馈信息。请让我们了解如何改进文档。要做到这一点，创建一个 Bugzilla ticket：

1. 进入 *Bugzilla* 网站。
2. 在组件下拉列表中选择 **Documentation**。
3. 在 Sub-Component 下拉菜单中，选择适当的子组件。
4. 选择相应的文档版本。
5. 在 **Summary** 和 **Description** 字段中填写您要改进的建议。包括文档相关部分的链接。
6. 可选：添加一个附件（若有）。
7. 点 **Submit Bug**。

第 1 章 简介

Red Hat Ceph Storage 是一个可大规模扩展、开放、软件定义的存储平台，它将最稳定版本的 Ceph 存储系统与 Ceph 管理平台、部署实用程序和支持服务相结合。

Red Hat Ceph Storage 文档位于 https://access.redhat.com/documentation/zh-cn/red_hat_ceph_storage/7。

第 2 章 致谢

Red Hat Ceph Storage 版本 7.1 包含 Red Hat Ceph Storage 团队的许多贡献。此外，Ceph 社区中个人和组织的贡献质量和数量有显著的增长。我们借此感谢 Red Hat Ceph Storage 团队的所有成员、Ceph 社区中的所有个人贡献者，并且包括但不限于：

- Intel®
- Fujitsu ®
- UnitedStack
- Yahoo ™
- Ubuntu Kylin
- Mellanox ®
- CERN ™
- Deutsche Telekom
- Mirantis ®
- SanDisk ™
- SUSE ®

第 3 章 新功能

本节列出了本 Red Hat Ceph Storage 版本中引入的所有主要更新、增强功能和新功能。

3.1. CEPHADM 实用程序

用户现在可以在 idmap.conf中配置各种 NFS 选项

有了这个增强，在 **idmap.conf** 中引入了配置 NFS 选项的功能，如 "Domain", "Nobody-User", "Nobody-Group" 和 like。

Bugzilla:2068026

现在，通过 NFS 的新 haproxy 协议模式可以进行客户端 IP 限制

在以前的版本中，客户端 IP 限制无法在通过 NFS 使用 haproxy 的设置中工作。

在这个版本中，Cephadm 部署 NFS 支持 haproxy 协议。如果用户将 **enable_haproxy_protocol: True** 添加到其 ingress 和 haproxy 规格，或者将 **--ingress-mode haproxy-protocol** 传递给 **ceph nfs cluster create** 命令，NFS 守护进程将使用 haproxy 协议。

Bugzilla:2068030

用户现在必须输入一个用户名和密码才能访问 Grafana API URL

在以前的版本中，可以连接到 Grafana API URL 的任何人都可以访问它，而无需任何凭证。

在这个版本中，Cephadm 部署 Grafana 使用用户名和密码设置，以便用户访问 Grafana API URL。

Bugzilla:2079815

现在，可以将带有 NFS 后端的 Ingress 服务设置为只使用 keepalived 为 NFS 守护进程创建一个虚拟 IP (VIP)来绑定到，而无需涉及 HAProxy 层

在这个版本中，NFS 后端的 ingress 服务可以设置为只使用 **keepalived** 为 NFS 守护进程创建一个虚拟 IP，以便 NFS 守护进程绑定到，而无需涉及 HAProxy 层。当 NFS 守护进程被移动且客户端不需要使用不同的 IP 来进行连接时，这非常有用。

Cephadm 部署 **keepalived** 以设置 VIP，然后让 NFS 守护进程绑定到该 VIP。这也可以通过 **ceph nfs cluster create** 命令使用 NFS 模块来设置，使用标志 **--ingress --ingress-mode keepalive-only --virtual-ip <VIP>**。

规格文件类似如下：

```
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
```

-

这包括 **keepalive_ony: true** 设置。

NFS 规格类似如下：

```
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
```

这包括应与入口规格中的 VIP 匹配的 **virtual_ip** 字段。

Bugzilla:2089167

HAProxy 守护进程仅在附带 keepalived 创建的 VIP 上绑定到其前端端口

在这个版本中，HAProxy 守护进程只会在附带的 keepalived 创建的 VIP 上绑定到其前端端口，而不是在 0.0.0.0 上绑定。部署的 cephadm HAProxy 会将其前端端口绑定到 VIP，允许 NFS 守护进程等其他服务绑定到同一节点上其他 IP 上的端口 2049。

Bugzilla:2176297

ingress 服务的 HAProxy 健康检查间隔现在可以自定义

在以前的版本中，在某些情况下，第二个默认健康检查间隔太频繁，并导致不必要的流量。

在这个版本中，ingress 服务的 HAProxy 健康检查间隔可自定义。通过应用包含 **health_check_interval** 字段的 ingress 规格，服务的每个 HAProxy 守护进程生成的 HAProxy 配置将包括健康检查间隔的值。

Ingress 规格文件：

```
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
```

间隔的有效单位为： **us** microseconds **ms** : milliseconds **s** : seconds **m** : minutes **h** : hours **d** : days

Bugzilla:2199129

Grafana 现在绑定到主机上的特定网络中的 IP，而不是始终绑定到 0.0.0.0

在这个版本中，使用 Grafana 规格文件，其中包含带有 Grafana 绑定到 IP 的网络部分的 Grafana 规格文件，以及规格的 "spec" 部分的 **only_bind_port_on_networks: true**，Cephadm 将 Grafana 守护进程配置为绑定到该网络中的 IP，而不是 0.0.0.0。这可让用户使用 Grafana 用于另一个服务但主机上的不同 IP

端口。如果它是一个规格更新，它并不会导致它们被移动，则可以运行 **ceph orch redeploy grafana** 来获取设置的更改。

Grafana 规格文件：

```
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

现在，在 cephadm-ansible 模块中使用所有 bootstrap CLI 参数

在以前的版本中，只有 bootstrap CLI 参数的子集可用，它被限制模块使用。

在这个版本中，所有 bootstrap CLI 参数都可用于 **cephadm-ansible** 模块。

Bugzilla:2246266

Prometheus scrape 配置添加到 nfs-ganesha exporter

在这个版本中，Prometheus 提取配置添加到 nfs-ganesha 导出器中。这样做的目的是将 nfs-ganesha prometheus exporter 公开的指标提取到 Ceph 中运行的 Prometheus 实例中，这会进一步被 Grafana Dashboards 使用。

Bugzilla:2263898

Prometheus 现在绑定到主机上的特定网络中的 IP，而不是始终绑定到 0.0.0.0

在这个版本中，使用包含 Prometheus 绑定到 IP 的网络的 Prometheus 规格文件，以及规范的 "spec" 部分的 **only_bind_port_on_networks: true**，Cephadm 将 Prometheus 守护进程配置为绑定到该网络内的 IP，而不是 0.0.0.0。这可让用户使用 Prometheus 用于另一个服务但主机上的不同 IP 端口。如果是一个规格更新，它不会导致它们被移动，则可以运行 **ceph orch redeploy prometheus** 来获取设置的更改。

Prometheus 规格文件：

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

Bugzilla:2264812

用户现在可以挂载快照(exports 在 .snap 目录中)

有了这个增强，用户可以挂载快照（位于 **.snap** 目录中）查看 RO 模式。使用 NFS MGR 模块创建的 NFS 导出现在包含 **cmount_path** 设置（这不能配置），并且应保留为"/"），允许挂载快照。

Bugzilla:2245261

zonegroup 主机名现在可以使用 ceph rgw realm bootstrap... 命令中提供的规格文件来设置

在这个版本中，在继续到 Ceph 对象网关多站点设置的自动化中，用户现在可以通过 bootstrap 命令 **ceph rgw realm bootstrap...** 传递的初始规格文件来设置 zonegroup 主机名。

例如，

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

如果用户将上述部分添加到 realm bootstrap 命令传递的 Ceph 对象网关规范文件的"特定"部分中，Cephadm 将在 Ceph 对象网关模块完成 realm/zonegroup/zone 的规范中定义的 zonegroup 中自动将这些主机名添加到规范中定义的 zonegroup 中。请注意，这可能需要几分钟时间来看 Cephadm 模块当前完成的其他活动。

3.2. CEPH 仪表板

CephFS 快照在 Ceph 仪表板上调度管理

在以前的版本中，CephFS 快照调度只能通过命令行界面管理。

在这个版本中，可以列出 CephFS 快照调度，创建、编辑、激活、取消激活并从 Ceph 仪表板中删除。

Bugzilla:2264145

Ceph 仪表板现在在 Ceph 仪表板中支持基于 NFSv3 的导出

在这个版本中，在 Ceph 仪表板中启用对基于 NFSv3 的导出管理的支持。

Bugzilla:2267763

添加了为 CephFS 管理 Ceph 用户的功能

在这个版本中，添加了为 CephFS 管理 Ceph 用户的功能。这提供了从 File System 视图管理用户对卷、子卷组和子卷的权限。

Bugzilla:2271110

添加了用于多站点同步状态的新 API 端点

在以前的版本中，多站点同步状态只能通过 CLI 命令获得。

在这个版本中，多站点状态通过 Ceph 仪表板中的 API 添加。多站点同步状态的新 API 端点是 **api/rgw/multisite/sync_status**。

Bugzilla:2258951

改进了 NVMe-oF 网关的监控

在这个版本中，为了提高 NVMe-oF 网关的监控，会根据发出的指标添加 NVMe-oF 网关的警报，在 NVMe-oF 网关中提取来自内嵌的 prometheus exporter 的指标。

Bugzilla:2276038

Ceph 仪表板中的 CephFS 克隆管理

在这个版本中，Ceph 控制面板中提供了 CephFS 克隆管理功能。用户可以通过 Ceph 控制面板创建和删除子卷克隆。

Bugzilla:2264142

Ceph 仪表板中的 CephFS 快照管理

在这个版本中，Ceph 控制面板中提供了 CephFS 快照管理功能。用户可以通过 Ceph 控制面板创建和删除子卷快照。

Bugzilla:2264141

每个用户/存储桶标记的 Performance Counters

在这个版本中，用户无法获取每个 Ceph 对象网关节点发生的操作的信息，也可以在 Ceph 控制面板中查看 Ceph 对象网关性能计数器。

标记的 Sync Performance Counters into Prometheus

在这个版本中，用户可以从 Prometheus 收集有关区域间复制健康状况的实时信息，以提高 Ceph 对象网关多站点同步操作的可观察性。

在 Ceph 仪表板中添加并编辑存储桶

在这个版本中，作为 Ceph 仪表板的 Ceph 对象网关改进的一部分，添加了 Ceph 仪表板中应用、列出和编辑 Buckets 的功能。

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

在 Ceph 仪表板中添加、列出、删除和应用存储桶策略

在这个版本中，作为 Ceph 仪表板的 Ceph 对象网关改进的一部分，添加了从 Ceph 仪表板添加、列出、删除和应用 bucket 策略的功能。

3.3. CEPH 文件系统

MDS 动态元数据负载均衡器默认为 off

在以前的版本中，通过增加 **max_mds** 文件系统设置，均衡均衡的负载均衡器行为会以不可靠的方式进行碎片树。

在这个版本中，MDS 动态元数据负载均衡器默认是 off。Operator 必须明确打开负载均衡器才能使用它。

Bugzilla:2227309

CephFS 支持静止子卷或目录树

在以前的版本中，多个客户端会在一致的快照屏障之间交错读取和写入，在客户端之间存在带外通信。这种通信导致客户端错误地认为它们已达到通过快照相互恢复的检查点。

在这个版本中，CephFS 支持静止子卷或目录树，以启用崩溃一致性快照。现在，客户端被强制在 MDS 执行快照前静止所有 I/O。这会在子树的所有客户端之间强制实施检查点。

Bugzilla:2235753

MDS Resident Segment Size (RSS)性能计数器使用更高的优先级进行跟踪

在这个版本中，MDS Resident Segment Size 性能计数器会根据优先级进行跟踪，以允许调用者消耗其值来生成有用的警告。这允许 Rook 识别 MDS RSS 大小并相应地操作。

Bugzilla:2256560

现在，只有没有 laggy OSD 时，laggy 客户端才会被驱除

在以前的版本中，MDS 中的监控性能转储有时会显示 OSD 为 laggy、**objecter.op_laggy** 和 **objecter.osd_laggy**，从而导致 laggy 客户端(dirty 数据无法清除大写)。

在这个版本中，如果 **defer_client_eviction_on_laggy_osds** 选项被设置为 true，并且客户端会因为 laggy OSD 造成滞后，则客户端驱除不会发生，直到 OSD 不再被滞后。

Bugzilla:2260003

CephFS-mirror 守护进程通过 perf dump 命令导出快照同步性能计数器

ceph-mds 守护进程会导出现有 **perf dump** 命令中包含的每个客户端性能计数器。

Bugzilla:2264177

引入了一个新的 dump dir 命令转储目录信息

有了这个增强，引进了 **dump dir** 命令，来转储目录信息并打印输出。

Bugzilla:2269687

对子卷的快照调度支持

在这个版本中，为子卷提供快照调度支持。所有快照调度命令都接受 **--subvol** 和 **--group** 参数，以引用适当的子卷和子卷组。如果在没有子卷组参数的情况下指定子卷，则会考虑默认的子卷组。另外，当引用子卷时，不需要指定有效的路径，且因为使用了参数解析的性质，只需要一个占位符字符串就足够了。

Example

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

Bugzilla:2238537

添加或修改 MDS caps 的 Ceph 命令说明为什么用户传递的 MDS caps 被拒绝

在以前的版本中，添加或修改 MDS caps printed "Error EINVAL: mds capability parse failed, stopped at 'allow w' of 'allow w' 的 Ceph 命令。

有了这个增强，命令说明为什么用户传递的 MDS caps 被拒绝，且 MDS caps 中的 Permission 标志必须以 'r' 或 'rw' 开始，或者 '*' 或 'all'。

Bugzilla:2247586

3.4. CEPH 对象网关

现在，添加了 admin 接口来管理存储桶通知

在以前的版本中，S3 REST API 用于管理存储桶通知。但是，如果管理员希望覆盖它们，则无法通过 radosgw-admin 工具进行此操作。

在这个版本中，添加了一个带有以下命令的管理界面来管理存储桶通知：

```
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

现在，当运行 ceph 计数器转储 时，带有 user 和 bucket 操作计数器的 RGW 在不同的部分中是不同的部分

在以前的版本中，所有标记为操作的 RGW 都位于 **ceph counter dump** 命令的输出的 **rgw_op'** 部分中，但会具有用户标签或存储桶标签。

在这个版本中，在执行 **ceph counter dump** 命令时，标记为 user 和 bucket 操作计数器的 RGW 分别位于 **rgw_op_per_user** 或 **rgw_op_per_bucket** 部分中。

Bugzilla:2265574

用户现在可以使用 -t 命令行选项将临时文件放入目录中

在以前的版本中，**/usr/bin/rgw-restore-bucket-index** 工具只使用 **/tmp**，该目录有时没有足够的可用空间来保存所有临时文件。

有了这个增强，用户可以指定一个目录，可以使用 **-t** 命令行选项放置临时文件，并在它们用尽空间时通知这些目录，因此知道要重新运行工具的调整。另外，用户可以定期检查工具的临时文件是否已耗尽存在临时文件的文件系统上的可用空间。

Bugzilla:2267715

现在支持使用 copy-object API 复制加密对象

在以前的版本中，在 Ceph 对象网关中，不支持使用 copy-object API 复制加密对象，因为其服务器端加密支持。

在这个版本中，支持使用 copy-object API 复制加密对象，依赖复制对象操作的工作负载也可以使用服务器端加密。

Bugzilla:2149450

添加了一个新的 Ceph 对象网关 admin-ops 功能，以允许读取用户元数据，但不读取其关联的授权密钥

在这个版本中，添加了一个新的 Ceph 对象网关 admin-ops 功能，以允许读取 Ceph 对象网关用户元数据，但不能读取其关联的授权密钥。这是为了减少自动化和报告工具的特权，并避免模拟用户或查看其密钥。

Bugzilla:2112325

云转换：添加新的受支持的 S3 兼容平台

在这个版本中，若要将对象存储移到云或其他内部 S3 端点，当前的生命周期转换和存储类模型已扩展。现在，云存档功能支持 S3 兼容平台，如 IBM Cloud Object Store (COS)和 IBM Storage Ceph。

使用 RGW 后端的 NFS

在这个版本中，带有 Ceph 对象网关后端的 NFS 使用现有功能重新获得。

3.5. 多站点 CEPH 对象网关

radosgw-admin sync status 命令中引入了重试机制

在以前的版本中，当多站点同步向远程区域发送请求时，它使用 round robin 策略来选择其其中一个区域端点。如果该端点不可用，**radosgw-admin sync status** 命令使用的 http 客户端逻辑不提供重试机制，因此会报告输入/输出错误。

在这个版本中，如果所选端点不可用，在 sync status 命令中引入了重试机制，则会选择不同的端点来提供请求。

Bugzilla:1995152

NewerNoncurrentVersions,ObjectSizeGreaterThan, 和 ObjectSizeLessThan 过滤器被添加到生命周期中

在这个版本中，对 **NewerNoncurrentVersions**、**ObjectSizeGreaterThan**、**ObjectSizeLessThan** 过滤器的支持被添加到生命周期中。

Bugzilla:2172162

现在支持 S3 复制 API

在这个版本中，支持用户 S3 复制 API。使用这些 API，用户可以在存储桶级别上设置复制策略。API 被扩展，使其包含用于指定源和目标区域名称的额外参数。

Bugzilla:2279461

bucket Granular Sync Replication GA （第 3 部分）

在这个版本中，通过存储桶粒度支持将存储桶或一组存储桶复制到不同的 Red Hat Ceph Storage 集群。可用性要求是 Ceph 对象网关多站点。

3.6. RADOS

在/off 中设置 noautoscale 标志可保留每个池的原始自动扩展模式配置

在以前的版本中，当设置了 **no autoscale** 标志时，**pg_autoscaler 不会在每个池的自动扩展模式配置** 中保留。因此，每当设置了 **noautoscale** 标志时，必须为每个池重复设置 **autoscale** 模式。

在这个版本中，**pg_autoscaler** 模块在设置 **noautoscale** 标志 后为自动扩展模式保留单独的池配置。在/off 中设置 **noautoscale** 标志仍然保留每个池的原始自动扩展模式配置。

Bugzilla:2136766

引入了 reset_purged_snaps_last OSD 命令

在这个版本中，引进了 **reset_purged_snaps_last** OSD 命令，以解决在 OSD 中缺少 **purged_snaps** 密钥(PSN)的情况，并存在于 monitor 中。**purged_snaps_last** 命令将为零，因此监控器将在下次引导时与 OSD 共享所有 **purged_snaps** 信息。

Bugzilla:2251188

启用 BlueStore 的 RocksDB 压缩

在这个版本中，为了确保元数据（特别是 OMAP）占用较少的空间，会修改 RocksDB 配置以启用其数据的内部压缩。

因此，在压缩过程中，* 数据库大小为较小的。X 平均 I/O 平均 I/O 越小，X CPU 用量较高

Bugzilla:2253313

OSD 现在对致命崩溃更具弹性

在以前的版本中，特殊的 OSD 层对象 "superblock" 会被覆盖，因为位于磁盘的开头，从而导致严重崩溃。

在这个版本中，OSD "superblock" 是冗余的，并在磁盘上迁移。其副本存储在数据库中。OSD 现在对致命损坏更具弹性。

Bugzilla:2079897

3.7. RADOS 块设备 (RBD)

改进了 rbd_diff_iterate2 （） API 性能

在以前的版本中，如果专用锁定在与 **fast-diff** 模式开头的时间(**fromsnapname == NULL**)以 fast-diff 模式(entire_**object == true**)时，如果启用了 fast-diff 镜像功能并有效，则 RBD diff-iterate 无法保证执行。

在这个版本中，**rbd_diff_iterate2 （）** API 性能有所改进，从而提高了 QEMU 实时磁盘同步和备份用例的性能，其中启用了 **fast-diff** 镜像功能。

Bugzilla:2258997

第 4 章 程序错误修复

本节论述了在这个 Red Hat Ceph Storage 发行版本中修复的用户有严重影响的错误。此外，部分还包括之前版本中发现的固定已知问题的描述。

4.1. CEPHADM 实用程序

将 --name NODE 标志与 cephadm shell 搭配使用，以启动已停止的 OSD 不再返回错误的镜像容器

在以前的版本中，在某些情况下，在使用 **cephadm shell --name NODE** 命令时，该命令会使用错误的工具版本启动容器。当用户在主机上具有较新的 ceph 容器镜像而不是其 OSD 使用的镜像时，会出现这种情况。

在这个版本中，Cephadm 在使用带有 **--name** 标志的 cephadm shell 命令时，Cephadm 决定已停止守护进程的容器镜像。用户不再有问题使用 **--name** 标志，命令可以正常工作。

Bugzilla:2258542

4.2. CEPH ANSIBLE 实用程序

Playbook 现在删除与正在运行的 RHEL 版本匹配的 RHCS 版本存储库

在以前的版本中，playbook 会尝试从 RHEL 9 中删除 Red Hat Ceph Storage 4 软件仓库，即使它们在 RHEL 9 中不存在。这会导致 playbook 失败。

在这个版本中，playbook 删除与运行的 RHEL 版本匹配的现有 Red Hat Ceph Storage 版本存储库，并删除正确的存储库。

Bugzilla:2258940

4.3. NFS GANESHA

现在，配置重新加载进程消耗的所有内存都已被释放

在以前的版本中，重新载入导出不会释放配置重新载入进程消耗的所有内存，从而导致内存占用增加。

在这个版本中，由配置重新载入进程消耗的所有内存都会被释放，从而减少内存占用量。

Bugzilla:2265322

4.4. CEPH 仪表板

用户可以在 Ceph 仪表板中使用多个主机创建卷

在这个版本中，用户可以在 Ceph 控制面板中使用多个主机创建卷。

Bugzilla:2241056

取消设置子卷大小不再设置为 'infinite'

在以前的版本中，取消设置子卷大小被设置为 'infinite'，从而导致更新失败。

在这个版本中，将大小设置为 'infinite' 的代码已被删除，更新可以正常工作。

Bugzilla:2251192

在内核挂载命令中添加缺少的选项

在以前的版本中，在内核挂载命令中缺少几个选项来附加文件系统，从而导致命令无法按预期工作。

在这个版本中，添加了缺少的选项，内核 mount 命令可以正常工作。

Bugzilla:2266256

Ceph 仪表板现在支持 NFS v3 和 v4-enabled 导出管理

在以前的版本中，Ceph 仪表板只支持启用了 NFSv4 的导出管理，不支持启用了 NFSv3 的导出。因此，通过 CLI 为 NFSv3 导出的任何管理都被破坏。

在这个版本中，通过有一个额外的复选框来启用对基于 NFSv3 的导出管理的支持。Ceph 仪表板现在支持 v3 和 v4-enabled 导出管理。

Bugzilla:2267814

现在，在创建区时 access/secret 密钥不会编译

在以前的版本中，在 Ceph 对象网关多站点中创建区域时，访问/secret 密钥被编译。因此，用户必须首先在区中设置非系统用户的密钥，并使用系统用户的密钥进行更新。

在这个版本中，在创建区时不会编译 access/secret 密钥。

Bugzilla:2275463

导入多站点配置不会在提交表单时抛出错误

在以前的版本中，多站点周期信息不包含 'realm' 名称。因此，在提交表单时导入多站点配置。

在这个版本中，从周期信息获取 'realm' 名称的检查已被删除，令牌导入可以正常工作。

Bugzilla:2275861

Ceph 对象网关指标标签名称与 Prometheus 标签命名格式一致，它们现在在 Prometheus 中可见

在以前的版本中，指标标签名称与 Prometheus 标签命名格式不一致，从而导致 Ceph 对象网关指标在 Prometheus 中不可见。

在这个版本中，连字符(-)被替换为 Ceph 对象网关指标标签名称中的下划线(_)，其中是否适用，所有 Ceph 对象网关指标现在在 Prometheus 中可见。

Bugzilla:2276340

全名现在可以在 Ceph 仪表板中包含点

在以前的版本中，在 Ceph 仪表板中，无法创建或修改带有点的全名，因为验证不正确。

在这个版本中，验证会被正确调整，以在 Ceph 仪表板中包含句点。

Bugzilla:2249812

4.5. CEPH 文件系统

现在，会批量添加带有 FSMap 更改的 MDS 元数据以确保一致性

在以前的版本中，监控器有时会在升级过程中丢失 MDS 元数据跟踪，并取消 PAXOS 事务会导致 MDS 元数据不再可用。

在这个版本中，带有 FSMap 更改的 MDS 元数据会添加到批处理中，以确保一致性。**ceph mds metadata** 命令现在可以正常工作。

Bugzilla:2144472

检测到 ENOTEMPTY 输出，并正确显示消息

在以前的版本中，当运行 **子卷组 rm** 命令时，卷插件中没有检测到 **ENOTEMPTY** 输出会导致常规的错误消息而不是特定消息。

在这个版本中，当 **子卷组中存在子卷且信息正确显示时，会检测到子卷组 rm** 命令的输出。

Bugzilla:2240138

MDS 现在会在请求清理过程中自动排队下一个客户端重播请求

在以前的版本中，有时 MDS 不会排队在 up:client-replay 状态中重播的下一个客户端请求，从而导致 MDS 挂起。

在这个版本中，下一个客户端重播请求会在请求清理过程中自动排队，MDS 会正常进行故障转移恢复。

Bugzilla:2243105

CephFS-mirroring 整体性能有所改进

在这个版本中，增量快照同步已被修正，这可以提高 cephfs-mirroring 的整体性能。

Bugzilla:2248639

loner 成员被设置为 true

在以前的版本中，对于 LOCK_EXCL_XSYN 状态的文件锁定，非外部客户端将发出空上限。但是，由于此状态的 loner 被设置为 false，所以可能会导致锁定程序向它们发出 Fcb caps，这是不正确的。这会导致一些客户端请求错误地撤销一些上限和无限等待，并导致请求较慢。

在这个版本中，loner 成员被设置为 true，因此对应的请求不会被阻止。

Bugzilla:2251258

每月快照的 snap-schedule 重复和保留规格从 m 改为 M

在以前的版本中，每月快照的 snap-schedule 重复规格和保留规格与其他 Ceph 组件不一致。

在这个版本中，规格从 m 改为 M，它现在与其他 Ceph 组件一致。例如，要保留 5 个每月快照，您需要发出以下命令：

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

Bugzilla:2264348

当在多mds 集群中复制一些内节点时，Ceph-mds 不再崩溃

在以前的版本中，由于 ceph-mds 中锁定断言不正确，当多mds 集群中复制某些 inode 时，ceph-mds 会崩溃。

在这个版本中，断言中的锁定状态会被验证，且不会观察崩溃。

Bugzilla:2265415

缺少的字段，如 日期,client_count, 过滤器添加到 --dump 输出中

在这个版本中，缺少的字段，如 日期,client_count,过滤器 被添加到 --dump 输出中。

Bugzilla:2272468

在恢复过程中，MDS 不再失败并显示 assert 功能

在以前的版本中，当恢复失败的等级时，MDS 有时会错误地报告元数据损坏，因此，使用 assert 函数会失败。

在这个版本中，启动过程已被修正，MDS 在恢复过程中不会失败。

Bugzilla:2272979

目标 mon_host 详情已从 peer List 和 mirror 守护进程状态中删除

在以前的版本中，快照镜像 peer-list 会显示比 peer 列表更多的信息。如果应该只显示一个 MON IP 或所有 MON 主机 IP，则此输出会导致混淆。

在这个版本中，mon_host 已从 fs snapshot mirror peer_list 命令中删除，目标 mon_host 详情已从 peer List 和 mirror 守护进程状态中删除。

Bugzilla:2277143

目标 mon_host 详情已从 peer List 和 mirror 守护进程状态中删除

在以前的版本中，quiesce 协议代码引入了一个回归。在终止客户端请求时，它将跳过为批处理操作选择新的批处理头。这会导致过时的批处理头请求永久保留在 MDS 缓存中，然后被视为较慢的请求。

在这个版本中，在终止请求时选择一个新的批处理头，且没有由批处理操作导致的请求速度。

Bugzilla:2277944

即使没有 MDS，也会进行文件系统升级

在以前的版本中，当所有 MDS 都停机时，监控器不允许 MDS 升级文件系统。因此，当 fail_fs 设置被设置为 'true' 时，升级会失败。

在这个版本中，监控器允许在没有 MDS 启动时进行升级。

Bugzilla:2244417

4.6. CEPH 对象网关

admin topic list 命令中不再显示自动生成的内部主题

在以前的版本中，自动生成的内部主题通过 topic list 命令向用户公开，因为用户可以看到很多主题，而不是创建的内容。

在这个版本中，admin topic list 命令中不会显示自动生成的内部主题，用户现在只会看到预期的主题列表。

Bugzilla:1954461

在 topic list 命令中不再显示已弃用的存储桶名称字段

在以前的版本中，如果拉取模式通知(pubsub)，通知存储在存储桶中。但是，虽然此模式已弃用，但主题 list 命令中仍会显示空 bucket name 字段。

在这个版本中，空 bucket name 字段会被删除。

Bugzilla:1954463

现在，通知会在生命周期转换时发送

在以前的版本中，在转换时分配的逻辑（与过期时间不同）。因此，在转换时不会看到通知。

在这个版本中，添加了新的逻辑，并在生命周期转换中发送通知。

Bugzilla:2166576

RGWCopyObjRequest 已被修复，重命名操作可以正常工作

在以前的版本中，在 zipper 转换后，RGWCopyObjRequest 初始化不正确，会破坏重命名操作。因此，许多 rgw_rename （） 场景无法复制源对象，因此由于辅助问题，也会删除源，即使副本失败也是如此。

在这个版本中，RGWCopyObjRequest 被修正，并为不同的重命名操作添加了几个单元测试情况。

Bugzilla:2217499

Ceph 对象网关无法再进行非法访问

在以前的版本中，在初始化前访问代表 Ceph 对象网关角色的变量，从而导致 segfault。

在这个版本中，操作会被重新排序，且没有非法访问。角色根据需要强制执行。

Bugzilla:2252048

现在，会针对错误的 CSV 对象存储显示错误消息

在以前的版本中，带有未关闭双引号的 CSV 文件会导致 assert，然后是 crash。

在这个版本中，引入了一个错误消息，它根据错误的 CSV 对象结构弹出。

Bugzilla:2252396

在 Ceph 仪表板中查询用户相关信息时，用户不再遇到 'user not found' 错误

在以前的版本中，在 Ceph 仪表板中，最终用户无法从 Ceph 对象网关检索与用户相关的信息，因为仪表板无法识别的完整 user_id 中存在命名空间，从而导致出现 "user not found" 错误。

在这个版本中，一个完全构建的用户 ID，其中包括 租户、命名空间 和 user_id，当 GET 请求发送到 admin ops 时，会单独返回每个字段来获取用户信息。最终用户现在可以检索正确的 user_id，可用于进一步从 Ceph 对象网关获取其他用户相关信息。

Bugzilla:2255255

Ceph 对象网关现在通过新流编码表单格式良好的有效负载来传递请求

在以前的版本中，Ceph 对象网关无法识别 STREAMING-AWS4-HMAC-SHA256-PAYLOAD 和 STREAMING-UNSIGNED-PAYLOAD-TRAILER 编码形式，从而导致请求失败。

在这个版本中，识别、解析以及适用的逻辑验证为新编码表单提供的新结尾请求签名。Ceph 对象网关现在通过具有新流编码形式格式良好的有效负载的请求。

Bugzilla:2256967

现在，radosgw admin bucket 和 bucket reshard stat 计算的检查 stat 计算是正确的

在以前的版本中，由于代码更改，radosgw-admin bucket check stat 计算和 bucket reshard stat 计算在存在从未指定版本转换为版本时不正确的计算。

在这个版本中，计算已被修正，不正确的存储桶 stat 输出不再生成。

Bugzilla:2257978

在多部分上传失败时，tail 对象不再丢失

在以前的版本中，在多部分上传过程中，如果因为情况（如超时）上传失败，上传被重启，第一次尝试清理会从后续尝试中删除 tail 对象。因此，生成的 Ceph 对象网关多部分对象会损坏，因为缺少一些 tail 对象。它将响应 HEAD 请求，但在 GET 请求期间失败。

在这个版本中，代码会正确清理第一次尝试。生成的 Ceph 对象网关多部分对象不再损坏，并可以被客户端读取。

Bugzilla:2262650

CompleteMultipartUpload 及其通知中的 ETag 值现在存在

在以前的版本中，与通知相关的更改会导致对象句柄与完成多部分上传对应，不包含生成的 ETag。因此，因为 CompleteMultipartUpload 及其通知，无法完成多部分上传 ETags。（计算并存储了正确的 ETag，因此后续操作包含正确的 ETag 结果。）

在这个版本中，CompleteMultipartUpload 会刷新对象，并按预期打印它。CompleteMultipartUpload 及其通知中存在 ETag 值。

Bugzilla:2266579

通过 swift 列出容器(bucket)不再会导致 Ceph 对象网关崩溃

在以前的版本中，swift-object-storage 调用路径缺少调用，用于更新其对应存储桶(zipper backport 问题)的对象句柄。因此，通过 swift 列出容器(bucket)会在为同一存储桶配置 S3 网站时导致 Ceph 对象网关崩溃。

在这个版本中，添加了所需的 zipper 逻辑，崩溃不再发生。

Bugzilla:2269038

现在，在没有生命周期策略的存储桶上处理生命周期不会崩溃

在以前的版本中，尝试在没有生命周期策略的情况下手动处理存储桶的生命周期，导致 radosgw-admin 程序崩溃。

在这个版本中，在处理处理前检查 null bucket 句柄，以避免崩溃。

Bugzilla:2270402

现在，可以修改 datapool 的区详情

rgw::zone_create （） 函数在创建区域时初始化默认放置目标和池名称。此功能之前还用于设置了 exclusive=false 的 radosgw-admin zone。但是，zone set 不允许修改 STANDARD 存储类的 data_pool。

在这个版本中，如果 default-placement 目标已存在，并且 datapool 的区域详情可以如预期修改，则 default-placement 目标不应被覆盖。

Bugzilla:2254480

现在，浮点数上的 modulo 操作现在返回正确的结果

在以前的版本中，对浮点数的 modulo 操作返回错误的结果。

在这个版本中，SQL 引擎已被改进，以处理浮点数并返回正确的结果。

Bugzilla:2254125

SQL 语句可以正确地返回不区分大小写的布尔值表达式的结果

在以前的版本中，SQL 语句包含一个带有声明部分大写字母的布尔值表达式，从而导致错误的解释和

错误的结果。

在这个版本中，语句的解释区分大小写，因此会为任何情况返回正确的结果。

Bugzilla:2254122

SQL 引擎返回正确的 NULL 值

在以前的版本中，SQL 语句包含来自 NULL 的 cast 类型，因此会返回错误的结果，而不是返回 NULL。

在这个版本中，SQL 引擎识别来自 NULL 的 cast 并返回 NULL。

Bugzilla:2254121

eTags 值现在包括在 CompleteMultipartUpload 及其通知中

在以前的版本中，与通知相关的更改会导致对象句柄，对应于完成多部分上传，不包含生成的 ETag。因此，CompleteMultipartUpload 及其通知不存在 ETags。（计算并存储了正确的 ETag，因此后续操作包含正确的 ETag 结果。）

在这个版本中，CompleteMultipartUpload 会刷新对象，并按预期打印它。ETag 值现在存在于 CompleteMultipartUpload 及其通知中。

Bugzilla:2249744

将对象名称中的嵌入式反斜杠(//)工作负载发送到 cloud-sync 不再会导致同步失败

在以前的版本中，在云同步过程中，当工作负载包含带有嵌入式反斜杠(/)的对象（即，使用虚拟目录路径）的对象时，在云同步过程中的 URLescaping 会导致同步失败。

在这个版本中，修正不正确的转义，对象名称中的嵌入式反斜杠(/)中的工作负载可以按预期发送到 cloud-sync。

Bugzilla:2249068

包含布尔值表达式的 SQL 语句返回布尔值类型

在以前的版本中，包含布尔值表达式（投射）的 SQL 语句将返回字符串类型而不是布尔值类型。

在这个版本中，引擎根据语句语法将字符串标识为布尔值表达式，引擎可以成功返回布尔值类型 (true/false)。

Bugzilla:2254582

现在，工作调度程序会考虑 should_work 函数中的下一个日期

在以前的版本中，在 should_work 函数中使用的逻辑决定了生命周期是否应在当前时间开始运行，不会考虑下一个日期。因此，当 AB < XY 时，任何自定义工作时间 "XY:TW-AB:CD" 都会破坏生命周期处理。

在这个版本中，工作调度程序会考虑下一个日期，各种自定义生命周期工作调度现在可以正常工作。

Bugzilla:2255938

merge_and_store_attrs （） 方法不再导致属性更新操作失败

在以前的版本中，merge_and_store_attrs （） 方法中的一个错误，它处理协调更改和未更改的存储桶实例属性，从而导致一些属性更新操作以静默方式失败。因此，存储桶子集的一些元数据操作会失败。例如，存储桶所有者更改会在设置了速率限制的存储桶上失败。

在这个版本中，merge_and_store_attrs （） 方法已被修复，所有受影响的场景都可以正常工作。

Bugzilla:2262919

checksum 和 malformed trailers 不再会导致崩溃

在以前的版本中，在 java AWS4Test.testMultipartUploadWithPauseAWS4 过程中的 AWSv4ComplMulti 异常会导致一些客户端输入崩溃，特别是使用校验和跟踪器的用户。

在这个版本中，在 do_aws4_auth_completion （） 中实施了异常处理程序。checksum 和 malformed trailers 不再会导致崩溃。

Bugzilla:2266092

改进了尾部块边界检测的实现

在以前的版本中，没有处理 0-length 尾随块边界格式的有效格式。因此，Ceph 对象网关无法正确识别尾部块的开头，从而导致 403 错误。

在这个版本中，实现了改进的块边界检测，匿名访问案例中不再发生意外的 403 错误。

Bugzilla:2266411

Kafka 消息和闲置超时的默认值不再会导致挂起

在以前的版本中，Kafka 消息和闲置超时的默认值会在等待 Kafka 代理时造成意外挂起。

在这个版本中，超时会被调整，它不再挂起。

Bugzilla:2269381

删除存储桶标记不再失败

在以前的版本中，RADOS SAL merge_and_store_attrs （） 中的一个不正确的逻辑会导致删除的属性不材料。这也会影响 DeleteLifecycle。因此，纯属性删除不会在某些代码路径中生效。

在这个版本中，存储存储桶标签的逻辑使用 RADOS SAL put_info （） 而不是 merge_and_store_attrs （）。现在，删除存储桶标记会如预期成功。

Bugzilla:2271806

现在，S3 PutACL 和 ACL 更改上的对象 mtime 可以被正确复制

在以前的版本中，S3 PutACL 操作不会更新对象 mtime。因此，应用后 ACL 不会改变，因为基于时间戳的对象更改检查会错误地返回 false。

在这个版本中，S3 PutACL 和 ACL 更改会正确复制对象 mtime。

Bugzilla:2271938

现在，所有转换情况都可以发送通知

在以前的版本中，因为池转换上的通知没有发送，在转换时分配通知的逻辑被错误地范围到 cloud-transition。

在这个版本中，通知分配添加到池转换范围内，所有转换情况都可以分配通知。

Bugzilla:2279607

在 2106 年后 RetainUntilDate 不再截断，对于新的 PutObjectRetention 请求，可以正常工作

在以前的版本中，PutObjectRetention 请求在 2106 年后指定 RetainUntilDate 请求会截断，从而导致早期用于对象锁定强制的日期。这不会影响 ' PutBucketObjectLockConfiguration' 请求，其中持续时间以天为单位。

在这个版本中，对于新的 PutObjectRetention 请求，RetainUntilDate 可以正常工作。之前存在的请求不会自动修复。要修复现有请求，请根据 x-amz-object-lock-retain-until-date 使用 HeadObject 请求来识别请求，并使用 RetainUntilDate 再次保存。

如需更多信息，请参阅 *S3 放置对象保留*

Bugzilla:2265890

bucket 生命周期处理规则不再停滞

在以前的版本中，per-shard bucket-lifecycle 规则的枚举包含与并发删除存储桶的生命周期规则相关的逻辑错误。因此，分片可能会进入一个状态，该状态将停止处理该分片，从而导致无法处理一些存储桶生命周期规则。

在这个版本中，枚举可以跳过删除的条目，且与此问题相关的生命周期处理停滞已被解决。

Bugzilla:2270334

删除版本存储桶中的对象会导致统计不匹配

由于版本存储桶混合使用当前和非当前对象，因此删除对象可能会导致在本地和远程站点上的 bucket 和用户统计差异。这不会导致在任一站点上出现对象泄漏，只是统计信息。

Bugzilla:1871333

4.7. 多站点 CEPH 对象网关

Ceph 对象网关在删除对象的过程中不再死锁

在以前的版本中，在对象删除过程中，Ceph 对象网关 S3 DeleteObjects 将与多站点部署一起运行，从而导致 Ceph 对象网关死锁并停止接受新请求。这是因为 DeleteObjects 请求一次处理多个对象删除。

在这个版本中，复制日志被序列化，死锁会被阻止。

Bugzilla:2249651

CURL 路径规范化现在在启动时被禁用

在以前的版本中，由于 CURL 执行的路径规范化(Ceph 对象网关复制堆栈的一部分)，对象名称在复制过程中被静默重新格式化。因此，名称包含嵌入式 . 和 .. 的对象不会被复制。

在这个版本中，CURL 路径规范化在启动时被禁用，受影响的对象会如预期复制。

Bugzilla:2265148

在主站点上转发请求的验证不再失败

在以前的版本中，如果使用 STS 返回的临时凭证为请求签名，S3 请求会发出到次请求。发生了故障，因为请求将使用与转发请求的会话令牌中的临时凭据不匹配的系统用户凭证转发到主和签名。由于不匹配的凭据，主站点上转发请求的身份验证会失败，这会导致 S3 操作失败。

在这个版本中，当请求从 secondary 转发到 primary 时，身份验证是通过在会话令牌中使用临时凭证

传递的。系统用户的凭据用于成功完成身份验证。

Bugzilla:2271399

4.8. RADOS

如果池中存储的零对象，Ceph 会报告 POOL_APP_NOT_ENABLED 警告

在以前的版本中，如果为 RGW 池启用了应用程序标签，Ceph 状态将无法报告池应用程序警告，从而导致 RGW 存储桶创建失败。

在这个版本中，Ceph 会报告 POOL_APP_NOT_ENABLED 警告，即使池存储有零个对象。

Bugzilla:2029585

为扩展集群的两个站点之间不均匀的 OSD 权重添加了检查

在以前的版本中，扩展集群部署后没有检查相等的 OSD 权重。因此，用户可以使 OSD 权重变得不等。

在这个版本中，会添加检查，以便在扩展集群的两个站点之间不均匀 OSD 权重。集群现在提供两个站点之间不均匀 OSD 权重的警告。

Bugzilla:2125107

当设置 norecover 标志时，自动扩展不再运行

在以前的版本中，当设置 norecover 标志时，自动扩展会运行，从而导致创建新 PG 和需要回填的 PG。在设置 norecover 标志时，在缺少或降级对象时允许运行自动扩展，以避免无限期地挂起客户端 I/O。

在这个版本中，当设置 norecover 标志时，自动扩展不会运行。

Bugzilla:2134786

ceph config dump 命令输出现在一致

在以前的版本中，没有用户用户格式化的输出的 ceph config dump 命令显示本地化选项名称及其值。一个规范化与本地化选项的示例如下所示：

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

但是，命令的用户打印（如 JSON）版本仅显示上例中所示的规范化选项名称。ceph config dump 命令结果在 with 和 without the pretty-print 选项之间不一致。

在这个版本中，输出一致，在使用 ceph config dump --format TYPE 命令时，始终显示本地化选项名称，并将 TYPE 作为 pretty-print 类型。

Bugzilla:2213766

MGR 模块不再每分钟占用一个 CPU 内核，CPU 使用率是正常的

在以前的版本中，从放置组 auto-scaler 模块获取 OSDMap 的昂贵调用会导致 MGR 模块每分钟占用一个 CPU 内核。因此，MGR 守护进程中的 CPU 使用率很高。

在这个版本中，从放置组 auto-scaler 模块进行的 OSD map 调用数量会减少。CPU 用量现在正常。

Bugzilla:2241030

确定 OSD 父（主机）的正确 CRUSH 位置

在以前的版本中，当启用 osd_memory_target_autotune 选项时，内存目标会在主机级别应用。这是在自动调整内存时使用主机掩码来完成的。但是，应用到内存目标的代码不会决定父主机的正确 CRUSH 位置，以便传播到主机的 OSD。因此，机器托管的任何 OSD 都不会被配置观察者获得通知，并且 osd_memory_target 因这些 OSD 集合而保持不变。

在这个版本中，OSD 父（主机）的正确 CRUSH 位置根据主机掩码决定。这允许更改传播到主机上的 OSD。当 auto-tuner 应用新的 osd_memory_target 且更改被反映时，由机器托管的所有 OSD 都会获得通知。

Bugzilla:2244604

在崩溃/关闭测试过程中，监控器不再处于选举状态

在以前的版本中，只有输入 stretch_mode 时，monitorMap 的 disallowed_leaders 属性才会有条件地填充。但是，有些情况下，监控被重新验证的实例不会立即进入 stretch_mode，因为它们处于概率状态。这会导致集群中的 monitor 之间 disallowed_leaders 不匹配。因此，监视器将无法选举领导机，选举将卡住，从而导致 Ceph 不响应。

在这个版本中，监控器不必处于 stretch_mode 来填充 disallowed_leaders 属性。在崩溃/关闭测试过程中，监控器不再处于选举状态。

Bugzilla:2248939

不再发生 'error getting attr on' 信息

在以前的版本中，在使用 --op list 时 ceph-objectstore-tool 会列出 pgmeta 对象，从而导致 "Error getting attr on" 信息。

在这个版本中，pgmeta 对象会被跳过，错误消息不再会出现。

Bugzilla:2251004

分配器中的 LBA 对齐不再使用，OSD 守护进程不会因为分配失败而断言

在以前的版本中，OSD 守护进程会断言，且无法重启，这有时会导致数据不可用或数据丢失。如果分配器进入 4000 请求并使用不同的分配单元配置，则 OSD 守护进程不会被视为没有问题。

在这个版本中，分配器中的 LBA 对齐不会被使用，OSD 守护进程不会因为分配失败而断言。

Bugzilla:2260306

使用 "libcephsqlite" 库的 sqlite 数据库不再可能会损坏，因为简短读取无法正确为零页。

在以前的版本中，"libcephsqlite"无法正确处理简短的读取，这可能会导致 sqlite 数据库崩溃。

在这个版本中，"libcephsqlite"零页可以正确地进行简短读取，以避免潜在的崩溃。

Bugzilla:2240139

4.9. RBD 镜像功能

现在，当对等站点在强制提升过程中停机时，镜像状态描述会显示 "orphan (force promote) "

在以前的版本中，在强制提升时，当对等站点停机时，镜像状态描述会显示 "local image linked to unknown peer"，这不是明确的描述。

在这个版本中，镜像守护进程被改进来显示镜像状态描述为 "orphan (force promote) "。

Bugzilla:2190366

rbd_support 模块不再无法从客户端重复的块列表中恢复

在以前的版本中，因为 rbd_support 模块中的递归死锁、rbd_support 模块中的竞争条件，以及 librbd cython 绑定中的 librbd cython 绑定中存在一个竞争条件，在 librbd cython 绑定中存在一个错误，有时会导致 ceph-mgr 崩溃。

在这个版本中，所有这些 3 个问题都已被修复，rbd_support' 模块不再无法从客户端重复的块列表中恢复

Bugzilla:2247531

第 5 章 已知问题

本节记录了本版本的 Red Hat Ceph Storage 中已知的问题。

5.1. CEPHADM 实用程序

在排空 OSD 时，cephadm 不会维护以前的 OSD 权重

在排空 OSD 时，cephadm 不会维护以前的 OSD 权重。因此，如果运行 ceph orch osd rm <osd-id > 命令，则 OSD 移除将停止，Cephadm 不会将 OSD 的 crush weight 设为其原始值。crush weight 将保持为 0。

作为临时解决方案，用户必须手动将 OSD 的 crush 权重调整为其原始值，或者完成移除 OSD 并部署新的 OSD。在取消 ceph orch osd rm 操作时，用户应该小心，因为在删除过程开始前 OSD 的 crush weight 不会返回到其原始值。

Bugzilla:2247211

重复使用 Ceph 对象网关 realm bootstrap 命令会导致将 zonegroup 主机名设置为失败

多次使用 Ceph 对象网关 realm bootstrap 命令设置 zonegroup 主机名会失败。因此，重复使用 Ceph 对象网关 realm bootstrap 命令重新创建 realm/zonegroup/zone 无法正常工作，zonegroup_hostnames 字段无法正常工作，主机名不会在 zonegroup 中设置。

作为临时解决方案，使用 radosgw-admin 工具手动设置 zonegroup 主机名。Bugzilla:2241321

5.2. CEPH 对象网关

处理大型 Parquet 对象的查询会导致 Ceph 对象网关进程停止

在以前的版本中，在某些情况下，在处理 Parquet 对象的查询时，该对象会在块后读取块，这些块可能非常大。这会导致 Ceph 对象网关将大型缓冲区加载到对于低端计算机而言太大的内存，特别是在 Ceph 对象网关与 OSD 进程共存时，这消耗大量内存。这种情况将触发操作系统来终止 Ceph 对象网关进程。

作为临时解决方案，将 Ceph 对象网关放在单独的节点上，从而为 Ceph 对象网关留出更多内存，使其可以成功完成处理。

Bugzilla:2275323

当前 RGW STS 实现不支持大于 1024 字节的加密密钥

当前的 RGW STS 实现不支持大于 1024 字节的加密密钥。

作为临时解决方案，在 Keycloak: realm 设置 - 密钥 中，编辑 'rsa-enc-generated' 供应商，使其具有优先级 90 而不是 100，keySize 为 1024，而不是 2048。

Bugzilla:2276931

Intel QAT Acceleration for Object Compression and Encryption

Intel QuickAssist Technology (QAT)被实施，以帮助减少节点 CPU 使用量，并在启用压缩和加密时提高 Ceph 对象网关的性能。在本发行版本中，QAT 只能在新设置(Greenfield)上配置，此功能有限制。QAT Ceph 对象网关守护进程不能在与非 QAT （常规）Ceph 对象网关守护进程相同的集群中配置。

Bugzilla:2284394

第 6 章 源

更新的 Red Hat Ceph Storage 源代码软件包位于以下位置：

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