Red Hat OpenShift Container Storage 4.8

4.8 发行注记

有关功能、功能增强、已知问题和其他重要发行信息发行注记

Last Updated: 2021-12-10

Red Hat OpenShift Container Storage 4.8 4.8 发行注记

有关功能、功能增强、已知问题和其他重要发行信息发行注记

Enter your first name here. Enter your surname here.
Enter your organisation's name here. Enter your organisational division here.
Enter your email address here.

法律通告

摘要

此发行注记介绍了 Red Hat OpenShift Container Storage 4.8 的新功能、功能增强、显著的技术变化，以及在此版本正式发行(GA)时存在的已知问题的信息。

目录

第 1 章 简介

Red Hat OpenShift Container Storage 是软件定义的存储，针对容器环境进行了优化。它在 OpenShift Container Platform 上作为操作器运行，为容器提供高度集成和简化的持久性存储管理。

Red Hat OpenShift Container Storage 集成到最新的 Red Hat OpenShift Container Platform 中，以解决平台服务、应用程序可移植性和持久性难题。它为下一代原生云应用程序提供了一个高度可扩展的后端，它建立在包括 Red Hat Ceph Storage、Rook.io Operator 和 NooBaa 的多云对象网关技术的新技术堆栈上。

Red Hat OpenShift Container Storage 提供了一个值得信任的企业级应用程序开发环境，它通过多种方式简化和增强整个应用程序生命周期的用户体验：

- 为数据库提供块存储。
- 用于持续集成、消息传递和数据聚合的共享文件存储.
- 云先开发、归档、备份和媒体存储的对象存储.
- 按指数扩展应用程序和数据.
- 以更快的速度附加和分离永久数据卷。
- 在多个数据中心或可用性区域扩展集群。
- 建立全面的应用程序容器注册表.
- 支持下一代 OpenShift 工作负载，如数据分析、Artificial Intelligence、Machine Learning、Deep Learnings(IoT)。
- 不仅动态置备应用程序容器，还有数据服务卷和容器，以及额外的 OpenShift Container Platform 节点 Elastic Block Store(EBS)卷和其他基础架构服务。

1.1. 关于此版本

Red Hat OpenShift Container Storage 4.8（RHBA-2021:3002 和 RHBA-2021:3003）现已正式发布。OpenShift Container Storage 4.8 的新增强、功能以及已知的问题包括在此文档中。

Red Hat OpenShift Container Storage 4.8 支持 Red Hat OpenShift Container Platform 版本 4.8。如需更多信息，请参阅 Red Hat OpenShift Container Storage 支持性和互操作性指南 。

随着 OpenShift Container Storage 4.8 的发布，版本 4.5 现已结束其生命周期。如需更多信息，请参阅 Red Hat OpenShift Container Platform 生命周期政策 。

第 2 章 新功能

本节论述了 Red Hat OpenShift Container Storage 4.8 中引入的新功能。

紧凑部署通用版本支持

OpenShift Container Storage 现在可以在三节点 OpenShift 紧凑裸机集群中安装，所有工作负载都在三个强大的 master 节点上运行。没有 worker 或存储节点。

有关如何在紧凑的裸机集群中配置 OpenShift Container Platform 的详情，请参考配置三节点集群 并为 Edge 部署提供三节点架构。

对象存储桶的缓存策略

在 Red Hat OpenShift Container Storage 的 Multicloud 对象网关中，现在可以创建一个缓存存储桶。缓存存储桶是带有 hub 目标和缓存目标的命名空间存储桶。如需更多信息，请参阅对象存储桶的缓存策略。

通过存储类常规可用性支持对持久性卷进行加密

您可以使用外部密钥管理系统(KMS)使用存储类加密来加密持久性卷（仅块）来存储设备加密密钥。持久卷加密仅适用于 RBD 持久卷。OpenShift Container Storage 4.7 或更高版本支持存储类加密。如需更多信息，请参阅如何使用持久性卷加密创建存储类。

持久性卷加密现在也支持快照和克隆。

密集置备 VMware 平台的存储

现在，除了 VMware 托管的 OpenShift Container Platform 的精简置备存储外，您还可以使用 thick-provisioned 存储来提高性能和安全性。当您需要在 OpenShift Container 存储中使用 thick-provisioned 存储时，您必须在 OpenShift Container Platform 中创建一个带有零精简或 enzeroedthick 磁盘格式的存储类。在创建 OpenShift Container Storage 集群服务时，您可以选择除默认的精简存储类外创建的存储类。

如需更多信息，请参阅创建 OpenShift Container Storage Cluster Service

新的池管理用户界面

新的管理功能为您提供一个简单易用的界面，以创建存储类或删除自动与之关联的池；或者，如果要更新现有池的特征（如压缩、副本），或将其删除。此功能不是现有存储类配置的替代。 *如需更多信息，请参阅《管理和分配存储资源指南》中的"块池"一章。*

支持 IBM Power 系统和 IBM Z 基础架构上的断开连接的环境

OpenShift Container Storage 4.8 现在可在 air-gapped 环境中部署，这些环境中没有互联网连接。

IBM Power 系统和 IBM Z 基础架构上的多云对象网关

IBM Power 系统和 IBM Z 基础架构上的 Red Hat OpenShift Container Storage 4.8 添加了对 Noobaa 的多云对象服务的支持，它为对象工作负载提供多云和混合功能。默认情况下，多云对象网关使用默认的后备存储，即云原生或 RGW。

IBM Power 系统上的加密存储数据

管理员现在可以选择在部署过程中加密 OpenShift Container Storage 4.8 集群中的所有数据。如需更多信息，请参阅数据加密选项。

支持 IBM Z 基础架构上的 DASD

现在，IBM Z 基础架构中的存储节点支持 DASD。

第 3 章 增强

本节论述了 Red Hat OpenShift Container Storage 4.8 中引入的主要改进。

添加了一个新警报，以便在一个或多个 OSD 请求处理时改进通知用户

此警报对于向 OpenShift Container Storage 管理员通知较慢的操作非常重要，这可能表示负载过大、存储设备缓慢或软件漏洞。用户可以检查 ceph 状态，找出造成缓慢的原因。

当 RADOS 对象网关(RGW)不可用或不健康时，会生成 ClusterObjectStoreState 警报消息。

在以前的版本中，如果 RADOS 对象网关(RGW)不可用或不健康，则不会生成 **ClusterObjectStoreState** 警报消息。在 OpenShift Container Storage operator 中实施修复后，用户现在可以在 RADOS 对象网关(RGW)不可用或不健康时看到 ClusterObjectStoreState 警报。

在池中启用或禁用压缩的功能

在 OpenShift Container Storage 4.8 之后，您可以使用用户界面在池中启用或禁用压缩作为第 2 天操作。

添加了使用 OpenShift Container Platform 用户界面创建命名空间存储桶的功能

命名空间存储桶可以使用 OpenShift Container Platform 用户界面添加。命名空间 bucket 提供内部云或 S3 兼容存储中现有对象 bucket 的聚合视图。有关使用用户界面添加命名空间存储桶的更多信息，请参阅使用 OpenShift Container Platform 用户界面添加命名空间存储桶。

在初始部署和本地存储设备的扩展过程中利用所有可用设备

对于附加模式部署中的所有本地存储设备，存储集群现在使用所有本地可用的存储设备。同样，在通过增加容量进行扩展期间，可以添加所有可用的存储设备。

如果 OSD 因节点排空以外的原因而停机，请防止在故障realm上添加 no-out 标记

当 OSD 由于磁盘失败而停机时，故障域中会添加 **no-out** 标志。这可防止 OSD 使用标准的 ceph mon_osd_down_out_interval 进行标记。在这个版本中，当 OSD 因为节点排空的原因而停机时，例如，当 pgs 不健康时，磁盘失败，则 rook 会在其他故障域中创建一个阻止的 PodDisruptionBudget 来防止进一步排空节点。在这种情况下，不会在节点上设置 **noout** 标志。如果 OSD 停机，但所有 pgs 都是 **active+clean**，则集群将被视为完全健康。默认 PodDisruptionBudget（带有 maxUnavailable=1）将被重新添加，并且阻止它们将被删除。

第 4 章 技术预览

技术预览功能的支持范围有限，如客户门户中详述的功能支持范围。

本节论述了 Red Hat OpenShift Container Storage 4.8 中在技术预览功能技术预览功能。

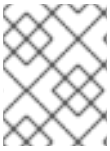
使用 abiter 进行灾难恢复

在这个版本中，Red Hat OpenShift Container Storage 提供 Metro-DR 扩展的集群(arbiter)功能，允许您在两个区域中扩展单个集群，并将第三个区作为存储集群创建过程中仲裁程序的位置。

如需更多信息，请参阅规划您的部署指南中的灾难恢复。

支持多网络插件(Multus)

支持使用多容器网络插件(multus)通过隔离网络来提高安全性和性能的功能。这个功能只在裸机和 VMWare 部署中测试。有关 multus 的更多信息，请参阅多网络插件(Multus)支持。



注意

如果删除了插件 pod，则在节点重启前无法访问数据。这是个已知问题。如需更多信息，请参阅 如果删除了插件 pod，则数据将无法访问，直到节点重启发生为止。

第 5 章 开发人员预览

本节论述了 Red Hat OpenShift Container Storage 4.7 中引入的开发人员技术预览功能。

开发人员预览功能受到开发人员预览支持的限制。开发人员预览版本不应在生产环境中运行。使用开发人员预览功能部署的集群被视为开发集群，不受红帽客户门户网站问题单管理系统的支持。如果您需要开发人员预览功能的帮助，请联络 ocs-devpreview@redhat.com 邮件列表和红帽开发团队成员将根据可用性和工作计划尽快为您提供协助。

每个主机组的数据隔离

工作负载可以使用特定的隔离 IO 路径，并分布到特定的存储节点，以限制集群部分故障或无处不在并隔离租户时的影响。

如需更多信息，请参阅知识库文章。

区域灾难恢复

Red Hat OpenShift Container Storage 在两个 OpenShift Container Storage 集群间提供多集群异步存储卷复制，为两个 Openshift Container Platform 集群提供多集群异步复制。任何有状态应用，包括其无状态的对等应用，都需要在对等群集上部署相同之前进行一些准备。

对象存储设备权重和 pod 关联性配置

现在，您可以高效地使用带有非 OpenShift Container Storage 设备的分区磁盘来最大限度地提高容量利用率，而不阻止该工作负载。您可以将对象存储设备(OSD)分配给主机计算机上的所有固态驱动器(SSD)，包括具有操作系统专用分区的分区设备。您可以减少 OSD 上的负载以提高效率，因为它共享相同的物理设备。要减少负载，您可以在 Storage Cluster CR 中配置 OSD 权重和 pod 关联性参数。

如需更多信息，请参阅知识库文章。

OpenShift Container Storage 组件部署的灵活性

现在，组件部署具有灵活性，能够在部署期间禁用组件多云对象网关、RGW 和 CephFS。也可以在 OpenShift Container Storage 部署后禁用或启用这些组件。这种灵活性有助于在使用 Amazon S3 时降低资源成本。

如需更多信息，请参阅知识库文章。

第 6 章 程序错误修复

本节论述了 Red Hat OpenShift Container Storage 4.8 中引入的显著程序错误修复。

无法同时启用仲裁器和灵活扩展。

当启用了仲裁程序和灵活的扩展时，存储集群会在 **READY** 状态下显示，即使存在错误 **仲裁器和 flexibleScaling 的日志或消息无法启用**。这是因为存储集群 CR 的 spec 不正确。在这个版本中，存储集群处于 "ERROR" 状态，并显示正确的错误消息。

(BZ#1946595)

当库需要清理时，存储桶总是被删除

在以前的版本中，在 OBC 创建失败时，lib-bucket-provisioner 在重试前会向置备程序发送删除请求以进行清理。NooBaa 置备程序会查看对象存储桶的重新声明策略，但在某些情况下不会删除底层存储桶。在这个版本中，在清理场景中，无论重新声明策略是什么，都应删除底层存储桶。

(BZ#1947796)

收集附加每个 OSD 的配置

在以前的版本中，无法找到每个 OSD 的详细配置。在这个版本中，**must-gather** 会收集 OSD 的所有配置，以进一步改进调试。

(BZ#1962755)

现在默认禁用 gRPC 指标

在以前的版本中，**cephcsi** 容器集公开远程过程调用(gRPC)指标以进行调试。**cephcsi** 节点插件容器集将主机端口 9091 用于 CephFS，9090 用于运行 **cephcsi** 节点插件容器集的节点。这意味着 **cephcsi** 容器集无法出现。在这个版本中，gRPC 指标会被默认禁用，**cephcsi** pod 在运行节点插件 Pod 的节点上不使用端口 9091 和 9090。

(BZ#1923819)

MDS 报告过大小缓存

升级时 rook 之前没有应用 mds_cache_memory_limit。这意味着没有应用该选项的 OpenShift Container Storage 4.2 集群没有使用正确的值更新，这通常是 pod 内存限值的一半。因此，备用重播中的 MDS 可能会报告过大的缓存。

(BZ#1944148)

现在，新恢复的 PVC 可以挂载到节点上

在以前的版本中，Ceph-CSI 驱动程序中的一个程序错误会导致在使用小于 8.2 的 Red Hat Enterprise Linux 版本的节点（没有深入扁平化功能）中使用删除的父快照挂载新恢复的 PVC 时出现错误。这个问题已通过在使用小于 8.2 的 Red Hat Enterprise Linux 版本的节点（没有深入扁平功能）挂载前扁平化新恢复的 PVC 来解决。

(BZ#1956232)

可靠的 mon 仲裁

在以前的版本中，如果在 mon 故障切换期间 Operator 重新启动，Operator 可能会错误地删除新的 mon。因此，当 Operator 删除新 mon 时，mon 仲裁会面临风险。在这个版本中，当 mon 故障切换正在进行时，Operator 将恢复状态，并在 Operator 重启后正确完成 mon 故障切换。现在，在节点排空和 mon 故障切换场景中，mon 仲裁更为可靠。

(BZ#1955831)

第 7 章 已知问题

本节论述了 Red Hat OpenShift Container Storage 4.8 中已知的问题。

仲裁程序节点不能使用 OpenShift Container Storage 节点标签标记

如果使用 OpenShift Container Storage 节点标签 **cluster.ocs.openshift.io/openshift-storage** 标记，仲裁程序节点将被视为有效的非仲裁节点。这意味着，非仲裁资源的放置不会确定。要临时解决这个问题，请不要使用 OpenShift Container Storage 节点标签标记仲裁节点，以便仅将仲裁资源放置在仲裁器节点上。

(BZ#1947110)

替换磁盘后，Ceph 状态为 HEALTH_WARN

替换磁盘后，即使所有 OSD pod 都已启动并在运行，也会看到 **最近出现 1 后台程序崩溃** 的警告。此警告会导致 Ceph 状态出现更改。Ceph 状态应当是 HEALTH_OK，而非 HEALTH_WARN。为解决这个问题，将 rsh 到 ceph-tools pod 并静默警告，Ceph 健康状况随后将返回到 HEALTH_OK。

(BZ#1896810)

在 CephCluster 资源中重置监控规格

每当 **ocs-operator** 重启或升级过程中，监控规格就会变为空。这不会影响功能，但如果您查找监控端点详情，您会发现它为空。

要解决这个问题，请在从 4.7 升级到 4.8 后更新 **rook-ceph-external-cluster-details** secret，以便将所有端点（如以逗号分隔的活跃和待机 MGR 的 IP 地址）的详情更新至"MonitoringEndpoint"数据键。这有助于避免因为新集群中和升级的集群中端点数量存在差异，未来出现任何问题。

(BZ#1984735)

没有obaa-db-pg-0 的问题

当托管节点停机时，NooBaa-db-pg-0 pod 不会迁移到其他节点。当节点停机时，NooBaa 无法正常工作，因为 noobaa-db-pg-0 pod 的迁移已被阻止。

(BZ#1783961)

如果删除了插件 pod，则数据将无法访问，直到节点重启发生为止

造成此问题的原因是，当 **csi-cephfsplugin** pod 重启时，挂载的 **netns** 被销毁，这会导致 in **csi-cephfs 插件锁定** 所有挂载的卷。只有使用 multus 启用的集群才会出现这个问题。

当您在删除后重启 who **csi-cephfs 插件** 的节点时会解决这个问题。

(BZ#1979561)

加密密码短语存储在源 KMS 中，以便从快照中恢复卷

当父级和恢复的 PVC 在 KMS 设置中具有不同的后端路径 StorageClass 时，恢复的 PVC 会进入 **Bound** 状态，加密密码则在快照的 KMS 设置的后端路径中创建。恢复的 PVC 无法附加到 Pod，因为检查加密密码短语时使用 2nd StorageClass 路径中链接的设置，其中加密密码短语无法在后端路径中找到。

为防止这个问题，PVC 在创建和恢复快照时应始终使用相同的 KMS 设置。

(BZ#1975730)

在使用 kv-v2 secret 引擎删除加密 PVC 后，密钥仍然会列在 Vault 中

HashiCorp Vault 为键值存储 v2 添加了一项功能，其中删除存储的密钥可让您在单独步骤中删除已删除密钥的元数据。在 Hashicorp Vault 中将键值 v2 存储用于 secret 时，删除卷不会从 KMS 中删除加密密码短语的元数据。虽然以后可以恢复加密的密码短语。KMS 不会自动清理这些部分删除的密钥。

您可以通过手动删除已删除密钥的元数据来解决这个问题。在元数据中设置了 **delete_time** 的任何键都可以假定在使用键值存储 v1 时已被删除，但可通过 v2 保持可用。

(BZ#1979244)

恢复 Snapshot/Clone 操作的大小大于父 PVC 会导致无限制循环

Ceph CSI 不支持恢复快照或创建大小大于父 PVC 的克隆。因此，恢复快照/Clone 操作的大小更大会导致无状态循环。要解决这个问题，请删除待处理的 PVC。要获得更大的 PVC，请根据您正在使用的操作完成以下操作之一： 如果使用快照，恢复现有快照以创建与父 PVC 大小相同的卷，然后将其附加到 pod 并将其扩展为所需的大小。如需更多信息，请参阅卷快照。如果使用 Clone，克隆父 PVC 以创建与父 PVC 大小相同的卷，然后将其附加到 pod，并将 PVC 扩展至所需的大小。如需更多信息，请参阅卷克隆。

(BZ#1870334)

PVC 从快照恢复，或者从密集置备的 PVC 克隆，且未置备为 thick

当使用 Sthick **provisioning** 启用的存储类**恢复 thick** 置备 PVC 快照时，恢复的卷不会被置备。恢复的 PVC 达到 **Bound** 状态，且没有 thick provisioning。这只有在使用 RHCS-5.x 时才能解决。较旧的 Ceph 版本不支持复制零填充的数据块（在密集置备时使用）。

目前，要解决基于 RHCS-4.x 的部署的问题，需要将 thick-provisioned 卷的 PVC-cloning 和 snapshot-restoring 标记为限制。新创建的卷将变为精简置备。

(BZ#1959793)

在密集置备进行时删除待处理的 PVC 和 RBD 置备程序 Pod，会保留过时的镜像和 OMAP 元数据

当 RBD PVC 被密集置备时，持久性卷声明(PVC)处于 **Pending** 状态。如果删除了 RBD 置备程序领导和 PVC 本身，RBD 镜像和 OMAP 元数据不会被删除。

要解决这个问题，在密集置备进行过程中不要删除 PVC。

(BZ#1962956)

当存储集群利用率达到 85% 甚至删除 PVC 后，置备尝试也不会停止。

如果在调配 RBD thick PVC 时存储集群利用率达到 85%，则调配尝试不会通过删除待处理的 PVC 来自动停止，而且 RBD 镜像在删除待处理的 PVC 后也不会被删除。

如果请求的大小超过可用的存储，则最好的方法是不启动置备。

(BZ#1965016)

使用 kv-v2 时，不会在卸载过程中删除 Vault 中的 OSD 的密钥

当 Vault K/V Secret 引擎是版本 2 时，**密钥加密密钥数据会在集群删除过程中从** Vault 进行软删除。这意味着可以检索任意版本的密钥，从而撤消删除。元数据仍然可见，因此可以恢复密钥。如果这会导致不便，则仍可使用 vault 命令及"destroy"参数手动删除密钥。

(BZ#1975323)

删除 CephBlockPool 会卡住，并阻止新池的创建

在 Multus 启用的集群中，Rook Operator 没有网络注解，因此无法访问 OSD 网络。这意味着，在池清理期间运行"rbd"类型命令时，命令将挂起，因为它无法联系 OSD。解决方法是使用 toolbox 手动删除 **CephBlockPool**。

(BZ#1983756)

无法通过用户界面为加密的 OpenShift Container Storage 集群执行设备替换操作

在加密的 OpenShift Container Storage 集群中，发现结果 CR 会发现由 Ceph OSD（对象存储守护进程）支持的设备与 Ceph 警报中报告的不同。单击警报时，用户会显示 Disk not found 消息。由于不匹配，console UI 无法为 OpenShift Container Storage 用户启用磁盘替换选项。要解决这个问题，请在替换设备指南中使用 CLI 流程来替换失败的设备。

(BZ#1906002)

对于将 volumeMode 作为块的 PVC 的 false Alerts

由于 Kubernetes 中的更改，OpenShift Container Platform 的 Prometheus 警报中出现了一个回归问题。这个变化会影响以下内容：

Alert: KubePersistentVolumeFillingUp.

PVC： volumeMode 中的 PVC: Block

匹配命名空间中的正则表达式："(openshift-.|kube-.|default|logging)"

metric: **kubelet_volume_stats_available_bytes**

因此，警报 kubelet_volume_stats_available_bytes 将 PVC 创建时的可用大小报告为 0，并为 volumeMode: Blocks 中的所有 PVC 触发一个假的警报：与正则表达式匹配："(openshift-.|kube-.|default|logging)"。这会影响为内部和内部附加模式以及 Amazon Web Services、VMware、Baremetal 等不同基础架构部署的 OSD 设备集创建的所有 PVC。这也会影响客户工作负载 PVC。

目前，这个问题在即将发布的 OpenShift Container Platform 4.8.z 的一个次发行版本中解决前，还没有可用的临时解决方案。因此，可以非常迅速、紧急地解决有关 OpenShift Container Storage 存储容量的任何警报。

(BZ#1984817)

关键警报通知在安装仲裁存储集群后发送，因为 ceph objectstore 的 ceph 对象用户无法在存储集群重新安装期间创建。

在包含 **CephCluster 和一个或多个 Ceph ObjectStores** 的存储集群中，如果在完全删除所有 **Ceph ObjectStore 资源前删除了 Ceph Cluster** 资源，Rook Operator 仍然可以在内存中保留 CephObjectStore(s)的连接详情。如果重新创建了相同的 **CephCluster** 和 CephObjectStore(s)，CephObjectStore(s)可能会进入 "Failed" 状态。

为避免此问题，请在删除 CephCluster 之前完全删除 CephObjectStore。

- 如果您不想等待 CephObjectStore 被删除，重启 Rook Operator（通过删除 Operator Pod）可以避免在卸载后出现问题。
- 如果您正遇到这个问题，重启 Rook Operator 将通过清除旧 **CephObjectStore** 连接详情的 Operator 内存来解决此问题。

(BZ#1974344)