



Red Hat OpenStack Platform 17.1

将 Red Hat Ceph Storage 和 Red Hat OpenStack Platform 与 director 一起部署

配置 director 以部署和使用 Red Hat Ceph Storage 集群

Red Hat OpenStack Platform 17.1 将 Red Hat Ceph Storage 和 Red Hat OpenStack Platform 与 director 一起部署

配置 director 以部署和使用 Red Hat Ceph Storage 集群

OpenStack Team
rhos-docs@redhat.com

法律通告

Copyright © 2024 Red Hat, Inc.

The text of and illustrations in this document are licensed by Red Hat under a Creative Commons Attribution–Share Alike 3.0 Unported license ("CC-BY-SA"). An explanation of CC-BY-SA is available at

<http://creativecommons.org/licenses/by-sa/3.0/>

. In accordance with CC-BY-SA, if you distribute this document or an adaptation of it, you must provide the URL for the original version.

Red Hat, as the licensor of this document, waives the right to enforce, and agrees not to assert, Section 4d of CC-BY-SA to the fullest extent permitted by applicable law.

Red Hat, Red Hat Enterprise Linux, the Shadowman logo, the Red Hat logo, JBoss, OpenShift, Fedora, the Infinity logo, and RHCE are trademarks of Red Hat, Inc., registered in the United States and other countries.

Linux[®] is the registered trademark of Linus Torvalds in the United States and other countries.

Java[®] is a registered trademark of Oracle and/or its affiliates.

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

MySQL[®] is a registered trademark of MySQL AB in the United States, the European Union and other countries.

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

The OpenStack[®] Word Mark and OpenStack logo are either registered trademarks/service marks or trademarks/service marks of the OpenStack Foundation, in the United States and other countries and are used with the OpenStack Foundation's permission. We are not affiliated with, endorsed or sponsored by the OpenStack Foundation, or the OpenStack community.

All other trademarks are the property of their respective owners.

摘要

本指南提供有关使用 Red Hat OpenStack Platform director 创建带有 Red Hat Ceph Storage 集群的 overcloud 的信息。这包括通过 director 自定义 Red Hat Ceph Storage 集群的说明。

目录

| | |
|---|-----------|
| 使开源包含更多 | 5 |
| 对红帽文档提供反馈 | 6 |
| 第 1 章 部署 OVERCLOUD 和 RED HAT CEPH STORAGE | 7 |
| 1.1. RED HAT CEPH STORAGE 集群 | 7 |
| 1.2. RED HAT CEPH STORAGE 节点和 RHEL 兼容性 | 7 |
| 1.3. RED HAT CEPH STORAGE 兼容性 | 7 |
| 1.4. 部署 RED HAT CEPH STORAGE | 7 |
| 1.5. RED HAT CEPH STORAGE 部署要求 | 8 |
| 1.6. 部署后验证 | 8 |
| 第 2 章 为部署准备 CEPH STORAGE 节点 | 10 |
| 2.1. 清理 CEPH STORAGE 节点磁盘 | 10 |
| 2.2. 注册节点 | 10 |
| 2.3. 验证可用的 RED HAT CEPH STORAGE 软件包 | 13 |
| 2.4. 为多磁盘 CEPH 集群定义根磁盘 | 13 |
| 2.5. 使用 OVERCLOUD-MINIMAL 镜像来避免使用红帽订阅授权 | 15 |
| 2.6. 为 RED HAT CEPH STORAGE 设计节点 | 17 |
| 第 3 章 配置 RED HAT CEPH STORAGE 集群 | 19 |
| 3.1. OPENSTACK OVERCLOUD CEPH DEPLOY 命令 | 19 |
| 3.2. CEPH 配置文件 | 19 |
| 3.3. 配置时间同步 | 20 |
| 3.4. 配置顶级域后缀 | 21 |
| 3.5. 配置 RED HAT CEPH STORAGE 集群名称 | 21 |
| 3.6. 使用网络数据文件配置网络选项 | 22 |
| 3.7. 使用配置文件配置网络选项 | 23 |
| 3.8. 为 OSD 配置 CRUSH 层次结构 | 24 |
| 3.9. 配置 CEPH 服务放置选项 | 25 |
| 3.10. 为 CEPH 节点配置 SSH 用户选项 | 26 |
| 3.11. 访问 CEPH STORAGE 容器 | 27 |
| 第 4 章 自定义 RED HAT CEPH STORAGE 集群 | 29 |
| 4.1. 配置选项 | 29 |
| 4.2. 生成服务规格（可选） | 29 |
| 4.3. RED HAT OPENSTACK PLATFORM WITH RED HAT CEPH STORAGE 的 CEPH 容器 | 30 |
| 4.4. 配置高级 OSD 规格 | 30 |
| 4.5. 从特定于节点的覆盖迁移 | 31 |
| 4.6. 启用 CEPH 在线加密 | 31 |
| 第 5 章 自定义存储服务 | 32 |
| 5.1. 配置自定义环境文件 | 32 |
| 5.2. RED HAT CEPH STORAGE 放置组 | 32 |
| 5.3. 启用 CEPH 元数据服务器 | 33 |
| 5.4. CEPH 对象网关对象存储 | 34 |
| 5.5. RED HAT OPENSTACK PLATFORM 对象存储的部署选项 | 35 |
| 5.6. 配置块存储备份服务以使用 CEPH | 35 |
| 5.7. 为 CEPH 节点配置多个绑定接口 | 35 |
| 第 6 章 使用原生 CEPHFS 部署共享文件系统服务 | 37 |
| 6.1. 原生 CEPHFS 带有原生驱动程序 | 37 |
| 6.2. 原生 CEPHFS 后端安全性 | 38 |

| | |
|---|-----------|
| 6.3. 原生 CEPHFS 部署 | 39 |
| 6.4. 要求 | 39 |
| 6.5. 文件共享 | 40 |
| 6.6. 原生 CEPHFS 的网络隔离 | 40 |
| 6.7. 部署原生 CEPHFS 环境 | 40 |
| 6.8. 原生 CEPHFS 后端环境文件 | 41 |
| 第 7 章 使用 CEPHFS-NFS 部署共享文件系统服务 | 43 |
| 7.1. 先决条件 | 43 |
| 7.2. CEPHFS-NFS 驱动程序 | 43 |
| 7.3. RED HAT CEPH STORAGE 服务和客户端访问 | 44 |
| 7.4. 使用 CEPHFS-NFS 容错的共享文件系统服务 | 45 |
| 7.5. CEPHFS-NFS 安装 | 45 |
| 7.6. 文件共享 | 46 |
| 7.7. CEPHFS-NFS 的网络隔离 | 46 |
| 7.8. 部署 CEPHFS-NFS 环境 | 47 |
| 7.9. CEPHFS-NFS 后端环境文件 | 48 |
| 第 8 章 启动 OVERCLOUD 部署 | 50 |
| 8.1. 启动 OVERCLOUD 部署 | 50 |
| 第 9 章 使用 DIRECTOR 为不同的工作负载定义性能层 | 52 |
| 9.1. 配置性能层 | 52 |
| 9.2. 验证 CRUSH 规则和池 | 54 |
| 第 10 章 将 RED HAT CEPH STORAGE 仪表板添加到 OVERCLOUD 部署中 | 56 |
| 10.1. 使用 CEPH 仪表板的任何位置 TLS | 57 |
| 10.2. 包含 CEPH 仪表板所需的容器 | 58 |
| 10.3. 部署 CEPH 仪表板 | 58 |
| 10.4. 使用可组合网络部署 CEPH 仪表板 | 59 |
| 10.5. 更改默认权限 | 60 |
| 10.6. 访问 CEPH 仪表板 | 61 |
| 第 11 章 管理 RED HAT CEPH STORAGE 集群的部署后操作 | 63 |
| 11.1. 禁用配置覆盖 | 63 |
| 11.2. 访问 OVERCLOUD | 63 |
| 11.3. 监控 RED HAT CEPH STORAGE 节点 | 63 |
| 11.4. 将 BLOCK STORAGE (CINDER)类型映射到您的新 CEPH 池 | 64 |
| 第 12 章 原生 CEPHFS 部署后配置和验证 | 66 |
| 12.1. 创建存储提供商网络 | 66 |
| 12.2. 配置存储提供商网络 | 67 |
| 12.3. 为存储供应商网络配置基于角色的访问控制 | 67 |
| 12.4. 配置默认共享类型 | 68 |
| 12.5. 验证隔离存储网络的创建 | 69 |
| 12.6. 验证 CEPH MDS 服务 | 69 |
| 12.7. 验证 CEPH 集群状态 | 69 |
| 12.8. 验证 MANILA-SHARE 服务状态 | 70 |
| 12.9. 验证 MANILA-API 服务确认调度程序和共享服务 | 71 |
| 第 13 章 CEPHFS NFS 部署后配置和验证 | 72 |
| 13.1. 创建存储提供商网络 | 72 |
| 13.2. 配置共享供应商 STORAGENFS 网络 | 73 |
| 13.3. 配置默认共享类型 | 73 |
| 13.4. 验证隔离存储 NFS 网络的创建 | 74 |

| | |
|---|-----------|
| 13.5. 验证 CEPH MDS 服务 | 74 |
| 13.6. 验证 CEPH 集群状态 | 75 |
| 13.7. 验证 NFS-GANESHA 和 MANILA-SHARE 服务状态 | 75 |
| 13.8. 验证 MANILA-API 服务确认调度程序和共享服务 | 76 |
| 第 14 章 重新引导环境 | 77 |
| 14.1. 重新引导 CEPH STORAGE (OSD) 集群 | 77 |
| 14.2. 重新引导 CEPH STORAGE OSD 以启用与 CEPH 监控服务的连接 | 78 |
| 第 15 章 扩展 CEPH STORAGE 集群 | 79 |
| 15.1. 扩展 CEPH STORAGE 集群 | 79 |
| 15.2. 缩减并替换 RED HAT CEPH STORAGE 节点 | 81 |
| 第 16 章 替换失败的磁盘 | 84 |
| 16.1. 替换磁盘 | 84 |

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。详情请查看 [CTO Chris Wright 的信息](#)。

对红帽文档提供反馈

我们感谢您对文档提供反馈信息。与我们分享您的成功秘诀。

在 JIRA 中提供文档反馈

使用 [Create Issue](#) 表单对文档提供反馈。JIRA 问题将在 Red Hat OpenStack Platform Jira 项目中创建，您可以在其中跟踪您的反馈进度。

1. 确保您已登录到 JIRA。如果您没有 JIRA 帐户，请创建一个帐户来提交反馈。
2. 点击以下链接打开 **Create Issue** 页面：[Create Issue](#)
3. 完成 **Summary** 和 **Description** 字段。在 **Description** 字段中，包含文档 URL、章节或章节号以及问题的详细描述。不要修改表单中的任何其他字段。
4. 点 **Create**。

第 1 章 部署 OVERCLOUD 和 RED HAT CEPH STORAGE

Red Hat OpenStack Platform (RHOSP) director 部署云环境，也称为 overcloud，以及 Red Hat Ceph Storage。director 使用通过 **tripleo-ansible** 软件包提供的 Ansible playbook 来部署 Ceph Storage 集群。director 还管理 Ceph 存储集群的配置和扩展操作。

有关 Red Hat Ceph Storage 的更多信息，请参阅 [Red Hat Ceph Storage 架构指南](#)。

如需有关 Red Hat OpenStack Platform 中服务的更多信息，请参阅使用 [CLI 工具配置基本的 overcloud](#)，[请参阅使用 director 安装和管理 Red Hat OpenStack Platform](#)。

1.1. RED HAT CEPH STORAGE 集群

Red Hat Ceph Storage 是一个分布式数据对象存储，旨在提高性能、可靠性和可扩展性。分布式对象存储使用无结构的数据来同时服务现代和旧对象接口。

Ceph Storage 部署为集群。Ceph Storage 集群由两个主要类型的守护进程组成：

- Ceph Object Storage Daemon (CephOSD)- CephOSD 执行数据存储、数据复制、重新平衡、恢复、监控和报告任务。
- Ceph Monitor (CephMon)- CephMon 使用集群的当前状态维护 cluster map 的主要副本。

有关 Red Hat Ceph Storage 的更多信息，请参阅 [Red Hat Ceph Storage 架构指南](#)。

1.2. RED HAT CEPH STORAGE 节点和 RHEL 兼容性

RHOSP 17.1 在 RHEL 9.2 上被支持。但是，映射到 Ceph Storage 角色的主机会更新到最新的主 RHEL 版本。

1.3. RED HAT CEPH STORAGE 兼容性

RHOSP 17.1 支持用于新部署的 Red Hat Ceph Storage 6。RHOSP 17.1 仅支持从 RHOSP 16.2 和 Red Hat Ceph Storage 4 升级的 Red Hat Ceph Storage 5。

1.4. 部署 RED HAT CEPH STORAGE

您可以在两个阶段部署 Red Hat Ceph Storage：

- 在部署 overcloud 之前，创建 Red Hat Ceph Storage 集群。
- 在 overcloud 部署期间配置 Red Hat Ceph Storage 集群。

创建 Ceph 存储集群可随时提供 Ceph RADOS 块设备(RBD)服务。另外，以下服务在适当的节点上运行：

- Ceph 监控器(CephMon)
- Ceph 管理器(CephMgr)
- Ceph OSD (CephOSD)

池和 cephx 密钥是在配置阶段创建的。

在配置阶段后，以下 Ceph Storage 组件将不可用：

- Ceph 控制面板(CephDashboard)
- Ceph 对象网关(CephRGW)
- Ceph MDS (CephMds)

Red Hat Ceph Storage 集群配置在 overcloud 部署期间完成。Ceph 对象网关和 Ceph 控制面板等守护进程和服务根据 overcloud 定义进行部署。Red Hat OpenStack Platform (RHOSP)服务被配置为 Ceph Storage 集群客户端。

1.5. RED HAT CEPH STORAGE 部署要求

在创建 Ceph Storage 集群之前，需要调配网络资源和裸机实例。在创建 Red Hat Ceph Storage 集群前配置以下内容：

- 使用 **openstack overcloud network provision** 命令和 **cli-overcloud-network-provision.yaml** ansible playbook 置备网络。
- 使用 **openstack overcloud node provision** 命令置备裸机实例，以使用 **cli-overcloud-node-provision.yaml** ansible playbook 置备裸机实例。

有关这些任务的更多信息，请参阅：

- [配置 Red Hat OpenStack Platform 网络](#)
- [配置裸机置备服务](#)

overcloud 环境中必须存在以下元素才能完成 Ceph 存储集群配置：

- 在 undercloud 主机上安装的 Red Hat OpenStack Platform director。请参阅使用 director 安装和管理 Red Hat OpenStack Platform 中的安装 director。
- 安装推荐的硬件来支持 Red Hat Ceph Storage。有关推荐硬件的更多信息，请参阅 Red Hat Ceph Storage 硬件指南。

1.6. 部署后验证

director 部署 Ceph Storage 集群，以使用 **cephadm** 命令执行的 **tripleo-ansible** 角色提供 Ceph RADOS 块设备(RBD)。

在 **cephadm** 完成 Ceph Storage 部署后验证以下内容：

- SSH 访问 CephMon 服务节点，以使用 **sudo cephadm shell** 命令。
- 所有 OSD 操作。



注意

检查操作 OSD 以了解环境问题，如未清理的磁盘。

- CephMon 服务节点的 **/etc/ceph** 目录中的 Ceph 配置文件和客户端管理密钥环文件。
- Ceph Storage 集群已准备好提供 RBD。

池、cephx 密钥、CephDashboard 和 CephRGW 使用 **openstack overcloud deploy** 命令在 overcloud 部署过程中配置。这是因为两个原因：

- 控制面板和 RGW 服务必须与 **haproxy** 集成。这使用 overcloud 部署。
- 池和 cephx 密钥的创建取决于部署哪些 OpenStack 客户端。

这些资源在 Ceph Storage 集群中通过 **openstack overcloud ceph deploy** 命令使用客户端管理密钥环文件和 `~/deployed_ceph.yaml` 文件输出创建。

有关 **cephadm** 的更多信息，请参阅 [Red Hat Ceph Storage 安装指南](#)。

第 2 章 为部署准备 CEPH STORAGE 节点

Red Hat Ceph Storage 节点是具有 IPMI 电源管理的裸机系统。director 会在每个节点上安装 Red Hat Enterprise Linux。

director 在内省和调配过程中通过 Provisioning 网络与每个节点通信。所有节点都通过原生 VLAN 连接到 Provisioning 网络。

有关 overcloud 部署前裸机置备的更多信息，[请参阅使用 director 安装和管理 Red Hat OpenStack Platform 指南中的置备和部署 overcloud。](#)

有关裸机部署的完整指南，[请参阅配置裸机置备服务。](#)

2.1. 清理 CEPH STORAGE 节点磁盘

Ceph Storage OSD 和日志分区需要工厂清理磁盘。在安装 Ceph OSD 服务前，所有数据和元数据都必须由这些磁盘中的 Bare Metal Provisioning 服务(ironic)清除。

您可以使用裸机置备服务将 director 配置为默认删除所有磁盘数据和元数据。当 director 配置为执行此任务时，裸机置备服务会执行额外步骤来在每次节点设置为 **可用时** 引导节点。



警告

裸机置备服务使用 **wipefs --force --all** 命令。这个命令会删除磁盘上的所有数据和元数据，但不执行安全清除。安全擦除需要更长的时间。

流程

1. 打开 **/home/stack/undercloud.conf** 并添加以下参数：

```
clean_nodes=true
```

2. 保存 **/home/stack/undercloud.conf**。
3. 更新 undercloud 配置。

```
openstack undercloud install
```

2.2. 注册节点

注册节点以启用与 director 的通信。

流程

1. 在 **/home/stack** 中创建节点清单 JSON 文件。
2. 输入每个节点的硬件和电源管理详情。
例如：

■

```
{
  "nodes":[
    {
      "mac":[
        "b1:b1:b1:b1:b1:b1"
      ],
      "cpu":"4",
      "memory":"6144",
      "disk":"40",
      "arch":"x86_64",
      "pm_type":"ipmi",
      "pm_user":"admin",
      "pm_password":"p@55w0rd!",
      "pm_addr":"192.0.2.205"
    },
    {
      "mac":[
        "b2:b2:b2:b2:b2:b2"
      ],
      "cpu":"4",
      "memory":"6144",
      "disk":"40",
      "arch":"x86_64",
      "pm_type":"ipmi",
      "pm_user":"admin",
      "pm_password":"p@55w0rd!",
      "pm_addr":"192.0.2.206"
    },
    {
      "mac":[
        "b3:b3:b3:b3:b3:b3"
      ],
      "cpu":"4",
      "memory":"6144",
      "disk":"40",
      "arch":"x86_64",
      "pm_type":"ipmi",
      "pm_user":"admin",
      "pm_password":"p@55w0rd!",
      "pm_addr":"192.0.2.207"
    },
    {
      "mac":[
        "c1:c1:c1:c1:c1:c1"
      ],
      "cpu":"4",
      "memory":"6144",
      "disk":"40",
      "arch":"x86_64",
      "pm_type":"ipmi",
      "pm_user":"admin",
      "pm_password":"p@55w0rd!",
      "pm_addr":"192.0.2.208"
    },
    {
      "mac":[
```

```
        "c2:c2:c2:c2:c2:c2"
    ],
    "cpu": "4",
    "memory": "6144",
    "disk": "40",
    "arch": "x86_64",
    "pm_type": "ipmi",
    "pm_user": "admin",
    "pm_password": "p@55w0rd!",
    "pm_addr": "192.0.2.209"
},
{
    "mac": [
        "c3:c3:c3:c3:c3:c3"
    ],
    "cpu": "4",
    "memory": "6144",
    "disk": "40",
    "arch": "x86_64",
    "pm_type": "ipmi",
    "pm_user": "admin",
    "pm_password": "p@55w0rd!",
    "pm_addr": "192.0.2.210"
},
{
    "mac": [
        "d1:d1:d1:d1:d1:d1"
    ],
    "cpu": "4",
    "memory": "6144",
    "disk": "40",
    "arch": "x86_64",
    "pm_type": "ipmi",
    "pm_user": "admin",
    "pm_password": "p@55w0rd!",
    "pm_addr": "192.0.2.211"
},
{
    "mac": [
        "d2:d2:d2:d2:d2:d2"
    ],
    "cpu": "4",
    "memory": "6144",
    "disk": "40",
    "arch": "x86_64",
    "pm_type": "ipmi",
    "pm_user": "admin",
    "pm_password": "p@55w0rd!",
    "pm_addr": "192.0.2.212"
},
{
    "mac": [
        "d3:d3:d3:d3:d3:d3"
    ],
    "cpu": "4",
    "memory": "6144",
```



```

    "disk":"40",
    "arch":"x86_64",
    "pm_type":"ipmi",
    "pm_user":"admin",
    "pm_password":"p@55w0rd!",
    "pm_addr":"192.0.2.213"
  }
]
}

```

3. 保存新文件。
4. 初始化 stack 用户：

```
$ source ~/stackrc
```

5. 将 JSON 清单文件导入到 director 并注册节点

```
$ openstack overcloud node import <inventory_file>
```

将 **<inventory_file>** 替换为第一步中创建的文件名称。

6. 将内核和 ramdisk 镜像分配给每个节点：

```
$ openstack overcloud node configure <node>
```

2.3. 验证可用的 RED HAT CEPH STORAGE 软件包

验证所有必需的软件包都可用，以避免 overcloud 部署失败。

2.3.1. 验证 cephadm 软件包安装

验证至少一个 overcloud 节点上安装 **cephadm** 软件包。**cephadm** 软件包用于引导 Ceph Storage 集群的第一个节点。

cephadm 软件包包含在 **overcloud-hardened-uefi-full.qcow2** 镜像中。**tripleo_cephadm** 角色使用 Ansible package 模块来确保镜像中存在。

2.4. 为多磁盘 CEPH 集群定义根磁盘

Ceph Storage 节点通常使用多个磁盘。director 必须在多个磁盘配置中识别根磁盘。在置备过程中，overcloud 镜像被写入根磁盘。

硬件属性用于识别根磁盘。有关可以用来识别根磁盘的属性的更多信息，请参阅 [标识根磁盘的属性](#)。

流程

1. 从每个节点的硬件内省验证磁盘信息：

```
(undercloud)$ openstack baremetal introspection data save <node_uuid> --file
<output_file_name>
```

- 将 **<node_uuid>** 替换为节点的 UUID。

- 将 `<output_file_name>` 替换为包含节点内省输出的文件的名称。
例如，一个节点的数据可能会显示 3 个磁盘：

```
[
  {
    "size": 299439751168,
    "rotational": true,
    "vendor": "DELL",
    "name": "/dev/sda",
    "wwn_vendor_extension": "0x1ea4dcc412a9632b",
    "wwn_with_extension": "0x61866da04f3807001ea4dcc412a9632b",
    "model": "PERC H330 Mini",
    "wwn": "0x61866da04f380700",
    "serial": "61866da04f3807001ea4dcc412a9632b"
  }
  {
    "size": 299439751168,
    "rotational": true,
    "vendor": "DELL",
    "name": "/dev/sdb",
    "wwn_vendor_extension": "0x1ea4e13c12e36ad6",
    "wwn_with_extension": "0x61866da04f380d001ea4e13c12e36ad6",
    "model": "PERC H330 Mini",
    "wwn": "0x61866da04f380d00",
    "serial": "61866da04f380d001ea4e13c12e36ad6"
  }
  {
    "size": 299439751168,
    "rotational": true,
    "vendor": "DELL",
    "name": "/dev/sdc",
    "wwn_vendor_extension": "0x1ea4e31e121cfb45",
    "wwn_with_extension": "0x61866da04f37fc001ea4e31e121cfb45",
    "model": "PERC H330 Mini",
    "wwn": "0x61866da04f37fc00",
    "serial": "61866da04f37fc001ea4e31e121cfb45"
  }
]
```

2. 使用唯一硬件属性为节点设置根磁盘：

```
(undercloud)$ openstack baremetal node set --property root_device='{<property_value>}'
<node-uuid>
```

- 将 `<property_value>` 替换为来自内省数据的唯一硬件属性值，以设置根磁盘。
- 将 `<node_uuid>` 替换为节点的 UUID。



注意

唯一的硬件属性是硬件内省步骤中唯一标识磁盘的任何属性。例如，以下命令使用磁盘序列号来设置根磁盘：

```
(undercloud)$ openstack baremetal node set --property
root_device='{"serial": "61866da04f380d001ea4e13c12e36ad6"}'
1a4e30da-b6dc-499d-ba87-0bd8a3819bc0
```

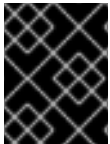
3. 将每个节点的 BIOS 配置为从网络首次引导，然后是根磁盘。

director 识别特定磁盘以用作根磁盘。运行 **openstack overcloud node provision** 命令时，director 准备 overcloud 镜像并将其写入根磁盘。

2.4.1. 标识根磁盘的属性

您可以定义多个属性以帮助 director 识别根磁盘：

- **model** (字符串)：设备识别码。
- **vendor** (字符串)：设备厂商。
- **serial** (字符串)：磁盘序列号。
- **hctl** (字符串)：SCSI 的 Host:Channel:Target:Lun。
- **size** (整数)：设备的大小 (以 GB 为单位)。
- **wwn** (字符串)：唯一的存储 ID。
- **wwn_with_extension** (字符串)：唯一存储 ID 附加厂商扩展名。
- **wwn_vendor_extension** (字符串)：唯一厂商存储标识符。
- **rotational** (布尔值)：旋转磁盘设备为 true (HDD)，否则为 false (SSD)。
- **name** (字符串)：设备名称，例如：`/dev/sdb1`。



重要

将 **name** 属性用于具有持久名称的设备。不要使用 **name** 属性为没有持久名称的设备设置根磁盘，因为该值可在节点引导时更改。

2.5. 使用 OVERCLOUD-MINIMAL 镜像来避免使用红帽订阅授权

Red Hat OpenStack Platform (RHOSP)部署的默认镜像是 **overcloud-hardened-uefi-full.qcow2**。**overcloud-hardened-uefi-full.qcow2** 镜像使用有效的 Red Hat OpenStack Platform (RHOSP)订阅。当您不希望使用您的订阅权利时，您可以使用 **overcloud-minimal** 镜像，以避免达到您付费红帽订阅的限制。例如，当您只置备带有 Ceph 守护进程的节点，或者想要调配您不想运行任何其他 OpenStack 服务的裸机操作系统(OS)时，这非常有用。有关如何获取 **overcloud-minimal** 镜像的详情，请参考 [获取 overcloud 节点的镜像](#)。



注意

overcloud-minimal 镜像仅支持标准 Linux 网桥。**overcloud-minimal** 镜像不支持 Open vSwitch (OVS)，因为 OVS 是一个需要 Red Hat OpenStack Platform 订阅权利的 OpenStack 服务。部署 Ceph Storage 节点不需要 OVS。使用 **linux_bond** 定义绑定，而不使用 **ovs_bond**。

流程

1. 打开 `/home/stack/templates/overcloud-baremetal-deploy.yaml` 文件。

2. 为您要使用 **overcloud-minimal** 镜像的节点添加或更新 **image** 属性。您可以将特定节点上的镜像设置为 **overcloud-minimal**，或将角色的所有节点设置为 **overcloud-minimal**。



注意

overcloud 最小镜像不是整个磁盘镜像。内核和 ramdisk 必须在 `/home/stack/templates/overcloud-baremetal-deploy.yaml` 文件中指定。

特定节点

```
- name: Ceph
  count: 3
  instances:
  - hostname: overcloud-ceph-0
    name: node00
    image:
      href: file:///var/lib/ironic/images/overcloud-minimal.raw
      kernel: file:///var/lib/ironic/images/overcloud-minimal.vmlinuz
      ramdisk: file:///var/lib/ironic/images/overcloud-minimal.initrd
  - hostname: overcloud-ceph-1
    name: node01
    image:
      href: file:///var/lib/ironic/images/overcloud-minimal.raw
      kernel: file:///var/lib/ironic/images/overcloud-minimal.vmlinuz
      ramdisk: file:///var/lib/ironic/images/overcloud-minimal.initrd
  - hostname: overcloud-ceph-2
    name: node02
    image:
      href: file:///var/lib/ironic/images/overcloud-minimal.raw
      kernel: file:///var/lib/ironic/images/overcloud-minimal.vmlinuz
      ramdisk: file:///var/lib/ironic/images/overcloud-minimal.initrd
```

特定角色的所有节点

```
- name: Ceph
  count: 3
  defaults:
    image:
      href: file:///var/lib/ironic/images/overcloud-minimal.raw
      kernel: file:///var/lib/ironic/images/overcloud-minimal.vmlinuz
      ramdisk: file:///var/lib/ironic/images/overcloud-minimal.initrd
  instances:
  - hostname: overcloud-ceph-0
    name: node00
  - hostname: overcloud-ceph-1
    name: node01
  - hostname: overcloud-ceph-2
    name: node02
```

3. 在 `roles_data.yaml` 角色定义文件中，将 `rhsm_enforce` 参数设置为 **False**。

```
rhsm_enforce: False
```

4. 运行 provisioning 命令：

```
(undercloud)$ openstack overcloud node provision \
--stack overcloud \
--output /home/stack/templates/overcloud-baremetal-deployed.yaml \
/home/stack/templates/overcloud-baremetal-deploy.yaml
```

5. 将 **overcloud-baremetal-deployed.yaml** 环境文件传递给 **openstack overcloud ceph deploy** 命令。

2.6. 为 RED HAT CEPH STORAGE 设计节点

要为 Red Hat Ceph Storage 指定节点，您必须创建一个新角色文件来配置 **CephStorage** 角色，并使用 **CephStorage** 的资源类配置裸机节点。

步骤

1. 以 **stack** 用户的身份登录 undercloud。
2. Source **stackrc** 文件：

```
[stack@director ~]$ source ~/stackrc
```

3. 生成一个名为 **roles_data.yaml** 的新角色数据文件，其中包含 **Controller**、**Compute** 和 **CephStorage** 角色：

```
(undercloud)$ openstack overcloud roles \
generate Controller Compute CephStorage -o /home/stack/templates/roles_data.yaml \
```

4. 打开 **roles_data.yaml**，并确保它有以下参数和部分：

| section/Parameter | 值 |
|-----------------------------------|-------------------------------------|
| role 注释 | Role: CephStorage |
| 角色名称 | 名称 : CephStorage |
| description | Ceph 节点角色 |
| HostnameFormatDefault | %stackname%-novaceph-%index% |
| deprecated_nic_config_name | ceph.yaml |

5. 将 overcloud 的 Ceph 节点添加到节点定义模板中，以注册它们。
6. 检查节点硬件：

```
(undercloud)$ openstack overcloud node introspect --all-manageable --provide
```

7. 使用自定义 Ceph 资源类标记您要为 Ceph 指定的每个裸机节点：

```
(undercloud)$ openstack baremetal node set \
--resource-class baremetal.CEPH <node>
```

将 `<node>` 替换为裸机节点的 ID。

- 将 **CephStorage** 角色添加到 `overcloud-baremetal-deploy.yaml` 文件中，并定义您要分配给节点的任何预先节点放置、资源类或其他属性：

```
- name: Controller
  count: 3
- name: Compute
  count: 3
- name: CephStorage
  count: 5
  defaults:
    resource_class: baremetal.CEPH
```

- 运行 provisioning 命令：

```
(undercloud)$ openstack overcloud node provision \
--stack stack \
--output /home/stack/templates/overcloud-baremetal-deployed.yaml \
/home/stack/templates/overcloud-baremetal-deploy.yaml
```

- 在一个单独的终端中监控置备进度。当置备成功时，节点状态将从 **available** 变为 **active**：

```
(undercloud)$ watch openstack baremetal node list
```

其他资源

- 有关节点注册的更多信息，请参阅 [第 2.2 节“注册节点”](#)。
- 有关检查节点硬件的更多信息，请参阅 [安装和管理 Red Hat OpenStack Platform 指南](#) 中的 [创建裸机节点硬件清单](#)。

第 3 章 配置 RED HAT CEPH STORAGE 集群

要为 Red Hat OpenStack Platform 环境部署 Red Hat Ceph Storage 集群，您必须首先为您的环境配置 Red Hat Ceph Storage 集群选项。

配置 Red Hat Ceph Storage 集群选项：

- [配置时间同步](#)
- [配置顶级域](#)
- [配置 Red Hat Ceph Storage 集群名称](#)
- [使用网络数据文件配置网络选项](#)
- [使用配置文件配置网络选项](#)
- [为 OSD 配置 CRUSH 层次结构](#)
- [配置 Ceph 服务放置选项](#)
- [为 Ceph 节点配置 SSH 用户选项](#)
- [配置容器 registry](#)

先决条件

在配置和部署 Red Hat Ceph Storage 集群前，请使用裸机置备服务(ironic)来置备裸机实例和网络。如需更多信息，请[参阅配置裸机置备服务](#)。

3.1. OPENSTACK OVERCLOUD CEPH DEPLOY 命令

如果使用 director 部署 Ceph 集群，则必须使用 **openstack overcloud ceph deploy** 命令。有关命令选项和参数的完整列表，请[参阅命令行界面参考中的 openstack overcloud ceph deploy](#)。

openstack overcloud ceph deploy --help 命令提供环境中当前可用的选项和参数。

3.2. CEPH 配置文件

标准格式初始化文件是执行 Ceph 集群配置的一种方法。此初始化文件用于配置 Ceph 集群。使用以下命令之一使用此文件：*** cephadm bootstrap --config <file_name> * openstack overcloud ceph deploy --config <file_name>** 命令。

示例

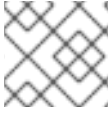
以下示例创建一个名为 **initial-ceph.conf** 的简单初始化文件，然后使用 **openstack overcloud ceph deploy** 命令来为其配置 Ceph 集群。它演示了如何将 messenger v2 协议配置为使用加密通过网络传递的所有数据的安全模式。

```
$ cat <<EOF > initial-ceph.conf
[global]
ms_cluster_mode = secure
ms_service_mode = secure
ms_client_mode = secure
EOF
$ openstack overcloud ceph deploy --config initial-ceph.conf ...
```

3.3. 配置时间同步

时间同步服务(chrony)会被默认启用进行时间同步。您可以执行以下任务来配置服务。

- [使用分隔列表配置时间同步](#)
- [使用环境文件配置时间同步](#)
- [禁用时间同步](#)



注意

使用分隔列表或环境文件配置时间同步。使用最适合您的管理实践的流程。

3.3.1. 使用分隔列表配置时间同步

您可以将时间同步服务(chrony)配置为使用分隔的列表来配置 NTP 服务器。

流程

1. 以 **stack** 用户身份登录 undercloud 节点。
2. 使用分隔列表配置 NTP 服务器：

```
openstack overcloud ceph deploy \
  --ntp-server "<ntp_server_list>"
```

将 **<ntp_server_list >** 替换为以逗号分隔的服务器列表。

```
openstack overcloud ceph deploy \
  --ntp-server "0.pool.ntp.org,1.pool.ntp.org"
```

3.3.2. 使用环境文件配置时间同步

您可以配置时间同步服务(chrony)，以使用定义 NTP 服务器的环境文件。

流程

1. 以 **stack** 用户身份登录 undercloud 节点。
2. 创建一个环境文件，如 **/home/stack/templates/ntp-parameters.yaml**，使其包含 NTP 服务器配置。
3. 添加 **NtpServer** 参数。**NtpServer** 参数包含以逗号分隔的 NTP 服务器列表。

```
parameter_defaults:
  NtpServer: 0.pool.ntp.org,1.pool.ntp.org
```

4. 使用环境文件配置 NTP 服务器：

```
openstack overcloud ceph deploy \
  --ntp-heat-env-file "<ntp_file_name>"
```


将 `<ntp_file_name>` 替换为您创建的环境文件的名称。

```
openstack overcloud ceph deploy \
  --ntp-heat-env-file "/home/stack/templates/ntp-parameters.yaml"
```

3.3.3. 禁用时间同步

时间同步服务(chrony)默认启用。如果您不想使用该服务，您可以禁用该服务。

流程

1. 以 **stack** 用户身份登录 undercloud 节点。
2. 禁用时间同步服务(chrony)：

```
openstack overcloud ceph deploy \
  --skip-ntp
```

3.4. 配置顶级域后缀

您可以配置顶级域(TLD)后缀。此后缀添加到短主机名中，以便为 overcloud 节点创建完全限定域名。



注意

TLS-e 配置需要一个完全限定域名。

流程

1. 以 **stack** 用户身份登录 undercloud 节点。
2. 配置顶级域后缀：

```
openstack overcloud ceph deploy \
  --tld "<domain_name>"
```

将 `<domain_name>` 替换为所需的域名。

```
openstack overcloud ceph deploy \
  --tld "example.local"
```

3.5. 配置 RED HAT CEPH STORAGE 集群名称

您可以使用您配置的名称部署 Red Hat Ceph Storage 集群。默认名称为 **ceph**。

流程

1. 以 **stack** 用户身份登录 undercloud 节点。
2. 使用以下命令配置 Ceph Storage 集群的名称：
openstack overcloud ceph deploy \ --cluster <cluster_name>

```
$ OpenStack overcloud ceph deploy \ --cluster central \
```



注意

目前不会创建密钥环文件。密钥环文件是在 overcloud 部署期间创建的。密钥环文件继承此流程中配置的集群名称。有关 overcloud 部署的更多信息，请参阅 [第 8.1 节“启动 overcloud 部署”](#)

在上例中，Ceph 集群名为 **central**。在部署过程中，将在 **/etc/ceph** 中创建 **中央 Ceph 集群** 的配置和密钥环文件。

```
[root@oc0-controller-0 ~]# ls -l /etc/ceph/
total 16
-rw-----. 1 root root 63 Mar 26 21:49 central.client.admin.keyring
-rw-----. 1 167 167 201 Mar 26 22:17 central.client.openstack.keyring
-rw-----. 1 167 167 134 Mar 26 22:17 central.client.radosgw.keyring
-rw-r--r--. 1 root root 177 Mar 26 21:49 central.conf
```

故障排除

如果您为 Ceph Storage 集群配置自定义名称，可能会显示以下错误：

monclient: get_monmap_and_config 无法识别要联系的监控器，因为

如果显示此错误，请在 Ceph 部署后使用以下命令：

cephadm shell --config <configuration_file> --keyring <keyring_file>

例如，如果在将集群名称配置为 **中央** 时显示这个错误，您可以使用以下命令：

```
cephadm shell --config /etc/ceph/central.conf \
  --keyring /etc/ceph/central.client.admin.keyring
```

以下命令也可以用作替代方案：

```
cephadm shell --mount /etc/ceph:/etc/ceph
export CEPH_ARGS='--cluster central'
```

3.6. 使用网络数据文件配置网络选项

网络数据文件描述了 Red Hat Ceph Storage 集群使用的网络。

流程

1. 以 **stack** 用户身份登录 undercloud 节点。
2. 创建一个 YAML 格式文件，用于定义名为 **network_data.yaml** 的自定义网络属性。



重要

使用网络隔离，标准网络部署由两个存储网络组成，它们映射到两个 Ceph 网络：

- 存储网络 **storage** 映射到 Ceph network, 网络 **public_network**。此网络处理存储流量，如从 Compute 节点到 Ceph 集群的 RBD 流量。
- 存储网络 **storage_mgmt** 映射到 Ceph 网络 **cluster_network**。此网络处理存储管理流量，如 Ceph OSD 之间的数据复制。

3. 使用 **openstack overcloud ceph deploy** 命令和 **--crush-hierarchy** 选项来部署配置。

```
openstack overcloud ceph deploy \
  deployed_metal.yaml \
  -o deployed_ceph.yaml \
  --network-data network_data.yaml
```



重要

openstack overcloud ceph deploy 命令使用 **--network-data** 选项指定的网络数据文件来确定要用作 **public_network** 和 **cluster_network** 的网络。该命令假设这些网络在网络数据文件中名为 **storage** 和 **storage_mgmt**，除非由 **--public-network-name** 和 **--cluster-network-name** 选项指定不同的名称。

在使用网络隔离部署时，您必须使用 **--network-data** 选项。如果没有使用这个选项，默认的 undercloud (192.168.24.0/24) 将同时用于 **public_network** 和 **cluster_network**。

3.7. 使用配置文件配置网络选项

可以使用配置文件来指定网络选项，作为网络数据文件的替代选择。



重要

使用此方法配置网络选项覆盖 **network_data.yaml** 中自动生成的值。在使用这种网络配置方法时，请确保设置了所有四个值。

流程

1. 以 **stack** 用户身份登录 undercloud 节点。
2. 创建标准格式初始化文件来配置 Ceph 集群。如果您已经创建了一个文件来包含其他配置选项，您可以将网络配置添加到其中。
3. 在文件的 **[global]** 部分添加以下参数：
 - **public_network**
 - **cluster_network**
 - **ms_bind_ipv4**



重要

确保 **public_network** 和 **cluster_network** 映射到与 **storage** 和 **storage_mgmt** 相同的网络。

以下是具有多个子网和自定义网络名称的网络配置的配置文件条目示例：

```
[global]
public_network = 172.16.14.0/24,172.16.15.0/24
cluster_network = 172.16.12.0/24,172.16.13.0/24
ms_bind_ipv4 = True
ms_bind_ipv6 = False
```

4. 使用命令 **openstack overcloud ceph deploy** 和 **--config** 选项来部署配置文件。

```
$ openstack overcloud ceph deploy \
--config initial-ceph.conf --network-data network_data.yaml
```

3.8. 为 OSD 配置 CRUSH 层次结构

您可以在 OSD 部署期间配置自定义受控复制(CRUSH)层次结构，将 OSD **location** 属性添加到 Ceph Storage 集群主机规格中。**location** 属性配置 OSD 放置到 CRUSH 层次结构中的位置。



注意

location 属性仅设置初始 CRUSH 位置。属性的后续更改将被忽略。

流程

1. 以 **stack** 用户身份登录 undercloud 节点。
2. 查找 **stackrc** undercloud 凭证文件：
\$ source ~/stackrc
3. 创建配置文件以定义自定义 CRUSH 层次结构，如 **crush_hierarchy.yaml**。
4. 在文件中添加以下配置：

```
<osd_host>:
  root: default
  rack: <rack_num>
<osd_host>:
  root: default
  rack: <rack_num>
<osd_host>:
  root: default
  rack: <rack_num>
```

- 将 **<osd_host>** 替换为部署 OSD 的节点的主机名，如 **ceph-0**。
 - 将 **<rack_num>** 替换为部署 OSD 的机架数量，如 **r0**。
5. 使用自定义 OSD 布局部署 Ceph 集群：



```
openstack overcloud ceph deploy \
  deployed_metal.yaml \
  -o deployed_ceph.yaml \
  --osd-spec osd_spec.yaml \
  --crush-hierarchy crush_hierarchy.yaml
```

Ceph 集群是使用自定义 OSD 布局创建的。

上面的示例文件将导致以下 OSD 布局。

| ID | CLASS | WEIGHT | TYPE | NAME | STATUS | REWEIGHT | PRI-AFF |
|----|-------|---------|------|--------------|--------|----------|---------|
| -1 | | 0.02939 | root | default | | | |
| -3 | | 0.00980 | rack | r0 | | | |
| -2 | | 0.00980 | host | ceph-node-00 | | | |
| 0 | hdd | 0.00980 | osd | osd.0 | up | 1.00000 | 1.00000 |
| -5 | | 0.00980 | rack | r1 | | | |
| -4 | | 0.00980 | host | ceph-node-01 | | | |
| 1 | hdd | 0.00980 | osd | osd.1 | up | 1.00000 | 1.00000 |
| -7 | | 0.00980 | rack | r2 | | | |
| -6 | | 0.00980 | host | ceph-node-02 | | | |
| 2 | hdd | 0.00980 | osd | osd.2 | up | 1.00000 | 1.00000 |



注意

Ceph 自动检测设备类，但 CRUSH 规则与池相关联。在 overcloud 部署期间，仍然使用 **CephCrushRules** 参数定义和创建池。

其他资源

如需更多信息，请参阅 *Red Hat Ceph Storage 安装指南* 中的 [Red Hat Ceph Storage 工作负载注意事项](#)。

3.9. 配置 CEPH 服务放置选项

您可以使用自定义角色文件定义哪些节点运行哪些 Ceph 服务。只有因为环境而不使用默认角色分配时才需要自定义角色文件。例如，部署超融合节点时，预部署计算节点应标记为 **osd**，服务类型为 **osd**，具有包含计算实例列表的放置列表。

roles_data.yaml 文件中的服务定义决定了哪个裸机实例运行哪个服务。默认情况下，Controller 角色具有 CephMon 和 CephMgr 服务，而 CephStorage 角色具有 CephOSD 服务。与大多数可组合服务不同，Ceph 服务不需要 heat 输出来确定如何配置服务。**roles_data.yaml** 文件始终决定 Ceph 服务放置，即使 Heat 运行前发生了部署的 Ceph 进程。

流程

1. 以 **stack** 用户身份登录 undercloud 节点。
2. 创建定义自定义角色的 YAML 格式文件。
3. 部署配置文件：

```
$ openstack overcloud ceph deploy \
  deployed_metal.yaml \
  -o deployed_ceph.yaml \
  --roles-data custom_roles.yaml
```

3.10. 为 CEPH 节点配置 SSH 用户选项

`openstack overcloud ceph deploy` 命令创建用户和密钥，并将其分发到主机，因此不需要执行本节中的流程。但是，它是一个受支持的选项。

Cephadm 使用 SSH 连接到所有受管的远程 Ceph 节点。Red Hat Ceph Storage 集群部署过程在所有 overcloud Ceph 节点上创建一个帐户和 SSH 密钥对。然后，为 Cephadm 提供密钥对，以便它可以与节点通信。

3.10.1. 在 Red Hat Ceph Storage 集群创建前创建 SSH 用户

您可以使用 `openstack overcloud ceph user enable` 命令在 Ceph 集群创建前创建 SSH 用户。

流程

1. 以 **stack** 用户身份登录 undercloud 节点。
2. 创建 SSH 用户：

```
$ OpenStack overcloud ceph user enable <specification_file>
```

- 将 `<specification_file>` 替换为 Ceph 规格文件的路径和名称，该文件描述了创建用户的集群，并安装公共 SSH 密钥。规范文件提供了用于确定要修改哪些节点以及是否需要私钥的信息。有关创建规格文件的更多信息，请参阅[生成服务规格](#)。



注意

默认用户名是 **ceph-admin**。要指定不同的用户名，请使用 `--cephadm-ssh-user` 选项指定不同的用户名。

```
OpenStack overcloud ceph user enable --cephadm-ssh-user
<custom_user_name>
```

建议您使用默认名称，而不使用 `--cephadm-ssh-user` 参数。

如果事先创建用户，请在执行 `openstack overcloud ceph deploy` 时使用参数 `--skip-user-create`。

3.10.2. 禁用 SSH 用户

禁用 SSH 用户会禁用 **cephadm**。禁用 **cephadm** 会删除服务管理 Ceph 集群的功能，并防止关联的命令正常工作。它还可防止 Ceph 节点 overcloud 扩展操作。它还会删除所有公钥和 SSH 私钥。

流程

1. 以 **stack** 用户身份登录 undercloud 节点。
2. 使用命令 `openstack overcloud ceph user disable --fsid <FSID> <specification_file>` 来禁用 SSH 用户。
 - 将 `<FSID>` 替换为集群的文件系统 ID。FSID 是集群的唯一标识符。FSID 位于 `deployed_ceph.yaml` 环境文件中。
 - 将 `<specification_file>` 替换为描述用户创建集群的 Ceph 规格文件的路径和名称。

**重要**

不建议 `openstack overcloud ceph user disable` 命令，除非需要禁用 `cephadm`。

**重要**

要在禁用后启用 SSH 用户和 Ceph 编配器服务，请使用 `openstack overcloud ceph user enable --fsid <FSID> <specification_file>` 命令。

**注意**

此命令需要 Ceph 规格文件的路径来确定：

- 哪些主机需要 SSH 用户。
- 哪些主机具有 `_admin` 标签，并且需要 SSH 私钥。
- 哪些主机需要 SSH 公钥。

有关规格文件以及如何生成它们的更多信息，请参阅生成服务规格。

3.11. 访问 CEPH STORAGE 容器

安装和管理 Red Hat OpenStack Platform 指南中的 [准备容器镜像](#)，其中包含有关如何准备 registry 和 undercloud 和 overcloud 配置以使用容器镜像的步骤和信息。使用本节中的信息，调整这些步骤来访问 Ceph Storage 容器。

从 overcloud 访问 Ceph Storage 容器有两个选项。

- [直接从远程 registry 下载容器](#)
- [在 undercloud 上缓存容器](#)

3.11.1. 直接从远程 registry 下载容器

您可以将 Ceph 配置为直接从远程 registry 下载容器。

`cephadm` 命令使用 `containers-prepare-parameter.yaml` 文件中配置的凭证来对远程 registry 进行身份验证并下载 Red Hat Ceph Storage 容器。

流程

1. 使用 [安装和管理 Red Hat OpenStack Platform 指南中的 `Preparing container -prepare-parameter.yaml` 文件](#) 创建 `containers-prepare-parameter.yaml` 文件。
2. 使用 `ContainerImageRegistryCredentials` 参数将远程 registry 凭证添加到 `containers-prepare-parameter.yaml` 文件中，如 [从私有 registry 获取容器镜像](#) 中所述。
3. 部署 Ceph 时，使用 `openstack overcloud ceph deploy` 命令传递 `containers-prepare-parameter.yaml` 文件。

```
openstack overcloud ceph deploy \
  --container-image-prepare containers-prepare-parameter.yaml
```



注意

如果您没有将 undercloud 上的容器缓存，如 undercloud 上的缓存容器 中所述，那么在部署 Ceph 时，您应该将相同的 **containers-prepare-parameter.yaml** 文件传递给 **openstack overcloud ceph deploy** 命令。这将在 undercloud 上缓存容器。

3.11.2. 在 undercloud 上缓存容器

在准备过程中修改镜像 的流程描述了使用以下命令：

```
sudo openstack tripleo container image prepare \  
-e ~/containers-prepare-parameter.yaml \  

```

如果您不使用 **--container-image-prepare** 选项为 **openstack overcloud ceph deploy** 命令提供身份验证凭据，并直接从远程 registry 下载 Ceph 容器，如部署 Ceph 之前，您必须运行 **sudo openstack tripleo container image prepare** 命令。

第 4 章 自定义 RED HAT CEPH STORAGE 集群

director 使用默认配置部署 Red Hat Ceph Storage。您可以自定义此默认配置。

先决条件

- 配置了存储网络的 Ceph Storage 节点。
- `openstack overcloud node provision -o ~/deployed_metal.yaml ...`

4.1. 配置选项

配置 Red Hat Ceph Storage 集群有几个选项。

流程

1. 以 **stack** 用户身份登录 undercloud 节点。
2. 可选：使用标准格式初始化(ini)文件来配置 Ceph 集群。

- a. 使用配置选项创建文件。
以下是简单配置文件的示例：

```
[global]
osd_crush_chooseleaf type = 0
log_file = /var/log/ceph/$cluster-$type.$id.log

[mon]
mon_cluster_log_to_syslog = true
```

- b. 保存配置文件。
- c. 使用 `openstack overcloud ceph deploy --config <configuration_file_name>` 命令来部署配置。
将 `<configuration_file_name>` 替换为您创建的文件名称。

```
$ openstack overcloud ceph deploy --config initial-ceph.conf
```

3. 可选：向 `cephadm bootstrap` 命令发送配置值：
`openstack overcloud ceph deploy --force \ --cephadm-extra-args '<optional_arguments>'`
将 `<optional_arguments>` 替换为提供给底层命令的配置值。



注意

在使用参数 `--log-to-file` 和 `--skip-prepare-host` 时，使用 `openstack overcloud ceph deploy --force \ --cephadm-extra-args '--log-to-file --skip-prepare-host'`

4.2. 生成服务规格（可选）

Red Hat Ceph Storage 集群服务规格是一个 YAML 文件，用于描述 Ceph Storage 服务的部署。在部署 Ceph Storage 集群前，它由 **tripleo** 自动生成。它通常不必单独生成。

可以创建自定义服务规格来自定义 Red Hat Ceph Storage 集群。

流程

1. 以 **stack** 用户身份登录 undercloud 节点。

2. 生成规格文件：

```
openstack overcloud ceph spec deployed_metal.yaml -o <specification_file>
```

- 将 **<specification_file>** 替换为要使用当前服务规格生成的文件的名称。



注意

deployed_metal.yaml 来自 **openstack overcloud node provision** 命令的输出。

3. 使用所需的配置编辑生成的文件。

4. 部署自定义服务规格：

```
openstack overcloud ceph deploy \
  deployed_metal.yaml \
  -o deployed_ceph.yaml \
  --ceph-spec <specification_file>
```

- 将 **<specification_file>** 替换为自定义服务规格文件的名称。

4.3. RED HAT OPENSTACK PLATFORM WITH RED HAT CEPH STORAGE 的 CEPH 容器

您必须具有 Ceph Storage 容器，才能将 Red Hat Openstack Platform (RHOSP) 配置为使用带有 NFS Ganesha 的 Red Hat Ceph Storage。如果外部 Ceph 存储集群仅提供块（通过 RBD）、对象（通过 RGW）或文件（通过原生 CephFS）存储，则不需要 Ceph Storage 容器。

RHOSP 17.1 将部署 Red Hat Ceph Storage 6.x (Ceph 软件包 17.x)。Ceph Storage 6.x 容器托管在 registry.redhat.io 上，这是需要身份验证的 registry。如需更多信息，请参阅 [容器镜像准备参数](#)。

4.4. 配置高级 OSD 规格

当默认规格不为您的 Ceph Storage 集群提供必要功能时，配置高级 OSD 规格。

流程

1. 以 **stack** 用户身份登录 undercloud 节点。

2. 创建定义高级 OSD 规格的 YAML 格式文件。

以下是自定义 OSD 规范的示例。

```
data_devices:
  rotational: 1
db_devices:
  rotational: 0
```

本例将创建一个 OSD 规格，所有轮转设备将是数据设备，并且所有非轮转设备都将用作共享设备。构建动态 Ceph 服务规格时，如果 `service_type` 是 `osd`，则规格文件中的任何内容都会附加到规格的部分中。

3. 保存规格文件。

4. 部署规格：

```
OpenStack overcloud ceph deploy \ --osd-spec <osd_specification_file>
```

将 `<osd_specification_file>` 替换为您创建的规格文件的名称。

```
$ openstack overcloud ceph deploy \ --osd-spec osd_spec.yaml \
```

其他资源

有关在服务规格中配置 OSD 相关属性的列表，请参阅 *Red Hat Ceph Storage Operations 指南* 中的 [部署 OSD 的高级服务规格和过滤器](#)。

4.5. 从特定于节点的覆盖迁移

在 Red Hat OpenStack Platform 17.0 之前，使用特定于节点的覆盖来管理非同构服务器硬件。现在，这是使用自定义 OSD 规格文件完成的。如需有关如何创建自定义 [OSD 规格文件的信息](#)，请参阅 [配置高级 OSD 规格](#)。

4.6. 启用 CEPH 在线加密

使用 messenger 版本 2 协议的 [安全模式](#)，为所有 Ceph Storage 流量启用加密。配置 Ceph 存储，如 *Red Hat Ceph Storage Data Hardening Red Hat OpenStack Platform* 中的 [Encryption and Key Management](#) 所述，以启用 Ceph 在线加密。

其他资源

有关 Ceph 进行线加密的更多信息，请参阅 [Red Hat Ceph Storage 架构指南](#) 中的 [Ceph 在线加密](#)。

第 5 章 自定义存储服务

director heat 模板集合包含必要的模板和环境文件，以启用基本的 Ceph Storage 配置。

director 使用 `/usr/share/openstack-tripleo-heat-templates/environments/cephadm/cephadm.yaml` 环境文件将配置添加到 **openstack overcloud ceph deploy** 部署的 Ceph Storage 集群中，并在部署期间将其与 overcloud 集成。

5.1. 配置自定义环境文件

director 将基本的默认设置应用到部署的 Red Hat Ceph Storage 集群。您必须在自定义环境文件中定义其他配置。

流程

1. 以 **stack** 用户身份登录 undercloud。
2. 创建用于定义自定义配置的文件。
vi /home/stack/templates/storage-config.yaml
3. 在文件中添加 **parameter_defaults** 部分。
4. 添加自定义配置参数。有关参数定义的更多信息，请参阅 [Overcloud 参数](#)。

```
parameter_defaults:
  CinderEnableScsiBackend: false
  CinderEnableRbdBackend: true
  CinderBackupBackend: ceph
  NovaEnableRbdBackend: true
  GlanceBackend: rbd
```



注意

自定义配置文件中定义的参数覆盖 `/usr/share/openstack-tripleo-heat-templates/environments/cephadm/cephadm.yaml` 中任何对应的默认设置。

5. 保存该文件。

其他资源

自定义配置在 overcloud 部署期间应用。

5.2. RED HAT CEPH STORAGE 放置组

放置组(PG)大规模促进动态高效的对象跟踪。如果 OSD 失败或 Ceph Storage 集群重新平衡，Ceph 可以移动或复制放置组的内容。这使得 Ceph 存储集群能够有效地重新平衡和恢复。

除非在 Ceph 配置文件中包含以下参数，否则放置组和副本数设置不会从默认值中更改：

- **osd_pool_default_size**
- **osd_pool_default_pg_num**
- **osd_pool_default_pgp_num**

使用 **openstack overcloud deploy** 命令部署 overcloud 时，会为每个启用的 Red Hat OpenStack Platform 服务创建一个池。例如，以下命令为计算服务(nova)、块存储服务(cinder)和 Image 服务(glance)创建池：

```
openstack overcloud deploy --templates \
  -e /usr/share/openstack-tripleo-heat-templates/environments/cephadm/cephadm-rbd-only.yaml
```

将 **-e environments/cinder-backup.yaml** 添加到命令中，创建一个名为 **backups** 的池：

```
openstack overcloud deploy --templates \
  -e /usr/share/openstack-tripleo-heat-templates/environments/cephadm/cephadm-rbd-only.yaml
  -e environments/cinder-backup.yaml
```

不需要为每个池配置放置组号；默认为启用 **pg_autoscale_mode** 属性。但是，建议您配置 **target_size_ratio** 或 **pg_num** 属性。这个减弱数据重新平衡。

要为每个池设置 **target_size_ratio** 属性，请使用类似以下示例的配置文件条目：

```
parameter_defaults:
  CephPools:
    - name: volumes
      target_size_ratio: 0.4
      application: rbd
    - name: images
      target_size_ratio: 0.1
      application: rbd
    - name: vms
      target_size_ratio: 0.3
      application: rbd
```

在本例中，每个服务使用的数据百分比将是：

- Cinder 卷 - 40%
- Glance 镜像 - 10%
- Nova vms - 30%
- 其他池的可用空间 - 20%

根据您的预期使用情况设置这些值。如果不覆盖 **CephPools** 参数，则每个池使用默认的放置组号。虽然自动扩展将根据使用情况自动调整此数量，但数据将在 Ceph 集群中移动。这使用计算资源。

如果您希望设置放置组数量而不是目标大小比率，请将示例中的 **target_size_ratio** 替换为 **pg_num**。根据您的预期使用情况，为每个池使用不同的整数。

请参阅 [Red Hat Ceph Storage 处理器、网络接口卡和电源管理接口的 Red Hat Ceph Storage 硬件指南](#)。

5.3. 启用 CEPH 元数据服务器

Ceph 元数据服务器(MDS)运行 **ceph-mds** 守护进程。此守护进程管理与 CephFS 中存储的文件相关的元数据。CephFS 可以原生使用，或者通过 NFS 协议使用。



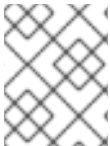
注意

红帽支持使用原生 CephFS 和 CephFS NFS 后端部署 Ceph MDS，用于共享文件系统服务 (manila)。

流程

- 要启用 Ceph MDS，请在部署 overcloud 时使用以下环境文件：

```
/usr/share/openstack-tripleo-heat-templates/environments/cephadm/ceph-mds.yaml
```



注意

默认情况下，Ceph MDS 部署在 Controller 节点上。您可以在自己的专用节点上部署 Ceph MDS。

其他资源

- [Red Hat Ceph Storage 文件系统指南](#)

5.4. CEPH 对象网关对象存储

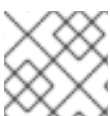
Ceph 对象网关(RGW)提供了一个接口，用于访问 Red Hat Ceph Storage 集群中的对象存储功能。

使用 director 部署 Ceph 时，director 会自动启用 RGW。这是 Object Storage 服务(swift)的直接替换。通常使用对象存储服务的服务可以使用 RGW，而无需额外的配置。Object Storage 服务作为升级的 Ceph 集群的对象存储选项保留可用。

不需要单独的 RGW 环境文件来启用它。有关其他对象存储选项的环境文件的更多信息，请参阅 [第 5.5 节 “Red Hat OpenStack Platform 对象存储的部署选项”](#)。

默认情况下，Ceph Storage 允许每个 Object Storage Daemon (OSD)有 250 个放置组。启用 RGW 时，Ceph Storage 会创建 RGW 所需的以下六个额外池：

- **.rgw.root**
- **<zone_name>.rgw.control**
- **<zone_name>.rgw.meta**
- **<zone_name>.rgw.log**
- **<zone_name>.rgw.buckets.index**
- **<zone_name>.rgw.buckets.data**



注意

在部署中，**<zone_name>** 替换为池所属的区的名称。

其他资源

- 有关 RGW 的更多信息，请参阅 Red Hat Ceph Storage [对象网关指南](#)。
- 有关使用 RGW 而不是 Swift 的更多信息，请参阅 [备份 BLock Storage 卷 指南](#)。

5.5. RED HAT OPENSTACK PLATFORM 对象存储的部署选项

部署 overcloud 对象存储有三个选项：

- Ceph 对象网关(RGW)
要按照 第 5.4 节 “Ceph 对象网关对象存储” 所述部署 RGW，请在 overcloud 部署期间包含以下环境文件：

```
-e environments/cephadm/cephadm.yaml
```

此环境文件同时配置 Ceph 块存储(RBD)和 RGW。

- Object Storage 服务 (swift)
要部署 Object Storage 服务(swift)而不是 RGW，请在 overcloud 部署期间包含以下环境文件：

```
-e environments/cephadm/cephadm-rbd-only.yaml
```

cephadm-rbd-only.yaml 文件配置 Ceph RBD，但不配置 RGW。



注意

如果在升级 Red Hat Ceph Storage 集群前使用 Object Storage 服务(swift)，您可以在升级过程中继续使用 Object Storage 服务(swift)而不是 RGW，将 **environments/ceph-ansible/ceph-ansible.yaml** 文件替换为 **environments/cephadm/cephadm-rbd-only.yaml**。如需更多信息，请参阅 [执行 Red Hat OpenStack Platform 的次要更新](#)。

Red Hat OpenStack Platform 不支持从 Object Storage 服务(swift)迁移到 Ceph 对象网关(RGW)。

- 没有对象存储
要使用 RBD 部署 Ceph，而不使用 RGW 或 Object Storage 服务(swift)，请在 overcloud 部署期间包含以下环境文件：

```
-e environments/cephadm/cephadm-rbd-only.yaml
-e environments/disable-swift.yaml
```

cephadm-rbd-only.yaml 文件配置 RBD，但不配置 RGW。**disable-swift.yaml** 文件确保 Object Storage 服务(swift)未部署。

5.6. 配置块存储备份服务以使用 CEPH

Block Storage Backup 服务(cinder-backup)被默认禁用。必须启用它才能与 Ceph 搭配使用。

流程

要启用块存储备份服务(cinder-backup)，请在部署 overcloud 时使用以下环境文件：

```
`/usr/share/openstack-tripleo-heat-templates/environments/cinder-backup.yaml`.
```

5.7. 为 CEPH 节点配置多个绑定接口

使用绑定接口组合多个 NIC，并为网络连接添加冗余。如果您在 Ceph 节点上有足够的 NIC，则可以在每个节点上创建多个绑定接口来扩展冗余功能。

为节点所需的每个网络连接使用一个绑定接口。这可为每个网络提供冗余和专用连接。

如需更多信息和流程，请参阅 *安装和管理 Red Hat OpenStack Platform 指南* 中的 [置备 overcloud 网络](#)。

第 6 章 使用原生 CEPHFS 部署共享文件系统服务

CephFS 是 Red Hat Ceph Storage 的高可扩展、开源分布式文件系统组件，它是一个统一的分布式存储平台。Ceph Storage 使用可靠的自主分布式对象存储(RADOS)实施对象、块和文件存储。CephFS 是兼容 POSIX 的，提供对 Ceph Storage 集群的文件访问。

共享文件系统服务(manila)允许用户在 CephFS 中创建共享，并使用原生 Ceph FS 协议访问它们。共享文件系统服务管理 OpenStack 中这些共享的生命周期。

在此发行版本中，director 可以使用原生 CephFS 后端在 overcloud 上部署共享文件系统服务。



重要

本章与部署和原生 CephFS 使用原生 CephFS 一起，通过原生 CephFS NAS 协议在 Red Hat OpenStack Platform (RHOSP)云中提供自助服务共享文件系统服务。这种类型的部署需要虚拟客户机虚拟机访问 Ceph 公共网络和基础架构。仅部署具有可信 OpenStack Platform 租户的原生 CephFS，因为它需要 permissive 信任模型，它不适用于一般目的 OpenStack Platform 部署。对于使用传统租户信任模型的常规 OpenStack Platform 部署，您可以通过 NFS 协议部署 CephFS。

6.1. 原生 CEPHFS 带有原生驱动程序

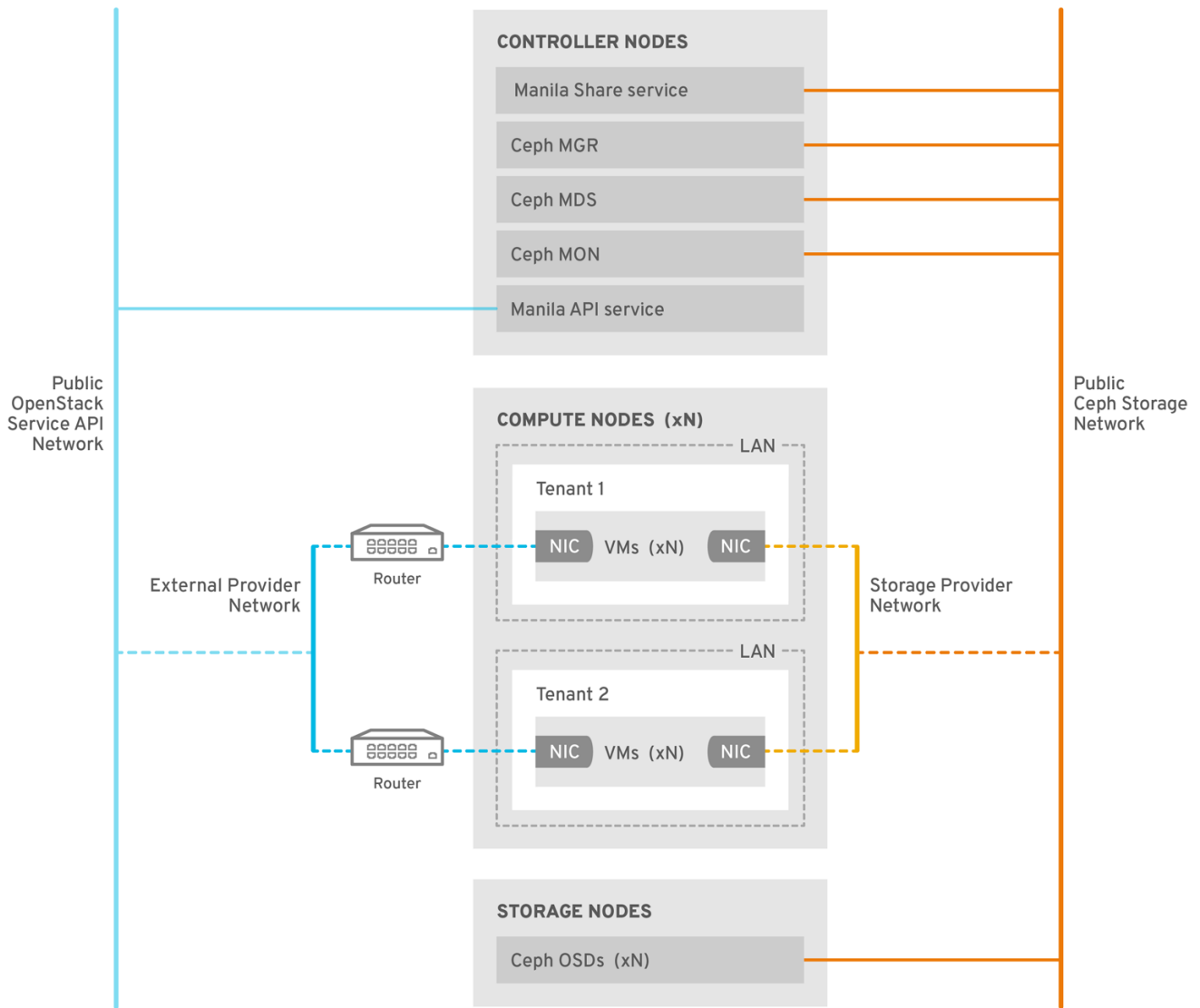
CephFS 原生驱动程序将 OpenStack 共享文件系统服务(manila)和 Red Hat Ceph Storage 结合。使用 Red Hat OpenStack(RHOSP)director 时，Controller 节点托管 Ceph 守护进程，如管理器、元数据服务器(MDS)，以及监控(MON)和共享文件系统服务。

Compute 节点可以托管一个或多个项目。项目（以前称为租户）在下图中以白色框表示。项目包含用户管理的虚拟机，由带有两个 NIC 的灰色方框表示。若要访问 ceph 和 manila 守护进程项目，可通过公共 Ceph 存储网络连接到守护进程。

在这个网络上，您可以访问 Ceph Object Storage Daemon(OSD)提供的存储节点上的数据。在项目引导时托管的实例或虚拟机(VM)带有两个 NIC：一个用于存储提供商网络，第二个则专用于外部提供商网络。

存储提供商网络将项目上运行的虚拟机连接到公共 Ceph 存储网络。Ceph 公共网络提供 Ceph 对象存储节点、元数据服务器(MDS)和控制器节点的后端访问。

通过利用原生驱动程序，CephFS 依赖于与客户端和服务端协作来强制实施配额，保证项目隔离和安全性。具有原生驱动程序的 CephFS 可在具有信任最终用户的环境中正常工作。此配置要求在用户控制下运行软件，才能正常工作。



OPENSTACK_476233_0818

6.2. 原生 CEPHFS 后端安全性

原生 CephFS 后端需要在 Red Hat OpenStack Platform (RHOSP) 租户中使用 permissive 信任模型。这种信任模型不适用于一般目的 OpenStack Platform 云，它们会故意阻止用户直接访问 OpenStack 平台所提供的服务背后的基础架构。

使用原生 CephFS，用户计算实例会直接连接到公开 Ceph 服务守护进程的 Ceph 公共网络。用户虚拟机上运行的 CephFS 客户端与 Ceph 服务守护进程协同工作，它们直接与 RADOS 交互，以读写文件数据块。

强制实施共享文件系统服务 (manila) 共享大小的 CephFS 配额在客户端端强制执行，如在 (RHOSP) 用户拥有的虚拟机上。用户虚拟机上的客户端侧软件可能不是最新的，这可能会使关键云基础架构易受攻击，或者意外地发生以 Ceph 服务端口为目标的软件。

仅在受信任的用户保持客户端软件最新保持客户端软件的环境中，部署原生 CephFS 作为后端。确保没有影响 Red Hat Ceph Storage 基础架构的软件在虚拟机上运行。

为了提供许多不可信用户的 RHOSP 部署，请部署 CephFS-NFS。有关使用 CephFS-NFS 的更多信息，[请参阅部署 Red Hat Ceph Storage 和 Red Hat OpenStack Platform 和 director](#)。

用户可能无法使客户端软件保持最新状态，它们可能无法从其虚拟机中排除有害的软件，而是使用

CephFS-NFS，他们只能访问 NFS 服务器的公共端，而不是 Ceph 基础架构本身。NFS 不需要同一类型的合作客户端，在最糟糕的情况下，用户虚拟机的攻击可能会破坏 NFS 网关，而不会损害 Ceph 存储基础架构。

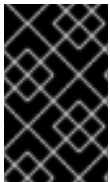
您可以向所有可信用户公开原生 CephFS 后端，但您必须了解以下安全措施：

- 将存储网络配置为提供商网络。
- 提供基于角色的访问控制(RBAC)策略来保护存储供应商网络。
- 创建私有共享类型。

6.3. 原生 CEPHFS 部署

在 Red Hat OpenStack Platform(RHOSP)环境中安装典型的原生 Ceph 文件系统(CephFS)包括以下组件：

- 运行容器化 Ceph 元数据服务器(MDS)、Ceph 监控器(MON)和共享文件系统服务(manila)服务的 RHOSP Controller 节点。这些服务在同一节点上共存，或者有一个或多个专用节点。
- Ceph Storage 集群，具有在 Ceph Storage 节点上运行的容器化对象存储守护进程(OSD)。
- 用作 Ceph 公共网络的隔离存储网络，这些公共网络可以和 Ceph 服务守护进程通信。为便于提供存储网络，存储网络作为提供商网络供用户连接其虚拟机并挂载 CephFS 共享。



重要

您不能将共享文件系统服务(manila)与 CephFS 原生驱动程序通过 Manila CSI 提供给 OpenShift Container Platform，因为红帽不支持这样的部署。如需更多信息，请联系红帽支持团队。

共享文件系统服务(manila)服务提供 API，使租户能够请求文件系统共享，由驱动程序模块实现。Red Hat CephFS 的驱动程序 `manila.share.drivers.cephfs.driver.CephFSDriver`，允许共享文件系统服务将原生 CephFS 用作后端。您可以在由 director 管理的集成部署中安装原生 CephFS。

当 director 使用 overcloud 上的 CephFS 后端部署共享文件系统服务时，它会自动创建所需的数据中心存储网络。但是，您必须在 overcloud 上创建对应的存储供应商网络。

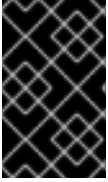
有关网络规划的更多信息，请参阅[使用 director 安装和管理 Red Hat OpenStack Platform 中的 Overcloud 网络](#)。

虽然您可以通过编辑 `/var/lib/config-data/puppet-generated/manila/manila/manila.conf` 文件来手动配置共享文件系统服务，但在以后的 overcloud 更新中，任何设置都可以被 Red Hat OpenStack Platform director 覆盖。红帽只支持部署由 director 管理的共享文件系统服务。

6.4. 要求

如果满足以下要求，您可以使用新的或现有的 Red Hat OpenStack Platform(RHOSP)环境部署原生 CephFS 后端：

- 使用 Red Hat OpenStack Platform 版本 17.0 或更高版本。
- 同时配置与原生 CephFS 后端相同的 Red Hat Ceph Storage 集群。有关如何部署 Ceph 存储的详情，请参考 [部署 Red Hat Ceph Storage](#) 和 [Red Hat OpenStack Platform 以及 director](#)。



重要

支持带有原生 CephFS 后端的 RHOSP 共享文件系统服务(manila)用于 Red Hat Ceph Storage 版本 5.2 或更高版本。有关如何确定系统上安装的 Ceph Storage 版本的更多信息，请参阅 [Red Hat Ceph Storage 发行版本以及对应的 Ceph 软件包版本](#)。

- 在 Controller 节点上安装共享文件系统服务。这是默认的行为。
- 仅将 CephFS 后端的单个实例用于共享文件系统服务。

6.5. 文件共享

共享文件系统服务(manila)、Ceph 文件系统(CephFS)和 CephFS-NFS 以不同的方式管理共享。

共享文件系统服务提供共享，其中共享是单个文件系统命名空间和带有定义大小的存储单元。共享文件系统存储允许多个客户端将数据连接到任何给定共享，但您必须为每个客户端通过共享文件系统服务访问控制 API 访问共享，然后才能进行连接。

CephFS 管理共享，如具有定义配额的目录，以及指向特定存储池或命名空间的布局。CephFS 配额将目录的大小限制为共享文件系统服务创建的共享的大小。

您可以使用元数据服务(MDS)身份验证功能控制对原生 CephFS 共享的访问。使用原生 CephFS，文件共享通过 CephFS 协议来置备并访问。使用使用 CephFS 用户名的 CephX 身份验证方案来执行访问控制。

6.6. 原生 CEPHFS 的网络隔离

原生 CephFS 部署使用 director 作为 Red Hat Ceph Storage 公共网络部署的隔离存储网络。客户端使用此网络与各种 Ceph Storage 基础架构服务守护进程通信。有关隔离网络的更多信息，请参阅[使用 director 安装和管理 Red Hat OpenStack Platform 中的网络隔离](#)。

6.7. 部署原生 CEPHFS 环境

当您准备好部署环境时，请使用 **openstack overcloud deploy** 命令以及配置原生 CephFS 后端所需的自定义虚拟环境和角色。

除其它所需选项外，**openstack overcloud deploy** 命令还具有以下选项。

| 操作 | 选项 | 其它信息 |
|------------------------------------|---|--|
| 使用 network_data.yaml 指定网络配置 | <code>[filename] -n /usr/share/openstack-tripleo-heat-templates/network_data.yaml</code> | 您可以使用自定义环境文件覆盖此网络数据文件中指定的默认网络值。这是使用隔离网络时可用的默认网络数据文件。您可以从 openstack overcloud deploy 命令中省略此文件，以进行 brevity。 |
| 部署 Ceph 守护进程。 | <code>-e /usr/share/openstack-tripleo-heat-templates/environments/cephadm/cephadm.yaml</code> | <i>Deploying Red Hat Ceph Storage 和 Red Hat OpenStack Platform together with director 中的 启动 overcloud 部署</i> |

| 操作 | 选项 | 其它信息 |
|--|--|--|
| 使用 ceph-mds.yaml 部署 Ceph 元数据服务器 | -e /usr/share/openstack-tripleo-heat-templates/environments/cephadm/ceph-mds.yaml | <i>Deploying Red Hat Ceph Storage 和 Red Hat OpenStack Platform together with director 中的 启动 overcloud 部署</i> |
| 使用原生 CephFS 后端部署 manila 服务。 | -e /usr/share/openstack-tripleo-heat-templates/environments/manila-cephfsnative-config.yaml | 环境文件 |

以下示例显示了 **openstack overcloud deploy** 命令，其中包含部署 Ceph 集群、Ceph MDS、原生 CephFS 后端和 Ceph 集群所需的网络的选项：

```
[stack@undercloud ~]$ openstack overcloud deploy \
...
-n /usr/share/openstack-tripleo-heat-templates/network_data.yaml \
-e /usr/share/openstack-tripleo-heat-templates/environments/network-isolation.yaml \
-e /home/stack/network-environment.yaml \
-e /usr/share/openstack-tripleo-heat-templates/environments/cephadm/cephadm.yaml \
-e /usr/share/openstack-tripleo-heat-templates/environments/cephadm/ceph-mds.yaml \
-e /usr/share/openstack-tripleo-heat-templates/environments/manila-cephfsnative-config.yaml
```

有关 **openstack overcloud deploy** 命令的更多信息，请参阅[使用 director 安装和管理 Red Hat OpenStack Platform 中的 置备和部署 overcloud](#)。

6.8. 原生 CEPHFS 后端环境文件

用于定义原生 CephFS 后端 **manila-cephfsnative-config.yaml** 的环境文件位于 undercloud 节点的以下路径中：**/usr/share/openstack-tripleo-heat-templates/environments/manila-cephfsnative-config.yaml**。

manila-cephfsnative-config.yaml 环境文件包含与部署共享文件系统服务相关的设置。后端默认设置应该适用于大多数环境。

示例显示了 director 在部署共享文件系统服务过程中使用的默认值：

```
[stack@undercloud ~]$ cat /usr/share/openstack-tripleo-heat-templates/environments/manila-cephfsnative-config.yaml

# A Heat environment file which can be used to enable a
# a Manila CephFS Native driver backend.
resource_registry:
  OS::TripleO::Services::ManilaApi: ../deployment/manila/manila-api-container-puppet.yaml
  OS::TripleO::Services::ManilaScheduler: ../deployment/manila/manila-scheduler-container-puppet.yaml
  # Only manila-share is pacemaker managed:
  OS::TripleO::Services::ManilaShare: ../deployment/manila/manila-share-pacemaker-puppet.yaml
  OS::TripleO::Services::ManilaBackendCephFs: ../deployment/manila/manila-backend-cephfs.yaml
```

```
parameter_defaults:
  ManilaCephFSBackendName: cephfs 1
  ManilaCephFSDriverHandlesShareServers: false 2
  ManilaCephFSCephFSAuthId: 'manila' 3
  ManilaCephFSCephFSEnableSnapshots: true 4
  ManilaCephFSCephVolumeMode: '0755' 5
  # manila cephfs driver supports either native cephfs backend - 'CEPHFS'
  # (users mount shares directly from ceph cluster), or nfs-ganesha backend -
  # 'NFS' (users mount shares through nfs-ganesha server)
  ManilaCephFSCephFSProtocolHelperType: 'CEPHFS' 6
```

parameter_defaults 标头表示配置的开头。具体来说，此标头中的设置可让您覆盖 **resource_registry** 中设置的默认值。这包括 **OS::Tripleo::Services::Services::ManilaBackendCephFs** 设置的值，它为 CephFS 后端设置默认值。

- 1 **ManilaCephFSBackendName** 设置 CephFS 后端的 manila 配置的名称。在这种情况下，默认后端名称为 **cephfs**。
- 2 **ManilaCephFSDriverHandlesShareServers** 控制共享服务器的生命周期。当设置为 **false** 时，驱动程序不会处理生命周期。这是 CephFS 后端唯一支持的选项。
- 3 **ManilaCephFSCephFSAuthId** 定义 director 为 manila 服务访问 Ceph 集群而创建的 Ceph 身份验证 ID。
- 4 **ManilaCephFSEnableSnapshots** 控制快照激活。Ceph Storage 4.1 及更新版本支持快照，但此参数的值默认为 **false**。您可以将值设为 **true** 以确保驱动程序向 manila 调度程序报告 **snapshot_support** 功能。
- 5 **ManilaCephFSCephVolumeMode** 控制对在原生 CephFS 后端中创建的 manila 共享设置的 UNIX 权限。默认值为 **755**。
- 6 **ManilaCephFSCephFSProtocolHelperType** 必须设置为 **CEPHFS** 以使用原生 CephFS 驱动程序。

有关环境文件的更多信息，请参阅 [使用 director 安装和管理 Red Hat OpenStack Platform 指南](#) 中的 [环境文件](#)。

第 7 章 使用 CEPHFS-NFS 部署共享文件系统服务

当您通过 NFS 网关(NFS-Ganesha)将共享文件系统服务(manila)与 Ceph 文件系统(CephFS)搭配使用时，您可以使用您用于块存储和对象存储的相同的 Red Hat Ceph Storage 集群通过 NFS 协议提供文件共享。

从 Red Hat OpenStack Platform (RHOSP)版本 13 开始完全支持 CephFS-NFS。RHOSP 17.0 及更新的带有 CephFS-NFS 的 RHOSP 共享文件系统服务(manila)支持用于 Red Hat Ceph Storage 版本 5.2 或更高版本。有关如何确定系统上安装的 Ceph Storage 版本的更多信息，请参阅 [Red Hat Ceph Storage 发行版本以及对应的 Ceph 软件包版本](#)。

CephFS 是 Red Hat Ceph Storage 的高可扩展开源分布式文件系统组件，它是一个统一的分布式存储平台。Ceph Storage 使用可靠的自主分布式对象存储(RADOS)实施对象、块和文件存储。CephFS 是兼容 POSIX 的，提供对 Ceph Storage 集群的文件访问。

共享文件系统服务允许用户在 CephFS 中创建共享，并通过用户空间 NFS 服务器软件 NFS-Ganesha 使用 NFS 4.1 访问它们。NFS-Ganesha 控制对共享的访问，并通过 NFS 4.1 协议将它们导出到客户端。共享文件系统服务管理 RHOSP 中这些共享的生命周期。当云管理员将服务配置为使用 CephFS-NFS 时，这些文件共享来自 CephFS 集群，但它们作为熟悉的 NFS 共享创建并访问。

有关共享文件系统服务的更多信息，请参阅 [配置持久性存储](#) 中的 [配置共享文件系统服务\(manila\)](#)。

7.1. 先决条件

- 您可以在 Controller 节点上安装共享文件系统服务，因为是默认行为。
- 您必须通过 RHOSP director 为存储流量创建一个 StorageNFS 网络。
- 您可以在 Controller 节点的 Pacemaker 集群上安装 NFS-Ganesha 网关服务。
- 您仅配置 CephFS 后端的一个实例，以使用共享文件系统服务。您可以将其他非 CephFS 后端用于单个 CephFS 后端。

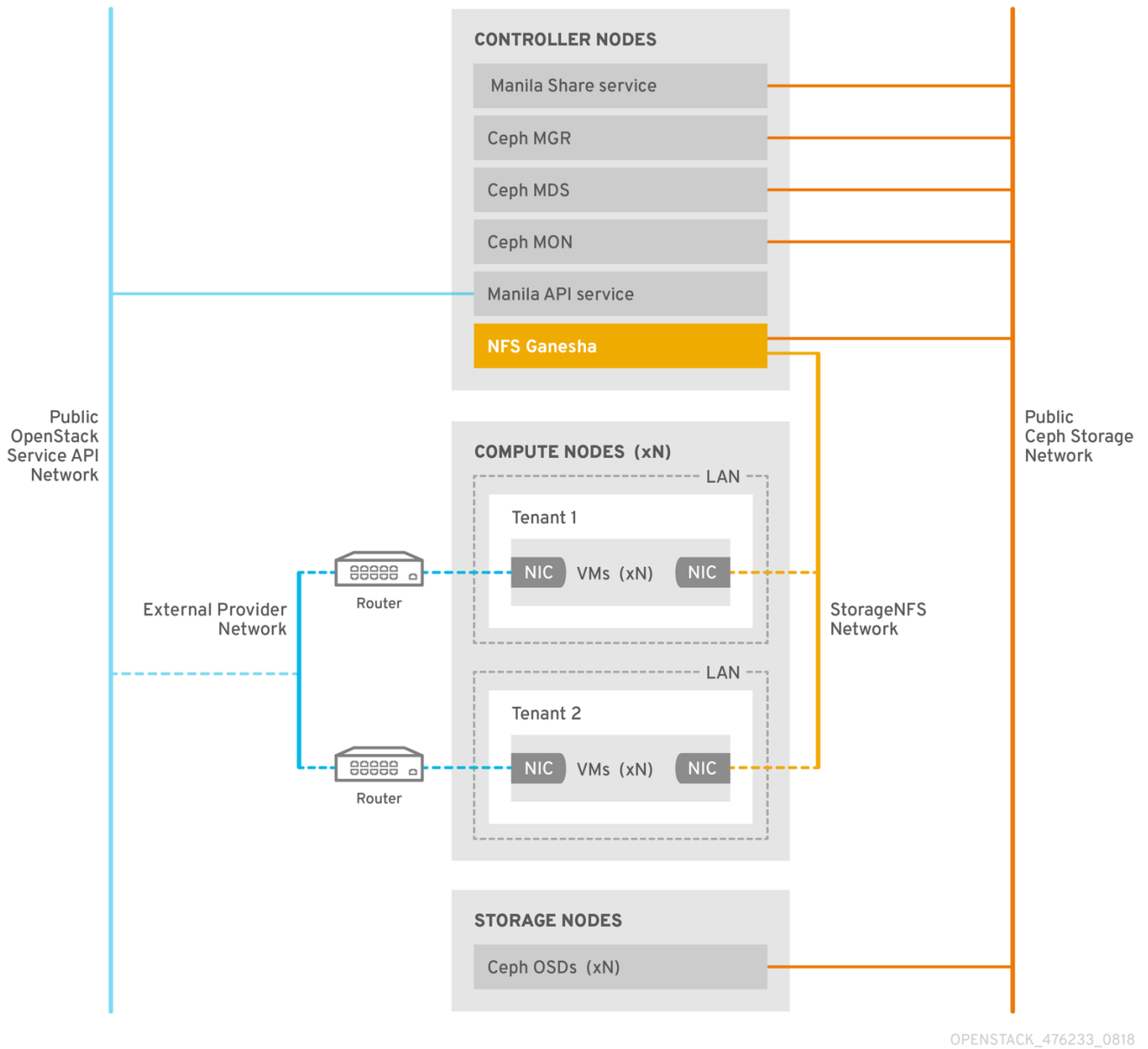
7.2. CEPHFS-NFS 驱动程序

共享文件系统服务(manila)中的 CephFS-NFS 后端由 Ceph 元数据服务器(MDS)、NFS 网关(NFS-Ganesha)和 Red Hat Ceph Storage 集群服务组件组成。

共享文件系统服务 CephFS-NFS 驱动程序使用 NFS-Ganesha 提供 CephFS 共享的 NFSv4 协议访问。Ceph MDS 服务将文件系统的目录和文件名称映射到存储在 RADOS 集群中的对象。NFS 网关可以提供与 Ceph 等不同存储后端的 NFS 文件共享。NFS-Ganesha 服务通过 Ceph 服务在 Controller 节点上运行。

使用隔离网络部署是可选的，但建议使用。在这种情况下，实例使用至少两个 NIC 引导：一个 NIC 连接到项目路由器，第二个 NIC 连接到 StorageNFS 网络，该网络直接连接到 NFS-Ganesha。实例使用 NFS 协议挂载共享。在 Ceph Object Storage Daemon (OSD)节点上托管的 CephFS 共享通过 NFS 网关提供。

NFS-Ganesha 通过防止用户实例直接访问 MDS 和其他 Ceph 服务来提高安全性。实例无法直接访问 Ceph 守护进程。



7.3. RED HAT CEPH STORAGE 服务和客户端访问

当使用 Red Hat Ceph Storage 提供对象和块存储时，您需要以下服务进行部署：

- Ceph 监视器(MON)
- Object Storage Daemon (OSD)
- RADOS 网关(RGW)
- Manager (管理者)

对于原生 CephFS，您还需要 Ceph 存储元数据服务(MDS)和 CephFS-NFS，您需要 NFS-Ganesha 服务作为使用 NFS 协议的原生 CephFS 的网关。

NFS-Ganesha 在自己的容器中运行，它们同时连接到 Ceph 公共网络和新的隔离网络 StorageNFS。如果使用 Red Hat OpenStack Platform (RHOSP) director 的可组合网络功能，您可以部署隔离网络并将其连接到 Controller 节点。作为云管理员，您可以将网络配置为 Networking(neutron)提供商网络。

NFS-Ganesha 通过 Ceph 公共网络访问 CephFS，并使用 StorageNFS 网络上的地址来绑定其 NFS 服务。

要访问 NFS 共享，您可以使用连接到 Storage NFS 网络的额外 NIC 置备 Compute (nova)实例。CephFS 共享的导出位置显示为标准 NFS IP:<path> 元组，它们使用 StorageNFS 网络上的 NFS-Ganesha 服务器 VIP。网络使用实例的 IP 地址对 NFS 共享执行访问控制。

网络(neutron)安全组阻止属于项目 1 的实例通过 StorageNFS 网络访问属于项目 2 的实例。项目共享相同的 CephFS 文件系统，但实施项目数据路径分离，因为实例只能访问导出树下的文件：

`/path/to/share1/.... /path/to/share2/....`

7.4. 使用 CEPHFS-NFS 容错的共享文件系统服务

当 Red Hat OpenStack Platform (RHOSP) director 启动 Red Hat Ceph Storage 服务守护进程时，它们管理自己的高可用性(HA)状态，通常有多个运行这些守护进程的实例。相反，在这个发行版本中，一个 NFS-Ganesha 实例一次只能提供文件共享。

为了避免 CephFS-NFS 共享的数据路径中出现单点故障，NFS-Ganesha 在由 Pacemaker-Corosync 集群管理的主动 - 被动配置中的 RHOSP Controller 节点上运行。NFS-Ganesha 使用虚拟服务 IP 地址来跨 Controller 节点充当虚拟服务。

如果一个 Controller 节点失败或者特定 Controller 节点上的服务失败，且无法在那个节点上恢复，Pacemaker-Corosync 会使用同一虚拟 IP 地址在不同的 Controller 节点上启动新的 NFS-Ganesha 实例。保留现有客户端挂载，因为它们使用虚拟 IP 地址进行共享导出位置。

失败后，使用默认 NFS mount-option 设置和 NFS 4.1 或更高版本，并重置 TCP 连接和客户端重新连接。I/O 操作在故障切换过程中临时停止响应，但它们不会失败。应用程序 I/O 也停止响应，但在故障转移完成后恢复。

新连接、新锁定状态等，直到服务器在达到 90 秒的宽限期后才会被拒绝，直到服务器等待客户端回收其锁定。如果所有客户端都回收其锁定，NFS-Ganesha 保留了客户端列表并提前退出宽限期。

7.5. CEPHFS-NFS 安装

Red Hat OpenStack Platform (RHOSP)环境中的典型 CephFS-NFS 安装包括以下配置：

- 运行以下 OpenStack Controller 节点：
 - Ceph 监视器(MON)
 - 容器化 Ceph 元数据服务器(MDS)
 - 共享文件系统服务(manila)
 - NFS-Ganesha
 其中一些服务可以在同一节点上共存，或者可以有一个或多个专用节点。
- 在 Ceph Storage 节点上运行的容器化对象存储守护进程(OSD)的 Red Hat Ceph Storage 集群
- 隔离存储NFS 网络，为 NFS 共享调配提供从项目到 NFS-Ganesha 服务的访问



重要

使用 CephFS-NFS 的共享文件系统服务完全支持通过 Manila CSI 向 Red Hat OpenShift Container Platform 提供共享。这个解决方案不适用于大规模部署。有关重要建议，请参阅 <https://access.redhat.com/articles/6667651>。

共享文件系统服务提供 API，允许项目请求文件系统共享，由驱动程序模块实现。如果将驱动程序用于 CephFS，`manila.share.drivers.cephfs.driver.CephFSDriver`，您可以使用 CephFS 后端的共享文件系统服务。RHOSP director 配置驱动程序来部署 NFS-Ganesha，以便 CephFS 共享通过 NFS 4.1 协议提供。

在准备 CephFS NFS 部署时，您将需要隔离的 **StorageNFS** 网络。您可以使用 director 创建这个隔离的 **StorageNFS** 网络。如需更多信息，请参阅 *使用 director 安装和管理 Red Hat OpenStack Platform* 中的 [配置 overcloud 网络](#)。

共享文件系统服务后端的手动配置选项

您可以通过编辑节点文件 `/etc/manila/manila.conf` 来手动配置共享文件系统服务。但是，RHOSP director 可以覆盖将来的 overcloud 更新中的任何设置。

您可以将 CephFS-NFS 添加到外部部署的 Ceph Storage 集群，它不由 director 配置。目前，您只能在 director 中定义一个 CephFS 后端。有关更多信息，请参阅 *Integrating the overcloud with an existing Red Hat Ceph Storage Cluster* 中的 [Integrating an overcloud with Ceph Storage](#) 部分。

7.6. 文件共享

共享文件系统服务(manila)、Ceph 文件系统(CephFS)和 CephFS-NFS 以不同的方式管理共享。

共享文件系统服务提供共享，其中共享是单个文件系统命名空间和带有定义大小的存储单元。共享文件系统存储允许多个客户端将数据连接到任何给定共享，但您必须为每个客户端通过共享文件系统服务访问控制 API 访问共享，然后才能进行连接。

CephFS 管理共享，如具有定义配额的目录，以及指向特定存储池或命名空间的布局。CephFS 配额将目录的大小限制为共享文件系统服务创建的共享的大小。

您可以通过指定客户端的 IP 地址来控制对 CephFS-NFS 共享的访问。使用 CephFS-NFS 时，文件共享通过 NFS 协议调配并访问。NFS 协议也管理安全性。

7.7. CEPHFS-NFS 的网络隔离

为安全起见，在使用 CephFS-NFS 时将 NFS 流量隔离到单独的网络，以便 NFS 服务器只能通过隔离的网络访问。Deployers 可以将隔离网络限制为云中选定的一组项目。Red Hat OpenStack (RHOSP) director 提供了部署专用 StorageNFS 网络的支持。

在部署 overcloud 以启用 CephFS-NFS 以用于共享文件系统服务前，您必须创建以下内容：

- 用于 NFS 流量的隔离网络，称为 **StorageNFS**
- 隔离网络上的虚拟 IP (VIP)
- 使用 StorageNFS 网络配置节点的 Controller 节点的自定义角色

有关创建隔离网络、VIP 和自定义角色的更多信息，请参阅 *使用 director 安装和管理 Red Hat OpenStack Platform* 中的 [配置 overcloud 网络](#)。



重要

可以省略为 NFS 流量创建隔离网络。但是，如果您在具有不受信任的客户端的生产部署中省略 StorageNFS 网络，则 director 可以在任何共享的、非隔离网络上（如外部网络）连接 Ceph NFS 服务器。共享网络通常可路由到云中的所有用户专用网络。当以这种方式通过路由网络访问 NFS 服务器时，您无法通过应用客户端 IP 访问规则来控制对共享文件系统服务共享的访问。用户必须允许使用通用 **0.0.0.0/0** IP 访问其共享。由于是通用的 IP，可以发现导出路径的任何人都可以挂载共享。

7.8. 部署 CEPHFS-NFS 环境

当您准备好部署环境时，请使用 **openstack overcloud deploy** 命令以及使用 NFS-Ganesha 运行 CephFS 所需的自定义虚拟环境和角色。

overcloud 部署命令除其它所需选项外，还具有以下选项：

| 操作 | 选项 | 附加信息 |
|---|--|--|
| 引用部署的网络，包括 StorageNFS 网络 | -e /home/stack/templates/overcloud-networks-deployed.yaml | 使用 <i>director</i> 安装和管理 Red Hat OpenStack Platform 中的配置 overcloud 网络 。如果您不想将 NFS 流量隔离到单独的网络，可以省略 StorageNFS 网络选项。 |
| 引用在部署的网络上创建的虚拟 IP，包括 StorageNFS 网络的 VIP | -e /home/stack/templates/overcloud-vip-deployed.yaml | 使用 <i>director</i> 安装和管理 Red Hat OpenStack Platform 中的配置 overcloud 网络 。如果您不想将 NFS 流量隔离到单独的网络，可以省略这个选项。 |
| 添加 roles_data.yaml 文件中定义的自定义角色。部署命令使用自定义角色将网络分配给 Controller 节点 | -r /home/stack/roles_data.yaml | 如果您不想将 NFS 流量隔离到单独的网络，可以省略这个选项。 |
| 部署 Ceph 守护进程。 | -e /usr/share/openstack-tripleo-heat-templates/environments/cephadm/cephadm.yaml | <i>Deploying Red Hat Ceph Storage 和 Red Hat OpenStack Platform together with director</i> 中的 启动 overcloud 部署 |
| 使用 ceph-mds.yaml 部署 Ceph 元数据服务器 | -e /usr/share/openstack-tripleo-heat-templates/environments/cephadm/ceph-mds.yaml | <i>Deploying Red Hat Ceph Storage 和 Red Hat OpenStack Platform together with director</i> 中的 启动 overcloud 部署 |
| 使用 CephFS-NFS 后端部署共享文件系统服务(manila)。使用 <i>director</i> 配置 NFS-Ganesha。 | -e /usr/share/openstack-tripleo-heat-templates/environments/manila-cephfs-ganesha-config.yaml | manila-cephfs-ganesha-config.yaml 环境文件 |

以下示例显示了使用 NFS-Ganesha、Ceph Storage 集群和 Ceph MDS 部署 CephFS 的 **openstack overcloud deploy** 命令：

```
[stack@undercloud ~]$ openstack overcloud deploy \
--templates /usr/share/openstack-tripleo-heat-templates \
-r /home/stack/roles_data.yaml \
-e /home/stack/templates/overcloud-networks-deployed.yaml \
-e /home/stack/templates/overcloud-vip-deployed.yaml \
-e /home/stack/containers-default-parameters.yaml \
-e /usr/share/openstack-tripleo-heat-templates/environments/network-isolation.yaml \
-e /home/stack/network-environment.yaml \
-e /usr/share/openstack-tripleo-heat-templates/environments/cephadm/cephadm.yaml \
-e /usr/share/openstack-tripleo-heat-templates/environments/cephadm/ceph-mds.yaml \
-e /usr/share/openstack-tripleo-heat-templates/environments/manila-cephfsganesha-config.yaml
```

有关 **openstack overcloud deploy** 命令的更多信息，请参阅使用 *director* 安装和管理 Red Hat OpenStack Platform 中的 [置备和部署 overcloud](#)。

7.9. CEPHFS-NFS 后端环境文件

用于定义 CephFS-NFS 后端 **manila-cephfsganesha-config.yaml** 的环境文件位于 undercloud 节点的以下路径中：**/usr/share/openstack-tripleo-heat-templates/environments/manila-cephfsganesha-config.yaml**。

manila-cephfsganesha-config.yaml 环境文件包含与部署共享文件系统服务(manila)相关的设置。后端默认设置可用于大多数环境。以下示例显示了 *director* 在部署共享文件系统服务过程中使用的默认值：

```
[stack@undercloud ~]$ cat /usr/share/openstack-tripleo-heat-templates/environments/manila-cephfsganesha-config.yaml
# A Heat environment file which can be used to enable a
# a Manila CephFS-NFS driver backend.
resource_registry:
  OS::TripleO::Services::ManilaApi: ../deployment/manila/manila-api-container-puppet.yaml
  OS::TripleO::Services::ManilaScheduler: ../deployment/manila/manila-scheduler-container-puppet.yaml
  # Only manila-share is pacemaker managed:
  OS::TripleO::Services::ManilaShare: ../deployment/manila/manila-share-pacemaker-puppet.yaml
  OS::TripleO::Services::ManilaBackendCephFs: ../deployment/manila/manila-backend-cephfs.yaml
  # ceph-nfs (ganesha) service is installed and configured by Director
  # but it's still managed by pacemaker
  OS::TripleO::Services::CephNfs: ../deployment/cephadm/ceph-nfs.yaml

parameter_defaults:
  ManilaCephFSBackendName: cephfs 1
  ManilaCephFSDriverHandlesShareServers: false 2
  ManilaCephFSCephFSAuthId: 'manila' 3
  # manila cephfs driver supports either native cephfs backend - 'CEPHFS'
  # (users mount shares directly from ceph cluster), or nfs-ganesha backend -
  # 'NFS' (users mount shares through nfs-ganesha server)
  ManilaCephFSCephFSProtocolHelperType: 'NFS'
```

parameter_defaults 标头表示配置的开头。要覆盖 **resource_registry** 中设置的默认值，请将此 **manila-cephfsganesha-config.yaml** 环境文件复制到本地环境文件目录 **/home/stack/templates/**，并根据需要编辑您的环境所需的参数设置。这包括 **OS::TripleO::Services::Services::ManilaBackendCephFs** 设置

的值，它为 CephFS 后端设置默认值。

- 1 **ManilaCephFSBackendName** 设置 CephFS 后端的 manila 配置名称。在这种情况下，默认后端名称为 **cephfs**。
- 2 **ManilaCephFSDriverHandlesShareServers** 控制共享服务器的生命周期。当设置为 **false** 时，驱动程序不会处理生命周期。这是唯一支持的选项。
- 3 **ManilaCephFSCephFSAuthId** 定义了 director 为 **manila** 服务访问 Ceph 集群而创建的 Ceph 身份验证 ID。

有关环境文件的更多信息，请参阅 *使用 director 安装和管理 Red Hat OpenStack Platform 指南* 中的[环境文件](#)。

第 8 章 启动 OVERCLOUD 部署

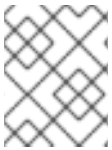
完成服务初始配置和自定义后，部署 overcloud。

8.1. 启动 OVERCLOUD 部署

部署 overcloud 以实现 Red Hat OpenStack Platform (RHOSP)环境的配置。

先决条件

- 在 undercloud 安装过程中，在 **undercloud.conf** 文件中设置 **generate_service_certificate=false**。否则，部署 overcloud 时，您必须注入信任定位符。



注意

如果要在 overcloud 部署期间添加 Ceph 仪表盘，请参阅 [第 10 章 将 Red Hat Ceph Storage 仪表盘添加到 overcloud 部署中](#)。

流程

使用 **openstack overcloud deploy** 命令部署 overcloud。有关所有命令参数的完整列表，请参阅 [命令行界面参考中的 openstack overcloud deploy](#)。

以下是使用该命令的示例：

```
$ openstack overcloud deploy --templates -r /home/stack/templates/roles_data_custom.yaml \
-e /usr/share/openstack-tripleo-heat-templates/environments/cephadm/cephadm.yaml \
-e /usr/share/openstack-tripleo-heat-templates/environments/cephadm/ceph-mds.yaml \
-e /usr/share/openstack-tripleo-heat-templates/environments/cinder-backup.yaml \
-e /home/stack/templates/storage-config.yaml \
-e /home/stack/templates/deployed-ceph.yaml \
-e /home/stack/templates/networks-deployed.yaml \
-e /home/stack/templates/deployed-metal.yaml \
-e /home/stack/templates/deployed-vips.yaml \
--ntp-server pool.ntp.org
```

示例命令使用以下选项：

- templates**
 - 从默认的 heat 模板集合 **/usr/share/openstack-tripleo-heat-templates/** 创建 overcloud。
- r /home/stack/templates/roles_data_custom.yaml**
 - 指定自定义角色定义文件。
- e /usr/share/openstack-tripleo-heat-templates/environments/cephadm/cephadm.yaml**
 - 设置 director 以完成之前部署的 Ceph Storage 集群。此环境文件默认部署 RGW。它还创建池、密钥和守护进程。如果您不需要部署 RGW 或对象存储，请查看 [第 5.5 节 “Red Hat OpenStack Platform 对象存储的部署选项”](#) 中介绍的选项
- e /usr/share/openstack-tripleo-heat-templates/environments/cephadm/ceph-mds.yaml**
 - 启用 Ceph 元数据服务器，如 [第 5.3 节 “启用 Ceph 元数据服务器”](#) 所述。

- **-e /usr/share/openstack-tripleo-heat-templates/environments/cinder-backup.yaml**
 - 启用块存储备份服务(**cinder-backup**), 如 第 5.6 节 “配置块存储备份服务以使用 Ceph” 所述。
- **-e /home/stack/templates/storage-config.yaml**
 - 添加包含自定义 Ceph Storage 配置的环境文件, 如下所述 第 5.1 节 “配置自定义环境文件”
- **-e /home/stack/templates/deployed-ceph.yaml**
 - 添加包含 Ceph 集群设置的环境文件, 如之前运行的 **openstack overcloud ceph deploy** 命令的输出一样。
- **-e /home/stack/templates/networks-deployed.yaml**
 - 添加包含 Ceph 集群网络设置的环境文件, 如 **openstack overcloud network provision** 的输出。
- **-e /home/stack/templates/deployed-metal.yaml**
 - 添加包含 Ceph 集群节点设置的环境文件, 如 **openstack overcloud node provision** 的输出。
- **-e /home/stack/templates/deployed-vips.yaml**
 - 添加包含 Ceph 集群网络 VIP 设置的环境文件, 如 **openstack overcloud network vip provision** 的输出。
- **--ntp-server pool.ntp.org**
 - 设置 NTP 服务器。

第 9 章 使用 DIRECTOR 为不同的工作负载定义性能层

Red Hat OpenStack Platform (RHOSP) director 部署 Red Hat Ceph Storage 性能层。Ceph Storage CRUSH 规则与 **CephPools** 参数组合，以使用设备类功能。这会构建不同的层，以适应具有不同性能要求的工作负载。

例如，您可以为常规工作负载定义 HDD 类，以及一个 SSD 类，仅通过 SSD 分发数据，以实现高性能负载。在这种情况下，当您创建新块存储卷时，您可以选择性能层(HDD 或 SSD)。

有关创建 CRUSH 规则的更多信息，[请参阅配置 CRUSH 层次结构](#)。



警告

在现有环境中定义性能层可能会导致 Ceph Storage 集群中的数据移动。director 在堆栈更新过程中使用 **cephadm**。**cephadm** 应用没有验证池是否存在并包含数据的逻辑。更改与池关联的默认 CRUSH 规则会导致数据移动。如果池包含大量数据，则会移动这些数据。

如果您需要帮助或建议添加或删除节点，请联系红帽支持。

Ceph Storage 自动检测磁盘类型，并将其分配到对应的设备类；HDD、SSD 或 NVMe；基于 Linux 内核公开的硬件属性。

先决条件

- 对于新部署，请使用 Red Hat Ceph Storage (RHCS) 版本 5.2 或更高版本。

9.1. 配置性能层

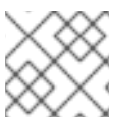
要部署不同的 Red Hat Ceph Storage 性能层，请创建一个新的环境文件，其中包含 CRUSH map 详情并将其包含在部署命令中。director 不会公开这个功能的特定参数，但您可以生成 **tripleo-ansible** 预期变量。



注意

性能层配置可以与 CRUSH 层次结构相结合。有关创建 [CRUSH 规则的信息](#)，[请参阅配置 CRUSH 层次结构](#)。

在示例流程中，每个 Ceph Storage 节点包含三个 OSD：**sdb** 和 **sdc** 是 spinning 磁盘，**sdv** 是一个 SSD。Ceph 自动检测正确的磁盘类型。然后，您要配置两个 CRUSH 规则(HDD 和 SSD)来映射到两个对应的设备类。



注意

HDD 规则是默认值，适用于所有池，除非您使用不同的规则配置池。

最后，您创建一个名为 **fastpool** 的额外池，并将它映射到 SSD 规则。这个池最终通过 Block Storage (cinder) 后端公开。任何消耗此块存储后端的工作负载都由 SSD 支持，以便只获得快速的性能。您可以将它用于数据或从卷引导。

WARNING

在现有环境中定义性能层可能会导致 Ceph 集群中进行大量数据移动。**Cephadm**, **director** 在堆栈更新期间触发，没有逻辑来验证池是否已在 Ceph 集群中定义，并且是否包含数据。这意味着，在现有环境中定义性能层可能会危险，因为与池关联的默认 CRUSH 规则的更改会导致数据移动。如果您需要帮助或建议添加或删除节点，请联系红帽支持。

流程

1. 以 **stack** 用户身份登录 undercloud 节点。
2. 创建一个环境文件，如 `/home/stack/templates/ceph-config.yaml`，使其包含 Ceph 配置参数和设备类变量。或者，您可以在现有环境文件中添加以下配置。
3. 添加 **CephCrushRules** 参数。**CephCrushRules** 参数必须包含您定义或 Ceph 自动检测到的每个类的规则。在创建新池时，如果没有指定规则，则会选择您希望 Ceph 用作默认值的规则。

```
CephCrushRules:
```

```
- name: HDD
  root: default
  type: host
  class: hdd
  default: true
- name: SSD
  root: default
  type: host
  class: ssd
  default: false
```

4. 添加 **CephPools** 参数：

- 使用 **rule_name** 参数指定不使用默认规则的每个池的层。在以下示例中，**fastpool** 池使用配置为快速层的 SSD 设备类来管理块存储卷。
- 使用 **CinderRbdExtraPools** 参数将 **fastpool** 配置为块存储后端。

```
CephPools:
```

```
- name: fastpool
  rule_name: SSD
  application: rbd
```

```
CinderRbdExtraPools: fastpool
```

5. 使用以下示例来确保环境文件包含正确的值：

```
parameter_defaults:
```

```
CephCrushRules:
```

```
- name: replicated_hdd
  default: true
  class: hdd
  root: default
```

```

type: host
CinderRbdExtraPools: fastpool
CephPools:
- name: fastpool
  rule_name: SSD
  application: rbd

```

6. 在 **openstack overcloud deploy** 命令中包含新的环境文件。

```

$ openstack overcloud deploy \
--templates \
...
-e <other_overcloud_environment_files> \
-e /home/stack/templates/ceph-config.yaml \
...

```

将 **<other_overcloud_environment_files>** 替换为属于部署的其他环境文件的列表。

重要

如果您将环境文件应用到现有的 Ceph 集群，则预先存在的 Ceph 池不会使用新规则更新。因此，部署完成后您必须输入以下命令将规则设置为指定池。

```
$ ceph osd pool set <pool> crush_rule <rule>
```

- 将 **<pool>** 替换为您要将新规则应用到的池的名称。
- 将 **<rule>** 替换为您通过 **crush_rules** 参数指定的规则名称之一。

对于您通过此命令更改的每个规则，请更新现有条目或在现有模板中的 **CephPools** 参数中添加新条目：

```

CephPools:
- name: <pool>
  rule_name: <rule>
  application: rbd

```

9.2. 验证 CRUSH 规则和池

验证 CRUSH 规则和池设置。

WARNING

在现有环境中定义性能层可能会导致 Ceph 集群中进行大量数据移动。**tripleo-ansible**, **director** 在堆栈更新期间触发，没有逻辑来检查 Ceph 集群中是否已定义池，并且是否包含数据。这意味着，在现有环境中定义性能层可能会危险，因为与池关联的默认 CRUSH 规则的更改会导致数据移动。如果您需要帮助或建议添加或删除节点，请联系红帽支持。

流程

1. 以 **tripleo-admin** 用户身份登录 overcloud Controller 节点。
2. 要验证您的 OSD 层是否已成功设置，请输入以下命令。

```
$ sudo cephadm shell ceph osd tree
```

3. 在生成的树视图中，验证 **CLASS** 列是否显示您设置的每个 OSD 的正确设备类。
4. 另外，使用以下命令验证 OSD 是否已正确分配给设备类。

```
$ sudo cephadm shell ceph osd crush tree --show-shadow
```

5. 将生成的层次结构与以下命令的结果进行比较，以确保对每个规则应用相同的值。

```
$ sudo cephadm shell ceph osd crush rule dump <rule_name>
```

- 将 <rule_name> 替换为您要检查的规则的名称。
6. 根据您在部署期间使用的 **crush_rules** 参数，验证您创建的规则名称和 ID 是否正确。

```
$ sudo cephadm shell ceph osd crush rule dump | grep -E "rule_(id|name)"
```

7. 验证 Ceph 池已绑定到在第 3 步中获得的正确 CRUSH 规则 ID。

```
$ sudo cephadm shell -- ceph osd dump | grep pool
```

8. 对于每个池，请确保规则 ID 与您预期的规则名称匹配。

第 10 章 将 RED HAT CEPH STORAGE 仪表板添加到 OVERCLOUD 部署中

Red Hat Ceph Storage Dashboard 默认禁用，但您可以使用 Red Hat OpenStack Platform (RHOSP) director 在 overcloud 中启用。Ceph 控制面板是内置的、基于 Web 的 Ceph 管理和监控应用，用于管理 Ceph 集群中的各个方面和对象。Red Hat Ceph Storage Dashboard 包括以下组件：

- Ceph Dashboard manager 模块提供用户界面，并嵌入平台前端 Grafana。
- Prometheus，监控插件。
- Alertmanager 将警报发送到仪表板。
- 节点导出器将 Ceph 集群数据导出到仪表板。



注意

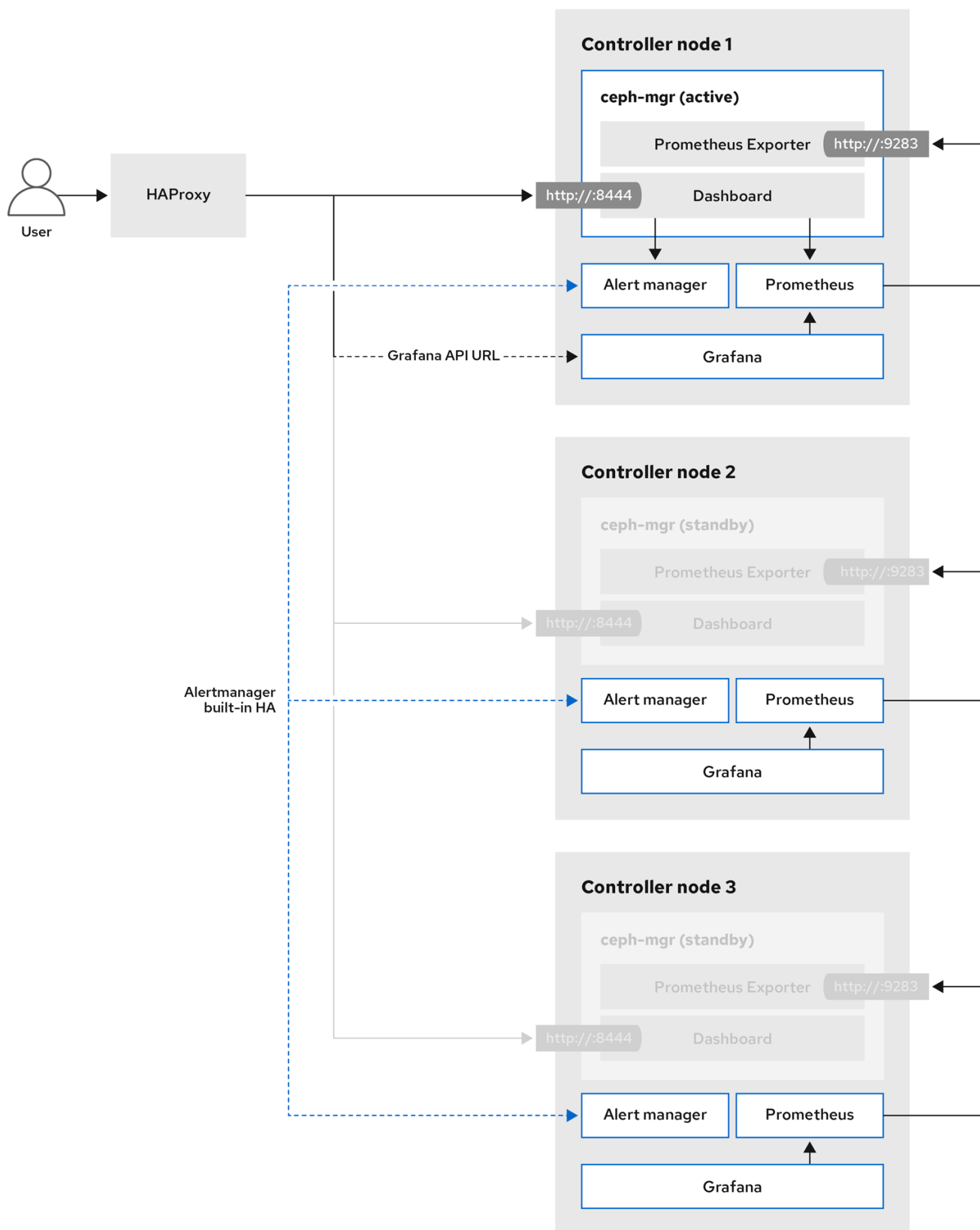
Ceph Storage 4.1 或更高版本支持此功能。有关如何确定系统上安装的 Ceph Storage 版本的更多信息，请参阅 [Red Hat Ceph Storage 发行版本以及对应的 Ceph 软件包版本](#)。



注意

Red Hat Ceph Storage Dashboard 始终与其他 Ceph 管理器组件位于同一个节点上。

下图显示了 Red Hat OpenStack Platform 上 Ceph 仪表板的架构：



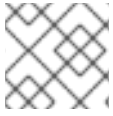
89_Ceph_0520

有关控制面板及其功能及限制的更多信息，请参阅 *Red Hat Ceph Storage Dashboard 指南* 中的 [控制面板功能](#)。

10.1. 使用 CEPH 仪表板的任何位置 TLS

控制面板前端与 TLS 随处框架完全集成。您可以在任何提供有所需环境文件的 TLS，并将它们包含在

overcloud 部署命令中。这会在 overcloud 部署期间触发 Grafana 和 Ceph 仪表板的证书请求，以及生成的证书和密钥文件传递给 **cephadm**。有关如何为仪表板以及其他 RHOSP 服务启用 TLS 的更多信息，请[参阅高级 Overcloud 自定义指南中的以下主题](#)：



注意

即使位于 TLS-everywhere 上下文中，访问 Ceph 仪表板的端口也会保持不变。

10.2. 包含 CEPH 仪表板所需的容器

在将 Ceph 控制面板模板添加到 overcloud 之前，您必须使用 **containers-prepare-parameter.yaml** 文件包含必要的容器。要生成 **containers-prepare-parameter.yaml** 文件来准备您的容器镜像，请完成以下步骤：

流程

1. 以 **stack** 用户身份登录 undercloud 主机。
2. 生成默认的容器镜像准备文件：

```
$ sudo openstack tripleo container image prepare default \
  --local-push-destination \
  --output-env-file containers-prepare-parameter.yaml
```

3. 编辑 **containers-prepare-parameter.yaml** 文件并进行修改以满足您的要求。以下示例 **containers-prepare-parameter.yaml** 文件包含与 Dashboard 服务相关的镜像位置和标签，包括 Grafana、Prometheus、Alertmanager 和 Node Exporter。根据您的特定场景编辑值：

```
parameter_defaults:
  ContainerImagePrepare:
    - push_destination: true
      set:
        ceph_alertmanager_image: ose-prometheus-alertmanager
        ceph_alertmanager_namespace: registry.redhat.io/openshift4
        ceph_alertmanager_tag: v4.12
        ceph_grafana_image: rhceph-6-dashboard-rhel9
        ceph_grafana_namespace: registry.redhat.io/rhceph
        ceph_grafana_tag: 6
        ceph_image: rhceph-6-rhel9
        ceph_namespace: registry.redhat.io/rhceph
        ceph_node_exporter_image: ose-prometheus-node-exporter
        ceph_node_exporter_namespace: registry.redhat.io/openshift4
        ceph_node_exporter_tag: v4.12
        ceph_prometheus_image: ose-prometheus
        ceph_prometheus_namespace: registry.redhat.io/openshift4
        ceph_prometheus_tag: v4.12
        ceph_tag: latest
```

有关使用 **containers-prepare-parameter.yaml** 文件的 registry 和镜像配置的更多信息，请[参阅自定义 Red Hat OpenStack Platform 部署指南中的容器镜像准备参数](#)。

10.3. 部署 CEPH 仪表板

包含 **ceph-dashboard** 环境文件，以部署 Ceph 控制面板。

完成此步骤后，生成的部署由带有 **grafana**、**prometheus**、**alertmanager** 和 **node-exporter** 容器的外部堆栈组成。Ceph Dashboard manager 模块是此堆栈的后端，它会嵌入 **grafana** 布局，为最终用户提供特定于集群的指标。



注意

如果要使用可组合网络部署 Ceph 控制面板，请参阅 [第 10.4 节 “使用可组合网络部署 Ceph 仪表板”](#)。



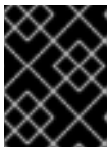
注意

Ceph 控制面板 admin 用户角色默认设置为只读模式。要更改 Ceph Dashboard admin 默认模式，请参阅 [第 10.5 节 “更改默认权限”](#)。

流程

1. 以 **stack** 用户身份登录 undercloud 节点。
2. 可选：Ceph Dashboard 网络默认设置为 provisioning 网络。如果要部署 Ceph 控制面板并通过其他网络访问它，请创建一个环境文件，例如：**ceph_dashboard_network_override.yaml**。将 **CephDashboardNetwork** 设置为现有 overcloud 路由网络之一，如 **external**：

```
parameter_defaults:
  ServiceNetMap:
    CephDashboardNetwork: external
```



重要

在初始部署后，不支持将 **CephDashboardNetwork** 值更改为从不同的网络访问 Ceph 控制面板。

3. 在 **openstack overcloud deploy** 命令中包含以下环境文件。如果选择更改默认网络，包含属于部署的所有环境文件，以及 **ceph_dashboard_network_override.yaml** 文件：

```
$ openstack overcloud deploy \
  --templates \
  -e <overcloud_environment_files> \
  -e /usr/share/openstack-tripleo-heat-templates/environments/cephadm/cephadm.yaml \
  -e /usr/share/openstack-tripleo-heat-templates/environments/cephadm/ceph-
  dashboard.yaml \
  -e ceph_dashboard_network_override.yaml
```

将 **<overcloud_environment_files>** 替换为属于部署的环境文件列表。

10.4. 使用可组合网络部署 CEPH 仪表板

您可以在可组合网络中部署 Ceph 控制面板，而不是在默认的 Provisioning 网络中部署。这消除了 Provisioning 网络中公开 Ceph 控制面板服务的需要。在可组合网络中部署仪表板时，您还可以实施单独的授权配置集。

您必须在部署前选择要使用的网络，因为您只能在首次部署 overcloud 时将控制面板应用到新网络。在部署前，使用以下步骤选择可组合网络。

完成此步骤后，生成的部署由带有 **grafana**、**prometheus**、**alertmanager** 和 **node-exporter** 容器的外部堆栈组成。Ceph Dashboard manager 模块是此堆栈的后端，它会嵌入 **grafana** 布局，为最终用户提供特定于集群的指标。

流程

1. 以 stack 用户身份登录 undercloud。
2. 生成特定于 Controller 的角色，使其包含 Dashboard 可组合网络：

```
$ openstack overcloud roles generate -o /home/stack/roles_data_dashboard.yaml
ControllerStorageDashboard Compute BlockStorage ObjectStorage CephStorage
```

- 在定义为命令输出的 YAML 文件中生成一个新的 **ControllerStorageDashboard** 角色。使用 **overcloud deploy** 命令时，您必须将此 YAML 文件包含在模板列表中。**ControllerStorageDashboard** 角色不包含 **CephNFS** 或 **network_data_dashboard.yaml**。
 - director 提供了一个定义可组合网络的网络环境文件。此文件的默认位置为 **/usr/share/openstack-tripleo-heat-templates/network_data_dashboard.yaml**。使用 **overcloud deploy** 命令时，您必须将此文件包含在 **overcloud** 模板列表中。
3. 在 **openstack overcloud deploy** 命令中包含以下环境文件，以及作为部署一部分的所有环境文件：

```
$ openstack overcloud deploy \
  --templates \
  -r /home/stack/roles_data.yaml \
  -n /usr/share/openstack-tripleo-heat-templates/network_data_dashboard.yaml \
  -e /usr/share/openstack-tripleo-heat-templates/environments/network-isolation.yaml \
  -e /usr/share/openstack-tripleo-heat-templates/environments/network-environment.yaml \
  -e <overcloud_environment_files> \
  -e /usr/share/openstack-tripleo-heat-templates/environments/cephadm/cephadm.yaml \
  -e /usr/share/openstack-tripleo-heat-templates/environments/cephadm/ceph-
  dashboard.yaml
```

将 **<overcloud_environment_files>** 替换为属于部署的环境文件列表。

10.5. 更改默认权限

Ceph Dashboard admin 用户角色默认设置为只读模式，以便安全地监控 Ceph 集群。要允许 admin 用户具有提升的特权，以便使用控制面板更改 Ceph 集群的元素，您可以使用 **CephDashboardAdminRO** 参数更改默认的 admin 权限。



警告

具有完整权限的用户可能会更改 director 配置的 Ceph 集群的元素。这可能会导致运行堆栈更新时与 director 配置的选项冲突。要避免这个问题，请不要使用 Ceph 控制面板更改 director 配置的选项，例如 Ceph OSP 池属性。

流程

1. 以 **stack** 用户的身份登录 undercloud。
2. 创建以下 **ceph_dashboard_admin.yaml** 环境文件：

```
parameter_defaults:
  CephDashboardAdminRO: false
```

3. 运行 `overcloud deploy` 命令以更新现有的堆栈，并包含您创建的所有其他环境文件，这些文件属于现有部署：

```
$ openstack overcloud deploy \
--templates \
-e <existing_overcloud_environment_files> \
-e ceph_dashboard_admin.yml
```

使用作为您现有部署一部分的环境文件列表替换 `<existing_overcloud_environment_files>`。

10.6. 访问 CEPH 仪表板

若要测试 Ceph 控制面板运行是否正确，请完成以下用于访问它的验证步骤，并检查它从 Ceph 集群显示的数据是否正确。仪表板应该可以被完全访问，显示的数字和图形应反映 `ceph -s` 命令显示的相同集群状态信息。

流程

1. 以 **stack** 用户身份登录 undercloud 节点。
2. 检索仪表板 admin 登录凭证：

```
[stack@undercloud ~]$ grep tripleo_cephadm_dashboard_admin_password <config-download>/<stack>/cephadm/cephadm-extra-vars-heat.yml
```

3. 检索 VIP 地址以访问 Ceph 控制面板：

```
[stack@undercloud-0 ~]$ grep tripleo_cephadm_dashboard_frontend_vip <config-download>/<stack>/cephadm/cephadm-extra-vars-ansible.yml
```

4. 使用 Web 浏览器指向前端 VIP 并访问控制面板。director 在 provisioning 网络中配置和公开仪表板，以便您可以使用您检索到的 VIP 在 TCP 端口 8444 上直接访问仪表板。确保满足以下条件：
 - Web 客户端主机是连接到 provisioning 网络的第 2 层。
 - 置备网络会被正确路由或代理，可以从 web 客户端主机访问它。如果没有满足这些条件，您仍然可以打开 SSH 隧道来访问 overcloud 上的 Dashboard VIP：

```
client_host$ ssh -L 8444:<dashboard_vip>:8444 stack@<your undercloud>
```

将 `<dashboard_vip>` 替换为您检索到的 control plane VIP 的 IP 地址。

5. 要访问仪表板，使用浏览器访问 <http://localhost:8444>，并使用以下详情进行登录：
 - **cephadm** 创建的默认用户：`admin`。

- `<config-download>/<stack>/cephadm/cephadm-extra-vars-heat.yml` 中的密码。

有关 Red Hat Ceph Storage 仪表板的更多信息，请参阅 [Red Hat Ceph Storage 管理指南](#)

第 11 章 管理 RED HAT CEPH STORAGE 集群的部署后操作

使用容器化 Red Hat Ceph Storage 部署 Red Hat OpenStack Platform (RHOSP) 环境后，您可以使用一些操作来管理 Ceph Storage 集群。

11.1. 禁用配置覆盖

最初部署 Ceph Storage 集群后，集群会被配置为允许在 overcloud 部署期间设置 RGW 等服务。overcloud 部署完成后，除非要扩展集群，否则不应使用 director 来更改集群配置。应使用 Ceph 命令执行集群配置更改。

流程

1. 以 **stack** 用户身份登录 undercloud 节点。
2. 打开文件 **deployed_ceph.yaml** 或您环境中使用的文件来定义 Ceph 存储集群配置。
3. 找到 **ApplyCephConfigOverridesOnUpdate** 参数。
4. 将 **ApplyCephConfigOverridesOnUpdate** 参数值更改为 **false**。
5. 保存该文件。

其他资源

有关 **ApplyCephConfigOverridesOnUpdate** 和 **CephConfigOverrides** 参数的更多信息，请参阅 [Overcloud 参数](#)。

11.2. 访问 OVERCLOUD

director 会生成脚本来配置和帮助认证 undercloud 与 overcloud 的交互。director 将此文件保存为 **stack** 用户主目录的 **overcloudrc** 文件。

流程

1. 运行以下命令来 source 文件：

```
$ source ~/overcloudrc
```

这会加载必要的环境变量，以便从 undercloud CLI 与 overcloud 交互。

2. 要返回与 undercloud 进行交互的状态，请运行以下命令：

```
$ source ~/stackrc
```

11.3. 监控 RED HAT CEPH STORAGE 节点

创建 overcloud 后，检查 Ceph 集群的状态以确认它正常工作。

流程

1. 以 **tripleo-admin** 用户身份登录 Controller 节点：

```
$ nova list
$ ssh tripleo-admin@192.168.0.25
```

2. 检查 Ceph 集群的运行状况：

```
$ sudo cephadm shell -- ceph health
```

如果 Ceph 集群没有问题，命令会返回 **HEALTH_OK**。这意味着 Ceph 集群是安全的使用。

3. 登录到运行 Ceph 监控服务的 overcloud 节点，并检查 Ceph 集群中的所有 OSD 的状态：

```
$ sudo cephadm shell -- ceph osd tree
```

4. 检查 Ceph Monitor 仲裁的状态：

```
$ sudo cephadm shell -- ceph quorum_status
```

这将显示参与仲裁以及哪个是领导的监控器。

5. 验证所有 Ceph OSD 是否正在运行：

```
$ sudo cephadm shell -- ceph osd stat
```

如需有关监控 Ceph 集群的更多信息，请参阅 *Red Hat Ceph Storage Administration Guide* 中的[监控 Ceph Storage 集群](#)。

11.4. 将 BLOCK STORAGE (CINDER)类型映射到您的新 CEPH 池

完成配置步骤后，使用 Block Storage (cinder)创建映射到您创建的 **fastpool** 层的类型，将性能层功能提供给 RHOSP 租户。

流程

1. 以 **stack** 用户身份登录 undercloud 节点。
2. 获取 **overcloudrc** 文件：

```
$ source overcloudrc
```

3. 检查块存储卷现有的类型：

```
$ cinder type-list
```

4. 创建新的块存储卷 fast_tier：

```
$ cinder type-create fast_tier
```

5. 检查是否创建了 Block Storage 类型：

```
$ cinder type-list
```

6. 当 **fast_tier** Block Storage 类型可用时，将 **fastpool** 设置为您创建的新层的块存储卷后端：

```
$ cinder type-key fast_tier set volume_backend_name=tripleo_ceph_fastpool
```

7. 使用新层创建新卷：

```
$ cinder create 1 --volume-type fast_tier --name fastdisk
```



注意

Red Hat Ceph Storage 文档为 Ceph Storage 集群的持续维护和操作提供了额外的信息和流程。请参阅 [Red Hat Ceph Storage 产品文档](#)。

第 12 章 原生 CEPHFS 部署后配置和验证

在创建 CephFS 共享、授予用户访问权限和挂载 CephFS 共享之前，您必须完成一些部署后配置任务。

- 将 Networking 服务(neutron)存储网络映射到隔离的数据中心存储网络。
- 通过基于自定义角色的访问控制(RBAC)，使存储提供商网络仅适用于可信租户。不要全局共享存储提供商网络。
- 创建私有共享类型。
- 授予特定可信租户的访问权限。

完成这些步骤后，租户计算实例可以创建、允许访问和挂载原生 CephFS 共享。

将原生 CephFS 部署为共享文件系统服务(manila)的后端，将以下新元素添加到 overcloud 环境中：

- 存储提供商网络
- Controller 节点上的 Ceph MDS 服务

云管理员必须先验证原生 CephFS 环境的稳定性，然后将其提供给服务用户。

有关将共享文件系统服务与原生 CephFS 搭配使用的更多信息，请参阅 [配置持久性存储指南](#) 中的 [配置共享文件系统服务\(manila\)](#)。

12.1. 创建存储提供商网络

您必须将新的隔离存储网络映射到 Networking(neutron)提供商网络。Compute 虚拟机附加到网络，以访问原生 CephFS 共享导出位置。

有关使用共享文件系统服务(manila)的网络安全性的信息，请参阅 [强化 Red Hat OpenStack Platform 中的强化共享文件系统服务](#)。

流程

openstack network create 命令定义存储 neutron 网络的配置。

1. 提供 overcloud 凭证文件：

```
$ source ~/<credentials_file>
```

- 将 **<credentials_file>** 替换为您的凭据文件的名称，如 **overcloudrc**。

2. 在 undercloud 节点上，创建存储网络：

```
(overcloud) [stack@undercloud-0 ~]$ openstack network create Storage --provider-network-type vlan --provider-physical-network datacentre --provider-segment 30
```

您可以使用以下选项输入这个命令：

- 对于 **--provider-physical-network** 选项，使用默认值 **datacentre**，除非您通过 `tripleo-heat-templates` 中的 `NeutronBridgeMappings` 为 `br-isolated` 网桥设置了另一个标签。
- 对于 **--provider-segment** 选项，在网络环境文件中使用为存储隔离网络设置的值。如果没有自定义，默认的环境文件为 `/usr/share/openstack-tripleo-heat-`

`templates/network_data.yaml`。除非修改了隔离网络定义，否则与存储网络值关联的 VLAN 为 30。

- 对于 `--provider-network-type` 选项，请使用值 `vlan`。

12.2. 配置存储提供商网络

在 neutron 提供商网络上创建对应的 **StorageSubnet**。确保子网与 undercloud 中的 `storage_subnet` 相同，并且存储子网的分配范围不会重叠。

要求

- 分配池的开头和结束 IP 范围
- 子网 IP 范围

流程

1. 在 undercloud 节点上输入以下命令：

```
[stack@undercloud ~]$ source ~/overcloudrc
```

2. 使用示例命令置备网络。更新这些值以适合您的环境。

```
(overcloud) [stack@undercloud-0 ~]$ openstack subnet create \
--allocation-pool start=172.17.3.10,end=172.17.3.149 \
--dhcp \
--network Storage \
--subnet-range 172.17.3.0/24 \
--gateway none StorageSubnet
```

- 对于 `--allocation-pool` 选项，将 `start=172.17.3.10,end=172.17.3.149` IP 值替换为您的网络的 IP 值。
- 对于 `--subnet-range` 选项，请将 `172.17.3.0/24` 子网范围替换为您的网络的子网范围。

12.3. 为存储供应商网络配置基于角色的访问控制

在识别可以使用存储网络的可信租户或项目后，通过网络服务(neutron)为它们配置基于角色的访问控制(RBAC)规则。

要求

需要访问存储网络的项目名称

流程

1. 在 undercloud 节点上输入以下命令：

```
[stack@undercloud ~]$ source ~/overcloudrc
```

2. 识别需要访问的项目：

```
(overcloud) [stack@undercloud-0 ~]$ openstack project list
```

```
+-----+
| ID           | Name |
+-----+
| 06f1068f79d2400b88d1c2c33eacea87 | demo |
| 5038dde12dfb44fdaa0b3ee4bfe487ce | service |
| 820e2d9c956644c2b1530b514127fd0d | admin |
+-----+
```

3. 使用所需项目创建网络 RBAC 规则：

```
(overcloud) [stack@undercloud-0 ~]$ openstack network rbac create \
--action access_as_shared Storage \
--type network \
--target-project demo
```

对需要访问存储网络的所有项目重复此步骤。

12.4. 配置默认共享类型

您可以使用共享文件系统服务(manila)来定义共享类型，以使用特定设置创建共享。共享类型的工作方式，如块存储卷类型。每种类型都有相关的设置，例如额外的规格。当您在创建共享期间调用类型时，设置将应用到共享文件系统。

为了防止原生 CephFS 后端不受不被信任的用户的影响，请不要创建默认的共享类型。当默认共享类型不存在时，用户会被强制指定共享类型，而受信任的用户可以使用具有独占访问权限的自定义私有共享类型。

如果必须为不受信任的租户创建默认的共享类型，您可以放弃本地 CephFS 后端进行置备。

流程

1. 提供 overcloud 凭证文件：

```
$ source ~/<credentials_file>
```

- 将 **<credentials_file>** 替换为您的凭据文件的名称，如 **overcloudrc**。

2. 在共享类型上设置额外的规格：

```
(overcloud) [stack@undercloud-0 ~]$ manila type-create default false
(overcloud) [stack@undercloud-0 ~]$ manila type-key default set share_backend_name='s!=cephfs'
```

3. 创建私有共享类型，并向可信租户提供对此共享类型的访问权限：

```
(overcloud) [stack@undercloud-0 ~]$ manila type-create --is-public false nativecephfstype false
(overcloud) [stack@undercloud-0 ~]$ manila type-key nativecephfstype set share_backend_name='cephfs'
(overcloud) [stack@undercloud-0 ~]$ manila type-access-add nativecephfstype <trusted_tenant_project_id>
```

- 将 **<trusted_tenant_project_id>** 替换为可信租户的 ID。

有关共享类型的更多信息，[请参阅 配置持久性存储](#)中的创建共享类型。

12.5. 验证隔离存储网络的创建

用于部署原生 CephFS 作为共享文件系统服务后端的 `network_data.yaml` 文件会创建存储 VLAN。使用此流程确认已成功创建了存储 VLAN。

流程

1. 登录到 overcloud 中的一个 Controller 节点。
2. 检查连接的网络，并验证 `network_data.yaml` 文件中的 VLAN 是否存在：

```
$ ip a
8: vlan30: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc noqueue state
UNKNOWN group default qlen 1000
    link/ether 52:9c:82:7a:d4:75 brd ff:ff:ff:ff:ff:ff
    inet 172.17.3.144/24 brd 172.17.3.255 scope global vlan30
        valid_lft forever preferred_lft forever
    inet6 fe80::509c:82ff:fe7a:d475/64 scope link
        valid_lft forever preferred_lft forever
```

12.6. 验证 CEPH MDS 服务

使用 `systemctl status` 命令验证 Ceph MDS 服务状态。

流程

- 在所有 Controller 节点上输入以下命令来检查 MDS 容器的状态：

```
$ systemctl status ceph-mds<@CONTROLLER-HOST>
```

例如：

```
$ systemctl status ceph-mds@controller-0.service

ceph-mds@controller-0.service - Ceph MDS
Loaded: loaded (/etc/systemd/system/ceph-mds@.service; enabled; vendor preset: disabled)
Active: active (running) since Tue 2018-09-18 20:11:53 UTC; 6 days ago
Main PID: 65066 (common)
Tasks: 16 (limit: 204320)
Memory: 38.2M
CGroup: /system.slice/system-ceph\x2dmds.slice/ceph-mds@controller-0.service
└─60921 /usr/bin/podman run --rm --net=host --memory=32000m --cpus=4 -v
/var/lib/ceph:/var/lib/ceph:z -v /etc/ceph:/etc/ceph:z -v
/var/run/ceph:/var/run/ceph:z -v /etc/localtime:/etc/localtime:ro>
```

12.7. 验证 CEPH 集群状态

验证 Ceph 集群状态，以确认集群为 active。

流程

1. 登录任何 Controller 节点。
2. 在 Ceph 监控守护进程中输入以下命令：

```
$ sudo podman exec ceph-mon-controller-0 ceph -s
cluster:
  id:   670dc288-cd36-4772-a4fc-47287f8e2ebf
  health: HEALTH_OK

services:
  mon: 3 daemons, quorum controller-1,controller-2,controller-0 (age 14h)
  mgr: controller-1(active, since 8w), standbys: controller-0, controller-2
  mds: cephfs:1 {0=controller-2=up:active} 2 up:standby
  osd: 15 osds: 15 up (since 8w), 15 in (since 8w)

task status:
  scrub status:
    mds.controller-2: idle

data:
  pools: 6 pools, 192 pgs
  objects: 309 objects, 1.6 GiB
  usage: 21 GiB used, 144 GiB / 165 GiB avail
  pgs: 192 active+clean
```



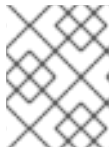
注意

有一个活跃 MDS，备用上有两个 MDS。

3. 要查看 Ceph 文件系统的详细状态，请输入以下命令：

```
$ sudo ceph fs ls

name: cephfs metadata pool: manila_metadata, data pools: [manila_data]
```



注意

在本示例输出中，**cephfs** 是 Ceph 文件系统的名称，director 会创建它来托管用户通过共享文件系统服务创建的 CephFS 共享。

12.8. 验证 MANILA-SHARE 服务状态

验证 manila-share 服务的状态。

流程

1. 在其中一个 Controller 节点上，确认 **openstack-manila-share** 已启动：

```
$ sudo pcs status resources | grep manila

* Container bundle: openstack-manila-share [cluster.common.tag/rhosp16-openstack-manila-share:pcmklatest]:
* openstack-manila-share-podman-0 (ocf::heartbeat:podman): Started controller-0
```

12.9. 验证 MANILA-API 服务确认调度程序和共享服务

完成以下步骤以确认 **manila-api** 服务确认调度程序和共享服务。

流程

1. 登录 undercloud。

2. 使用以下命令：

```
$ source /home/stack/overcloudrc
```

3. 输入以下命令确认启用了 **manila-scheduler** 和 **manila-share**：

```
$ manila service-list
```

```
| Id | Binary          | Host          | Zone | Status | State | Updated_at |
```

```
| 2 | manila-scheduler | hostgroup    | nova | enabled | up | 2018-08-08T04:15:03.000000 |
```

```
| 5 | manila-share    | hostgroup@cephfs | nova | enabled | up | 2018-08-08T04:15:03.000000 |
```

第 13 章 CEPHFS NFS 部署后配置和验证

在创建 NFS 共享、授予用户访问权限和挂载 NFS 共享之前，您必须完成两个部署后配置任务。

- 将网络服务(neutron)StorageNFS 网络映射到隔离的数据中心存储 NFS 网络。如果您不想将 NFS 流量隔离到单独的网络，可以省略这个选项。
- 创建默认共享类型。

完成这些步骤后，租户计算实例可以创建、允许访问和挂载 NFS 共享。

当您为 CephFS-NFS 部署为共享文件系统服务(manila)的后端时，您可以在 overcloud 环境中添加以下新元素：

- StorageNFS 网络
- 控制器上的 Ceph MDS 服务
- 控制器上的 NFS-Ganesha 服务

作为云管理员，您必须先验证 CephFS-NFS 环境的稳定性，然后才能将它提供给服务用户。

13.1. 创建存储提供商网络

将新隔离存储 NFS 网络映射到网络(neutron)提供商网络。Compute 虚拟机附加到网络，以访问由 NFS-Ganesha 网关提供的共享导出位置。

有关使用共享文件系统服务(manila)的网络安全性的信息，请参阅 [强化 Red Hat OpenStack Platform 中的强化共享文件系统服务](#)。

流程

openstack network create 命令定义 StorageNFS neutron 网络的配置。

1. 提供 overcloud 凭证文件：

```
$ source ~/<credentials_file>
```

- 将 **<credentials_file>** 替换为您的凭据文件的名称，如 **overcloudrc**。

2. 在 undercloud 节点上，创建 StorageNFS 网络：

```
(overcloud) [stack@undercloud-0 ~]$ openstack network create StorageNFS --share --
provider-network-type vlan --provider-physical-network datacentre --provider-segment 70
```

您可以使用以下选项输入这个命令：

- 对于 **--provider-physical-network** 选项，使用默认值 **datacentre**，除非您通过 **tripleo-heat-templates** 中的 **NeutronBridgeMappings** 为 **br-isolated** 网桥设置了另一个标签。
- 对于 **--provider-segment** 选项，在 **heat** 模板中为 StorageNFS 隔离网络设置的 VLAN 值 **/usr/share/openstack-tripleo-heat-templates/network_data_ganesha.yaml**。此值为 70，除非部署者修改了隔离的网络定义。
- 对于 **--provider-network-type** 选项，请使用值 **vlan**。

13.2. 配置共享供应商 STORAGE NFS 网络

在 neutron-shared 提供商网络上创建对应的 **StorageNFSSubnet**。确保子网与 **network_data.yml** 文件中的 **storage_nfs** 网络定义相同，并确保 **StorageNFS** 子网的分配范围不会重叠。不需要网关，因为 **StorageNFS** 子网专用于提供 NFS 共享。

先决条件

- 分配池的开头和结束 IP 范围。
- 子网 IP 范围。

13.2.1. 配置共享供应商 StorageNFS IPv4 网络

在 neutron-shared IPv4 提供商网络上创建对应的 **StorageNFSSubnet**。

流程

1. 登录 overcloud 节点。
2. 提供 overcloud 凭据。
3. 使用示例命令置备网络并进行以下更改：
 - a. 将 **start=172.17.0.4,end=172.17.0.250** IP 值替换为您的网络的 IP 值。
 - b. 将 **172.17.0.0/20** 子网范围替换为您的网络子网范围。

```
[stack@undercloud-0 ~]$ openstack subnet create --allocation-pool
start=172.17.0.4,end=172.17.0.250 \
--dhcp --network StorageNFS --subnet-range 172.17.0.0/20 \
--gateway none StorageNFSSubnet
```

13.2.2. 配置共享供应商 StorageNFS IPv6 网络

在 neutron-shared IPv6 提供商网络上创建对应的 **StorageNFSSubnet**。

流程

1. 登录 overcloud 节点。
2. 使用示例命令来置备网络，根据需要更新值。
 - 将 **fd00:fd00:fd00:7000::/64** 子网范围替换为您的网络的子网范围。

```
[stack@undercloud-0 ~]$ openstack subnet create --ip-version 6 --dhcp --network StorageNFS --
subnet-range fd00:fd00:fd00:7000::/64 --gateway none --ipv6-ra-mode dhcpv6-stateful --ipv6-
address-mode dhcpv6-stateful StorageNFSSubnet -f yaml
```

13.3. 配置默认共享类型

您可以使用共享文件系统服务(manila)来定义共享类型，以使用特定设置创建共享。共享类型的工作方式，如块存储卷类型。每种类型都有相关的设置，例如额外的规格。当您在创建共享期间调用类型时，设置将应用到共享文件系统。

使用 Red Hat OpenStack Platform(RHOSP)director, 您必须在打开云前创建一个默认共享类型, 供用户访问。

流程

- 使用 NFS 为 CephFS 创建默认共享类型 :

```
$ manila type-create default false
```

有关共享类型的更多信息, 请参阅 [配置持久性存储](#) 中的创建共享类型。

13.4. 验证隔离存储 NFS 网络的创建

network_data_ganesha.yaml 文件用于将 CephFS-NFS 部署为共享文件系统服务后端, 以创建 StorageNFS VLAN。完成以下步骤以验证隔离存储 NFS 网络是否存在。

流程

1. 登录 overcloud 中的一个控制器。
2. 输入以下命令检查连接的网络, 并验证 **network_data_ganesha.yaml** 中设置的 VLAN 存在 :

```
$ ip a
15: vlan310: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc noqueue state
UNKNOWN group default qlen 1000
    link/ether 32:80:cf:0e:11:ca brd ff:ff:ff:ff:ff:ff
    inet 172.16.4.4/24 brd 172.16.4.255 scope global vlan310
        valid_lft forever preferred_lft forever
    inet 172.16.4.7/32 brd 172.16.4.255 scope global vlan310
        valid_lft forever preferred_lft forever
    inet6 fe80::3080:cfff:fe0e:11ca/64 scope link
        valid_lft forever preferred_lft forever
```

13.5. 验证 CEPH MDS 服务

使用 **systemctl status** 命令验证 Ceph MDS 服务状态。

流程

- 在所有 Controller 节点上输入以下命令来检查 MDS 容器的状态 :

```
$ systemctl status ceph-mds<@CONTROLLER-HOST>
```

例如 :

```
$ systemctl status ceph-mds@controller-0.service

ceph-mds@controller-0.service - Ceph MDS
Loaded: loaded (/etc/systemd/system/ceph-mds@.service; enabled; vendor preset: disabled)
Active: active (running) since Tue 2018-09-18 20:11:53 UTC; 6 days ago
Main PID: 65066 (conmon)
Tasks: 16 (limit: 204320)
Memory: 38.2M
```

```
CGroup: /system.slice/system-ceph\x2dmds.slice/ceph-mds@controller-0.service
└─60921 /usr/bin/podman run --rm --net=host --memory=32000m --cpus=4 -v
/var/lib/ceph:/var/lib/ceph:z -v /etc/ceph:/etc/ceph:z -v
/var/run/ceph:/var/run/ceph:z -v /etc/localtime:/etc/localtime:ro>
```

13.6. 验证 CEPH 集群状态

完成以下步骤以验证 Ceph 集群状态。

流程

1. 登录到活跃的 Controller 节点。
2. 使用以下命令：

```
$ sudo ceph -s

cluster:
  id: 3369e280-7578-11e8-8ef3-801844eeec7c
  health: HEALTH_OK

services:
  mon: 3 daemons, quorum overcloud-controller-1,overcloud-controller-2,overcloud-
controller-0
  mgr: overcloud-controller-1(active), standbys: overcloud-controller-2, overcloud-controller-0
  mds: cephfs-1/1/1 up {0=overcloud-controller-0=up:active}, 2 up:standby
  osd: 6 osds: 6 up, 6 in
```

有一个活跃 MDS，两个备用的 MDS。

3. 要更详细地检查 Ceph 文件系统的状态，请输入以下命令，并将 **<cephfs>** 替换为 Ceph 文件系统的名称：

```
$ sudo ceph fs ls

name: cephfs, metadata pool: manila_metadata, data pools: [manila_data]
```

13.7. 验证 NFS-GANESHA 和 MANILA-SHARE 服务状态

完成以下步骤，以验证 NFS-Ganesha 和 manila-share 服务的状态。

流程

1. 从其中一个 Controller 节点输入以下命令，确认 **ceph-nfs** 和 **openstack-manila-share** 已启动：

```
$ pcs status

ceph-nfs      (systemd:ceph-nfs@pacemaker): Started overcloud-controller-1

podman container: openstack-manila-share [192.168.24.1:8787/rhosp-rhel8/openstack-
manila-share:pcmklatest]
  openstack-manila-share-podman-0    (ocf::heartbeat:podman): Started overcloud-
controller-1
```

13.8. 验证 MANILA-API 服务确认调度程序和共享服务

完成以下步骤以确认 **manila-api** 服务确认调度程序和共享服务。

流程

1. 登录 undercloud。

2. 使用以下命令：

```
$ source /home/stack/overcloudrc
```

3. 输入以下命令确认启用了 **manila-scheduler** 和 **manila-share**：

```
$ manila service-list
```

```
| Id | Binary          | Host          | Zone | Status | State | Updated_at |
```

```
| 2 | manila-scheduler | hostgroup    | nova | enabled | up | 2018-08-08T04:15:03.000000 |
```

```
| 5 | manila-share    | hostgroup@cephfs | nova | enabled | up | 2018-08-08T04:15:03.000000 |
```


第 14 章 重新引导环境

可能需要重新引导环境。例如，当您需要修改物理服务器或从电源中断中恢复时。在这样的情形中，务必要确保 Ceph Storage 节点正确引导。

您必须按照以下顺序引导节点：

1. **首先引导所有 Ceph 监控节点** - 这样可确保 Ceph 监控服务在高可用性 Ceph 集群中处于活动状态。默认情况下，Ceph Monitor 服务安装在 Controller 节点上。如果 Ceph Monitor 与自定义角色中的 Controller 分开，请确保此自定义 Ceph Monitor 角色处于活动状态。
2. **引导所有 Ceph Storage 节点** - 这样可确保 Ceph OSD 集群可以连接到 Controller 节点上的活跃 Ceph Monitor 集群。

14.1. 重新引导 CEPH STORAGE (OSD) 集群

完成以下步骤以重新引导 Ceph Storage (OSD) 节点集群。

先决条件

- 在运行 **ceph-mon** 服务的 Ceph monitor 或 Controller 节点上，检查 Red Hat Ceph Storage 集群状态是否健康，pg 状态为 **active+clean**：

```
$ sudo cephadm -- shell ceph status
```

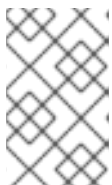
如果 Ceph 集群处于健康状态，它会返回 **HEALTH_OK** 状态。

如果 Ceph 集群状态不健康，它将返回 **HEALTH_WARN** 或 **HEALTH_ERR** 的状态。有关故障排除指南，请参阅 [Red Hat Ceph Storage 5 故障排除指南](#) 或 [Red Hat Ceph Storage 6 故障排除指南](#)。

流程

1. 登录到运行 **ceph-mon** 服务的 Ceph Monitor 或 Controller 节点，并临时禁用 Ceph Storage 集群重新平衡：

```
$ sudo cephadm shell -- ceph osd set noout
$ sudo cephadm shell -- ceph osd set norebalance
```



注意

如果您有多堆栈或分布式计算节点(DCN)架构，您必须在设置 **noout** 和 **norebalance** 标志时指定 Ceph 集群名称。例如：**sudo cephadm shell -c /etc/ceph/<cluster>.conf -k /etc/ceph/<cluster>.client.keyring**。

2. 选择第一个要重新引导的 Ceph Storage 节点并登录到该节点。
3. 重新引导节点：

```
$ sudo reboot
```

4. 稍等片刻，直到节点启动。

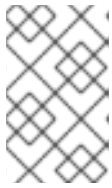
5. 登录节点并检查 Ceph 集群状态：

```
$ sudo cephadm -- shell ceph status
```

确认 **pgmap** 报告的所有 **pgs** 的状态是否都正常 (**active+clean**)。

6. 注销节点，重新引导下一个节点，并检查其状态。重复此过程，直到您已重新引导所有 Ceph Storage 节点。
7. 完成后，登录到运行 **ceph-mon** 服务的 Ceph Monitor 或 Controller 节点，并启用 Ceph 集群重新平衡：

```
$ sudo cephadm shell -- ceph osd unset noout
$ sudo cephadm shell -- ceph osd unset norebalance
```



注意

如果您有多堆栈或分布式计算节点(DCN)架构，您必须在取消设置 **noout** 和 **norebalance** 标志时指定 Ceph 集群名称。例如：**sudo cephadm shell -c /etc/ceph/<cluster>.conf -k /etc/ceph/<cluster>.client.keyring**

8. 执行最后的状态检查，确认集群报告 **HEALTH_OK**：

```
$ sudo cephadm shell ceph status
```

14.2. 重新引导 CEPH STORAGE OSD 以启用与 CEPH 监控服务的连接

如果发生所有 overcloud 节点同时引导的情况，Ceph OSD 服务可能无法在 Ceph Storage 节点上正确启动。在这种情况下，重新引导 Ceph Storage OSD，以便它们可以连接到 Ceph 监控服务。

流程

- 验证 Ceph Storage 节点集群的 **HEALTH_OK** 状态：

```
$ sudo ceph status
```

第 15 章 扩展 CEPH STORAGE 集群

您可以通过添加或删除存储节点来扩展 Ceph Storage 集群的大小。

15.1. 扩展 CEPH STORAGE 集群

随着容量和性能要求的变化，您可以扩展 Ceph 存储集群来满足不断增长的需求。在这样做前，请确保您有足够的节点供更新的部署使用。然后，您可以在 Red Hat OpenStack Platform (RHOSP) 环境中注册和标记新节点。

这个过程会产生以下操作：

- 存储网络和防火墙规则在新的 **CephStorage** 节点上配置。
- **ceph-admin** 用户在新的 **CephStorage** 节点上创建。
- **ceph-admin** 用户公共 SSH 密钥分发到新的 **CephStorage** 节点，以便 **cephadm** 可以使用 SSH 添加额外的节点。
- 如果添加了新的 **CephMon** 或 **CephMgr** 节点，**ceph-admin** 私钥也会分发到该节点。
- 应用更新的 Ceph 规格，**cephadm** 将调度新节点来加入 Ceph 集群。

流程

1. 以 **stack** 用户身份登录 undercloud 主机。
2. 查找 **stackrc** undercloud 凭证文件：

```
$ source ~/stackrc
```

3. 修改 **~/overcloud-baremetal-deploy.yaml**，将 CephStorage 节点添加到部署中。以下示例文件表示具有三个 CephStorage 节点的原始部署。

```
- name: CephStorage
  count: 3
  instances:
    - hostname: ceph-0
      name: ceph-0
    - hostname: ceph-1
      name: ceph-2
    - hostname: ceph-2
      name: ceph-2
```

以下示例修改此文件以添加三个额外的节点。

```
- name: CephStorage
  count: 6
  instances:
    - hostname: ceph-0
      name: ceph-0
    - hostname: ceph-1
      name: ceph-2
    - hostname: ceph-2
```

```

name: ceph-2
- hostname: ceph-3
name: ceph-3
- hostname: ceph-4
name: ceph-4
- hostname: ceph-5
name: ceph-5

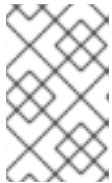
```

- 使用 **openstack overcloud node provision** 命令及更新的 `~/overcloud-baremetal-deploy.yaml` 文件。

```

$ openstack overcloud node provision \
--stack overcloud \
--network-config \
--output ~/overcloud-baremetal-deployed.yaml \
~/overcloud-baremetal-deploy.yaml

```



注意

此命令将置备配置的节点，并输出 `~/overcloud-baremetal-deployed.yaml` 的更新副本。新版本更新 **CephStorage** 角色。**DeployedServerPortMap** 和 **HostnameMap** 还包含新的存储节点。

- 使用 **openstack overcloud ceph spec** 命令生成 Ceph 规范文件。

```

$ openstack overcloud ceph spec ~/overcloud-baremetal-deployed.yaml \
--osd-spec osd_spec.yaml \
--roles-data roles_data.yaml \
-o ceph_spec.yaml

```



注意

openstack overcloud ceph spec 中使用的文件应该已经可用。它们在以下位置创建：

- 此步骤的上一步中创建了 **overcloud-baremetal-deployed.yaml** 文件。
- 在 [配置高级 OSD](#) 规格中创建了 **osd_spec.yaml** 文件。使用 **--osd-spec** 参数提供 OSD 规格是可选的。
- roles_data.yaml** 文件是 [在为 Red Hat Ceph Storage 设计节点](#) 时创建的。假设新节点被分配给此文件中的一个角色。

此命令的输出将是 **ceph_spec.yaml** 文件。

- 使用 **openstack overcloud ceph user enable** 命令在集群中的所有节点上创建 **ceph-admin** 用户。**ceph-admin** 用户必须在所有节点上存在，以便通过 Ceph 编配器启用对节点的 SSH 访问。

```

$ openstack overcloud ceph user enable ceph_spec.yaml

```



注意

使用上一步中创建的 **ceph_spec.yaml** 文件。

7. 使用 Red Hat [Ceph Storage Operations 指南中的使用服务规格部署 Ceph 守护进程](#)，将规格文件应用到 Red Hat Ceph Storage 集群。这个规格文件现在描述了添加了新节点的集群的操作状态。
8. 使用 **openstack overcloud deploy** 命令和更新的 `~/overcloud-baremetal-deployed.yaml` 文件。

```
$ openstack overcloud deploy --templates \
  -e /usr/share/openstack-tripleo-heat-templates/environments/cephadm/cephadm.yaml \
  -e deployed_ceph.yaml
  -e overcloud-baremetal-deploy.yaml
```

15.2. 缩减并替换 RED HAT CEPH STORAGE 节点

在某些情况下，您可能需要缩减 Red Hat Ceph Storage 集群或替换 Red Hat Ceph Storage 节点。在这两种情况下，您必须禁用并重新平衡您要从 overcloud 中删除的 Red Hat Ceph Storage 节点，以防止数据丢失。



流程

如果 Red Hat Ceph Storage 集群没有丢失 OSD 的容量，请不要继续这个过程。

1. 以 **tripleo-admin** 用户身份登录 overcloud Controller 节点。
2. 使用 **sudo cephadm shell** 命令启动 Ceph shell。
3. 使用 **ceph osd tree** 命令识别服务器要删除的 OSD。
在以下示例中，我们需要识别 **ceph-2** 主机的 OSD。

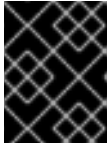
```
[ceph: root@oc0-controller-0 /]# ceph osd tree
ID CLASS WEIGHT  TYPE NAME        STATUS REWEIGHT PRI-AFF
-1      0.58557 root default
-7      0.19519 host ceph-2
  5  hdd 0.04880  osd.5      up    1.00000 1.00000
  7  hdd 0.04880  osd.7      up    1.00000 1.00000
  9  hdd 0.04880  osd.9      up    1.00000 1.00000
 11  hdd 0.04880  osd.11     up    1.00000 1.00000
```

4. 将 Ceph 集群规格导出到 YAML 文件。

```
[ceph: root@oc0-controller-0 /]# ceph orch ls --export > spec.yaml
```

5. 编辑导出的规格文件，以便适用的主机从 **service-type: osd hosts** 列表中移除，且适用的主机将移除 **placement: hosts** 值。
6. 保存编辑的文件。
7. 应用修改后的 Ceph 规格文件。

```
[ceph: root@oc0-controller-0 /]# ceph orch apply -i spec.yaml
```



重要

如果您在删除 OSD 前没有导出和导入 Ceph 规范文件，Ceph 管理器将尝试重新创建 OSD。

8. 使用 `ceph orch osd rm --zap <osd_list>` 命令来移除 OSD。

```
[ceph: root@oc0-controller-0 /]# ceph orch osd rm --zap 5 7 9 11
Scheduled OSD(s) for removal
[ceph: root@oc0-controller-0 /]# ceph orch osd rm status
OSD_ID HOST STATE PG_COUNT REPLACE FORCE DRAIN_STARTED_AT
7 ceph-2 draining 27 False False 2021-04-23 21:35:51.215361
9 ceph-2 draining 8 False False 2021-04-23 21:35:49.111500
11 ceph-2 draining 14 False False 2021-04-23 21:35:50.243762
```

9. 使用 `ceph orch osd status` 命令，检查移除 OSD 的状态。

```
[ceph: root@oc0-controller-0 /]# ceph orch osd rm status
OSD_ID HOST STATE PG_COUNT REPLACE FORCE DRAIN_STARTED_AT
7 ceph-2 draining 34 False False 2021-04-23 21:35:51.215361
11 ceph-2 draining 14 False False 2021-04-23 21:35:50.243762
```



警告

在这个命令返回任何结果前，不要继续下一步。

10. 使用 `ceph orch host drain <HOST>` 命令来排空任何剩余的守护进程。

```
[ceph: root@oc0-controller-0 /]# ceph orch host drain ceph-2
```

11. 使用 `ceph orch host rm <HOST>` 命令来移除主机。

```
[ceph: root@oc0-controller-0 /]# ceph orch host rm ceph-2
```

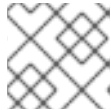
12. 结束 Ceph shell 会话。
13. 从 `tripleo-admin` 帐户注销。
14. 以 `stack` 用户身份登录 `undercloud` 节点。
15. 使用以下方法修改 `~/overcloud-baremetal-deploy.yaml`：
 - 减少角色中的 `count` 属性，以缩减。
 - 为被取消置备的每个节点添加一个 `instances` 条目。每个条目都必须包含以下内容：
 - `baremetal` 节点的名称。
 - 分配给该节点的主机名。

- **provisioned: false** 值。
以下示例将删除 **overcloud-compute-1** 节点。

```
- name: Compute
  count: 1
  instances:
    - hostname: overcloud-compute-0
      name: node10
      # Removed from deployment due to disk failure
      provisioned: false
    - hostname: overcloud-compute-1
      name: node11
```

16. 使用 **openstack overcloud node delete** 命令删除该节点。

```
openstack overcloud node delete \
--stack overcloud \
--baremetal-deployment ~/overcloud-baremetal-deploy.yaml
```



注意

在删除节点前，将提供一个确认提示要删除的节点列表。

注意

如果缩减 Ceph 集群是临时的，并且删除的节点稍后会被恢复，扩展操作在之前设置了 **provisioned: false** 的节点上增加 **count**，并设置 **provisioned: true** 如果节点永远不会重复使用，则可以无限期设置 **provisioned: false**，扩展操作可以指定新的 instances 条目。

+ 以下文件示例提供了每个实例的一些示例。

+

```
- name: Compute
  count: 2
  instances:
    - hostname: overcloud-compute-0
      name: node10
      # Removed from deployment due to disk failure
      provisioned: false
    - hostname: overcloud-compute-1
      name: node11
    - hostname: overcloud-compute-2
      name: node12
```

第 16 章 替换失败的磁盘

如果 Ceph Storage 集群中的磁盘失败，您可以替换它。

16.1. 替换磁盘

有关替换失败磁盘的信息，请参阅 *Red Hat Ceph Storage 安装指南* 中的 [添加 OSD](#)。