



# Red Hat OpenStack Platform 17.1

## 部署超融合基础架构

了解并配置 Red Hat OpenStack Platform overcloud 上的超融合基础架构



# Red Hat OpenStack Platform 17.1 部署超融合基础架构

---

了解并配置 Red Hat OpenStack Platform overcloud 上的超融合基础架构

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

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

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

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

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

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

All other trademarks are the property of their respective owners.

## 摘要

本文档描述了超融合的 Red Hat OpenStack Platform 实施，它将在同一主机上并置计算和 Ceph 存储服务。

# 目录

使开源包含更多 .....	4
对红帽文档提供反馈 .....	5
<b>第 1 章 配置和部署 RED HAT OPENSTACK PLATFORM 超融合基础架构 .....</b>	<b>6</b>
1.1. HYPERCONVERGED 基础架构概述 .....	6
<b>第 2 章 部署 HCI 硬件 .....</b>	<b>7</b>
2.1. 清理 CEPH STORAGE 节点磁盘 .....	7
2.2. 注册节点 .....	7
2.3. 验证可用的 RED HAT CEPH STORAGE 软件包 .....	10
2.4. 为 HCI 环境部署软件镜像 .....	10
2.5. 为 HCI 设计节点 .....	11
2.6. 为多磁盘 CEPH 集群定义根磁盘 .....	13
<b>第 3 章 为 HCI 配置 RED HAT CEPH STORAGE 集群 .....</b>	<b>16</b>
3.1. 部署先决条件 .....	16
3.2. OPENSTACK OVERCLOUD CEPH DEPLOY 命令 .....	16
3.3. HCI 的 CEPH 配置覆盖 .....	16
3.4. 配置时间同步 .....	17
3.5. 配置顶级域后缀 .....	18
3.6. 配置 RED HAT CEPH STORAGE 集群名称 .....	19
3.7. 使用网络数据文件配置网络选项 .....	20
3.8. 使用配置文件配置网络选项 .....	20
3.9. 为 OSD 配置 CRUSH 层次结构 .....	21
3.10. 配置 CEPH 服务放置选项 .....	22
3.11. 为 CEPH 节点配置 SSH 用户选项 .....	23
3.12. 访问 CEPH STORAGE 容器 .....	24
<b>第 4 章 为 HCI 自定义 RED HAT CEPH STORAGE 集群 .....</b>	<b>26</b>
4.1. 配置选项 .....	26
4.2. 生成服务规格（可选） .....	26
4.3. RED HAT OPENSTACK PLATFORM WITH RED HAT CEPH STORAGE 的 CEPH 容器 .....	27
4.4. 配置高级 OSD 规格 .....	27
4.5. 从特定于节点的覆盖迁移 .....	28
4.6. 启用 CEPH 在线加密 .....	28
<b>第 5 章 为 HCI 自定义存储服务 .....</b>	<b>29</b>
5.1. 为 HCI 配置计算服务资源 .....	29
5.2. 配置自定义环境文件 .....	29
5.3. 启用 CEPH 元数据服务器 .....	30
5.4. CEPH 对象网关对象存储 .....	30
5.5. RED HAT OPENSTACK PLATFORM 对象存储的部署选项 .....	31
5.6. 配置块存储备份服务以使用 CEPH .....	32
5.7. 为 CEPH 节点配置多个绑定接口 .....	32
5.8. 为 HCI 启动 OVERCLOUD 部署 .....	32
<b>第 6 章 验证 HCI 配置 .....</b>	<b>34</b>
6.1. 验证 HCI 配置 .....	34
<b>第 7 章 扩展超融合节点 .....</b>	<b>35</b>
7.1. 在 HCI 环境中扩展超融合节点 .....	35
7.2. 在 HCI 环境中缩减超融合节点 .....	35

<b>附录 A. 附加信息</b> .....	<b>36</b>
A.1. 配置指南	36
A.2. 配置超融合基础架构环境的指南和资源	37



## 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 配置和部署 RED HAT OPENSTACK PLATFORM 超融合基础架构

## 1.1. HYPERCONVERGED 基础架构概述

Red Hat OpenStack Platform (RHOSP)超融合基础架构(HCI)由超融合节点组成。在 RHOSP HCI 中，计算和存储服务在这些超融合节点上并置，以优化资源使用。您可以部署仅使用超融合节点的 overcloud，或使用普通 Compute 和 Red Hat Ceph Storage 节点混合使用超融合节点。



### 注意

您必须使用 Red Hat Ceph Storage 作为存储提供程序。

### 提示

使用 BlueStore 作为 HCI 部署的后端，以使用 BlueStore 内存处理功能。

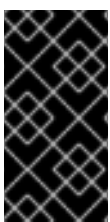
超融合基础架构使用部署 Red Hat Ceph 和 OpenStack 中的部署流程的一种变体构建，如 [部署 Red Hat Ceph 和 OpenStack 与 director 一起部署](#)。在这个部署场景中，RHOSP director 部署您的云环境，director 会调用 overcloud 和 Red Hat Ceph Storage。您可以管理和扩展 Ceph 集群本身与 overcloud 配置分开。



### 重要

RHOSP HCI 环境不支持实例 HA。要在 RHOSP HCI 环境中使用实例 HA，您必须使用 ComputeInstanceHA 角色指定 Compute 节点的子集以使用实例 HA。Red Hat Ceph Storage 服务不能托管在托管 Instance HA 的 Compute 节点上。

Red Hat OpenStack Platform 17.1 仅支持用于新部署的 Red Hat Ceph Storage 6。在新的部署场景中不支持 Red Hat Ceph Storage 5。



### 重要

支持的超融合基础架构环境中的所有 HCI 节点都必须使用与 Red Hat OpenStack Platform 控制器使用的版本相同的 Red Hat Enterprise Linux 版本。如果要在同一超融合基础架构环境中的 HCI 节点上使用混合状态的多个 [Red Hat Enterprise 版本](#)，请联系[红帽客户体验和参与度](#)团队，讨论支持例外。

有关 HCI 配置指导，请参阅 [配置指南](#)。

## 第 2 章 部署 HCI 硬件

本节介绍超融合节点的准备和配置的步骤和信息。

### 先决条件

- 您已阅读了 *使用 director 部署 Red Hat Ceph 和 OpenStack* 中的 [部署 overcloud 和 Red Hat Ceph Storage](#)。

### 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"
  }
]
}

```

- 保存新文件。
- 初始化 stack 用户：

```
$ source ~/stackrc
```

- 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。

- 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 HCI 环境部署软件镜像

为 HCI 环境配置的节点必须使用 **overcloud-hardened-uefi-full.qcow2** 软件镜像。使用此软件镜像需要 Red Hat OpenStack Platform (RHOSP) 订阅。

### 流程

- 打开 **/home/stack/templates/overcloud-baremetal-deploy.yaml** 文件。
- 为需要 **overcloud-hardened-uefi-full** 镜像的节点添加或更新 **image** 属性。您可以将镜像设置为在特定节点上使用，或设置为使用特定角色的所有节点：

#### 特定节点

```

- name: Ceph
  count: 3
  instances:
    - hostname: overcloud-ceph-0

```

```

name: node00
image:
  href: file:///var/lib/ironic/images/overcloud-minimal.qcow2
- hostname: overcloud-ceph-1
  name: node01
  image:
    href: file:///var/lib/ironic/images/overcloud-hardened-uefi-full.qcow2
- hostname: overcloud-ceph-2
  name: node02
  image:
    href: file:///var/lib/ironic/images/overcloud-hardened-uefi-full.qcow2

```

### 为特定角色配置所有节点

```

- name: ComputeHCI
  count: 3
  defaults:
    image:
      href: file:///var/lib/ironic/images/overcloud-hardened-uefi-full.qcow2
  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 deploy` 命令。

## 2.5. 为 HCI 设计节点

要为 HCI 指定节点，您必须创建一个新角色文件来配置 `ComputeHCI` 角色，并使用 `ComputeHCI` 的资源类配置裸机节点。

### 步骤

1. 以 `stack` 用户的身份登录 `undercloud`。
2. 查找 `stackrc` 凭证文件：

```
[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** 和 **ComputeHCI** 角色：

```
(undercloud)$ openstack overcloud roles generate Controller ComputeHCI -o
~/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	<b>Role: ComputeHCI</b>
角色名称	<b>Name: ComputeHCI</b>
<b>description</b>	<b>HCI 角色</b>
<b>HostnameFormatDefault</b>	<b>%stackname%-novaceph-%index%</b>
<b>deprecated_nic_config_name</b>	<b>ceph.yaml</b>

5. 将 overcloud 的 ComputeHCI 节点添加到节点定义模板 **node.json** 或 **node.yaml** 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 HCI 资源类标记您要为 HCI 指定的每个裸机节点：

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

将 **<node>** 替换为裸机节点的 ID。

8. 将 **ComputeHCI** 角色添加到您的 **/home/stack/templates/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: ComputeHCI
  count: 1
  defaults:
    resource_class: baremetal.HCI
```

9. 打开 **baremetal.yaml** 文件，并确保它包含 HCI 所需的网络配置。以下是配置示例：

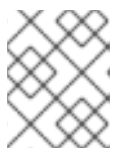
```
- name: ComputeHCI
  count: 3
  hostname_format: compute-hci-%index%
  defaults:
    profile: ComputeHCI
    network_config:
```



```

template: /home/stack/templates/three-nics-vlans/compute-hci.j2
networks:
- network: ctlplane
  vif: true
- network: external
  subnet: external_subnet
- network: internalapi
  subnet: internal_api_subnet01
- network: storage
  subnet: storage_subnet01
- network: storage_mgmt
  subnet: storage_mgmt_subnet01
- network: tenant
  subnet: tenant_subnet01

```



### 注意

**ComputeHCI** 角色中的网络配置包含 **storage\_mgmt** 网络。CephOSD 节点使用此网络进行冗余数据副本。**Compute** 角色的网络配置不包含此网络。

如需更多信息，[请参阅配置裸机置备服务](#)。

#### 10. 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

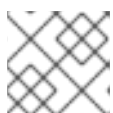
```

#### 11. 在一个单独的终端中监控置备进度。

```

(undercloud)$ watch openstack baremetal node list

```



### 注意

默认情况下，**watch** 命令每 2 秒续订一次。**n** 选项将续订计时器设置为不同的值。

#### 12. 要停止 监视 进程，请输入 **Ctrl-c**。

#### 13. 验证：置备成功后，节点状态从 **available** 变为 **active**。

### 其他资源

- 有关注册节点的更多信息，请参阅 [安装和管理 Red Hat OpenStack Platform 指南](#) 中的 [为 overcloud 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](#)。
- 有关 **ComputeHCI** 角色和 **storage\_mgmt** 网络中的网络配置的更多信息，[请参阅配置裸机置备服务](#)。

## 2.6. 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。

## 2.6.1. 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。

## 第 3 章 为 HCI 配置 RED HAT CEPH STORAGE 集群

本章论述了如何为 HCI 环境配置和部署 Red Hat Ceph Storage 集群。

### 3.1. 部署先决条件

在尝试配置和部署 Red Hat Ceph Storage 集群前确认已执行以下内容：

- 使用裸机置备服务(ironic)置备裸机实例及其网络。有关置备裸机实例的更多信息，[请参阅配置裸机置备服务](#)。

### 3.2. 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。

### 3.3. HCI 的 CEPH 配置覆盖

标准格式初始化文件是 Ceph 集群配置的一个选项。然后，此初始化文件用于使用 **cephadm bootstrap --config <file\_name>** 或 **openstack overcloud ceph deploy --config <file\_name>** 命令配置 Ceph 集群。

在超融合节点上并置 Ceph OSD 和计算服务会面临 Red Hat Ceph Storage 和 Compute 服务之间的资源争用风险。这是因为服务不知道 colocation。资源争用可能会导致服务降级，这可偏移超线程的优势。

可以使用初始化文件调优资源分配来管理资源争用。下面创建一个名为 **initial-ceph.conf** 的初始化文件，然后使用 **openstack overcloud ceph deploy** 命令配置 HCI 部署。

```
$ cat <<EOF > initial-ceph.conf
[osd]
osd_memory_target_autotune = true
osd_numa_auto_affinity = true
[mgr]
mgr/cephadm/autotune_memory_target_ratio = 0.2
EOF
$ openstack overcloud ceph deploy --config initial-ceph.conf
```

**osd\_memory\_target\_autotune** 选项设为 **true**，以便 OSD 守护进程根据 **osd\_memory\_target** 配置选项调整其内存消耗。**autotune\_memory\_target\_ratio** 默认为 **0.7**。这表示系统中 RAM 总量的 70% 是来自非自动 tuned Ceph 守护进程所消耗的任何内存的起点。然后，剩余的内存被 OSD 划分，假设所有 OSD 都 **osd\_memory\_target\_autotune** 设为 **true**。对于 HCI 部署，将 **mgr/cephadm/autotune\_memory\_target\_ratio** 设置为 0.2，以确保更多内存可用于 Compute 服务。**0.2** 值是一个繁琐的起点。部署后，如果需要，使用 **ceph** 命令更改此值。

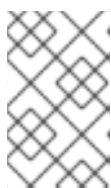
两个 NUMA 节点系统可以在一个 NUMA 节点上托管延迟敏感的 Nova 工作负载，并在其他 NUMA 节点上托管 Ceph OSD 工作负载。要将 Ceph OSD 配置为使用不供计算工作负载使用的特定 NUMA 节点，请使用以下 Ceph OSD 配置之一：

- **osd\_numa\_node** 将关联性设置为 numa 节点
- **osd\_numa\_auto\_affinity** 自动将关联性设置为存储和网络匹配的 NUMA 节点

如果 NUMA 节点和磁盘控制器上都有网络接口是 NUMA 节点 0，请将 NUMA 节点 0 上的网络接口用于存储网络，并在 NUMA 节点 0 上托管 Ceph OSD 工作负载。在 NUMA 节点 1 上托管 Nova 工作负载，并使其在 NUMA 节点 1 上使用网络接口。将 `osd_numa_auto_affinity` 设置为 `true` 来实现此配置。或者，`osd_numa_node` 可以直接设置为 `0`，并且没有为 `osd_numa_auto_affinity` 设置值，使其默认为 `false`。

当 OSD 离线导致超融合集群回填时，回填过程可能会减慢。在速度较慢的恢复过程中，回填活动对并置计算工作负载的影响较少。Red Hat Ceph Storage 有以下默认值来控制回填活动的速度：

- `osd_recovery_op_priority = 3`
- `osd_max_backfills = 1`
- `osd_recovery_max_active_hdd = 3`
- `osd_recovery_max_active_ssd = 10`



### 注意

不需要在初始化文件中传递这些默认值，因为它们是默认值。如果 initial 配置需要默认值，请在部署前使用所需值将它们添加到初始化文件中。部署后，使用命令 `'ceph config set osd'`。

## 3.4. 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。

### 3.4.1. 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"
```

### 3.4.2. 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"
```

### 3.4.3. 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
```

## 3.5. 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"
```

### 3.6. 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 5.8 节“为 HCI 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'
```

### 3.7. 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**。

### 3.8. 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
```

## 3.9. 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.0		up	1.00000	1.00000	
-5	0.00980	rack	r1				
-4	0.00980	host	ceph-node-01				
1	hdd 0.00980	osd.1		up	1.00000	1.00000	
-7	0.00980	rack	r2				
-6	0.00980	host	ceph-node-02				
2	hdd 0.00980	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](#)。

### 3.10. 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
```

### 3.11. 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。

#### 3.11.1. 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**。

#### 3.11.2. 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。

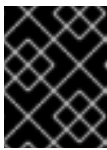
## 流程

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。

## 3.12. 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](#)

### 3.12.1. 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。

### 3.12.2. 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。

## 第 4 章 为 HCI 自定义 RED HAT CEPH STORAGE 集群

Red Hat OpenStack Platform (RHOSP) director 使用默认配置来部署容器化 Red Hat Ceph Storage。您可以通过覆盖默认设置来自定义 Ceph Storage。

### 先决条件

- 服务器已部署，并且配置了其存储网络。
- **openstack overcloud node provision -o ~/deployed\_metal.yaml ...** 将部署的裸机文件作为输出。

### 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：

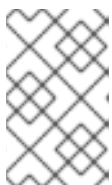
```
[global]
osd_crush_chooseleaf type = 0
log_file = /var/log/ceph/$cluster-$type.$id.log

[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）

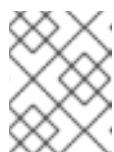
Red Hat Ceph Storage 集群服务规格是一个 YAML 文件，用于描述 Ceph Storage 服务的部署。在部署 Ceph Storage 集群前，它由 **tripleo** 自动生成。它通常不必单独生成。

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 为 HCI 自定义存储服务

Red Hat OpenStack Platform (RHOSP) director 提供了必要的 heat 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，为 **openstack overcloud ceph deploy** 部署的 Ceph 集群添加额外的配置。

有关 RHOSP 中容器化服务的更多信息，请参阅使用 [CLI 工具配置基本 overcloud](#)，[请参阅使用 director 安装和管理 Red Hat OpenStack Platform](#)。

### 5.1. 为 HCI 配置计算服务资源

在超融合节点上并置 Ceph OSD 和计算服务会面临 Red Hat Ceph Storage 和 Compute 服务之间的资源争用风险。这是因为服务不知道 colcation。资源争用可能会导致服务降级，这可偏移超线程的优势。

配置计算服务使用的资源可以缓解资源争用，并提高 HCI 性能。

#### 流程

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

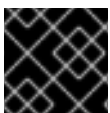
```
$ source ~/stackrc
```

3. 将 **NovaReservedHostMemory** 参数添加到 **ceph-overrides.yaml** 文件。以下是用法示例。

```
parameter_defaults:
  ComputeHCIParameters:
    NovaReservedHostMemory: 75000
```

**NovaReservedHostMemory** 参数覆盖 `/etc/nova/nova.conf` 中的 **reserved\_host\_memory\_mb** 的默认值。此参数设置为停止向虚拟机提供 Ceph OSD 需要的内存的 Nova 调度程序。

除了虚拟机监控程序的默认保留内存外，上面的示例还为每个主机每 10 个 OSD 保留 5 GB 的 OSD。在 IOPS 优化的集群中，您可以通过为每个 OSD 保留更多内存来提高性能。5 GB 编号作为起点提供，您可以根据需要进一步优化。



#### 重要

使用 **openstack overcloud deploy** 命令时包含此文件。

### 5.2. 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。

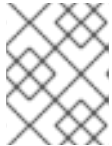
#### 流程

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:
  CinderEnableiscsiBackend: 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。

**5.3. 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。

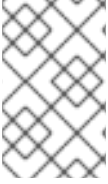
如需更多信息和流程, 请参阅 [安装和管理 Red Hat OpenStack Platform 指南](#)中的[置备 overcloud 网络](#)。

## 5.8. 为 HCI 启动 OVERCLOUD 部署

要实现您对 Red Hat OpenStack Platform (RHOSP)环境所做的更改, 您必须部署 overcloud。

### 先决条件

- 在 undercloud 安装前, 在 **undercloud.conf** 文件中设置 **generate\_service\_certificate=false**。否则, 您必须在 overcloud 上配置 SSL/TLS, 如 [强化 Red Hat OpenStack Platform 中的 overcloud 公共端点上启用 SSL/TLS](#) 所述。



## 注意

如果要在 overcloud 部署期间添加 Ceph 仪表盘，请参阅 [部署 Red Hat Ceph Storage 和 Red Hat OpenStack Platform](#) 中的 [将 Red Hat Ceph Storage 仪表盘添加到 overcloud 部署](#) 中。

## 流程

- 部署 overcloud。部署命令需要额外的参数，例如：

```
$ 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 \
--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。
- **-e /usr/share/openstack-tripleo-heat-templates/environments/cephadm/ceph-mds.yaml** - 启用 Ceph 元数据服务器。
- **-e /usr/share/openstack-tripleo-heat-templates/environments/cinder-backup.yaml** - 启用块存储备份服务。
- **-e /home/stack/templates/storage-config.yaml** - 添加包含自定义 Ceph Storage 配置的环境文件。
- **-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 `openstack overcloud ceph deploy` 命令的输出一样。
- **--ntp-server pool.ntp.org** - 设置 NTP 服务器。



## 注意

如需完整的选项列表，请运行 `openstack help overcloud deploy` 命令。

## 其他资源

- 如需更多信息，请参阅 [安装和管理 Red Hat OpenStack Platform](#) 指南中的使用 [CLI 工具配置基本 overcloud](#)。

## 第 6 章 验证 HCI 配置

部署完成后，验证 HCI 环境是否已正确配置。

### 6.1. 验证 HCI 配置

部署 HCI 环境后，使用指定的配置验证部署是否成功。

#### 流程

1. 启动 ceph shell。
2. 确认 NUMA 和内存目标配置：

```
[ceph: root@oc0-controller-0 /]# ceph config dump | grep numa
osd                                advanced osd_numa_auto_affinity      true
[ceph: root@oc0-controller-0 /]# ceph config dump | grep autotune
osd                                advanced osd_memory_target_autotune  true
[ceph: root@oc0-controller-0 /]# ceph config get mgr
mgr/cephadm/autotune_memory_target_ratio
0.200000
```

3. 确认特定的 OSD 配置：

```
[ceph: root@oc0-controller-0 /]# ceph config get osd.11 osd_memory_target
4294967296
[ceph: root@oc0-controller-0 /]# ceph config get osd.11 osd_memory_target_autotune
true
[ceph: root@oc0-controller-0 /]# ceph config get osd.11 osd_numa_auto_affinity
true
```

4. 确认特定的 OSD 回填配置：

```
[ceph: root@oc0-controller-0 /]# ceph config get osd.11 osd_recovery_op_priority
3
[ceph: root@oc0-controller-0 /]# ceph config get osd.11 osd_max_backfills
1
[ceph: root@oc0-controller-0 /]# ceph config get osd.11 osd_recovery_max_active_hdd
3
[ceph: root@oc0-controller-0 /]# ceph config get osd.11 osd_recovery_max_active_ssd
10
```

5. 确认 Compute 节点上的 **reserved\_host\_memory\_mb** 配置。

```
$ sudo podman exec -ti nova_compute /bin/bash
bash-5.1$ grep reserved_host_memory_mb /etc/nova/nova.conf
```

## 第 7 章 扩展超融合节点

要扩展或缩减 HCI 节点，则应用与扩展 Compute 节点或 Red Hat Ceph Storage 节点的不同原则和方法。

### 7.1. 在 HCI 环境中扩展超融合节点

要在 HCI 环境中扩展超融合节点，请遵循扩展非超融合节点的步骤。有关更多信息，请参阅 [向 overcloud 添加节点](#)。



#### 注意

在标记新节点时，请记住使用右侧类型。

### 7.2. 在 HCI 环境中缩减超融合节点

要在 HCI 环境中缩减超融合节点，您必须重新平衡 HCI 节点上的 Ceph OSD 服务，从 HCI 节点迁移实例，并从 overcloud 中删除 Compute 节点。

#### 流程

1. 在 HCI 节点上禁用并重新平衡 Ceph OSD 服务。此步骤是必需的，因为当您删除 HCI 或 Red Hat Ceph Storage 节点时，director 不会自动重新平衡 Red Hat Ceph Storage 集群。有关更多信息，请参阅 [部署 Red Hat Ceph Storage 和 Red Hat OpenStack Platform 中的扩展 Ceph Storage 集群](#)以了解更多信息。
2. 从 HCI 节点迁移实例。如需更多信息，请参阅配置 [计算服务创建指南中的在 Compute 节点之间迁移虚拟机](#)。
3. 从 overcloud 移除 Compute 节点。如需更多信息，请参阅 [删除 Compute 节点](#)。

## 附录 A. 附加信息

### A.1. 配置指南

以下配置指南旨在为创建超融合基础架构环境提供框架。这个指南并不适用于为每个 Red Hat OpenStack Platform 安装提供确定的配置参数。请联系红帽 [客户体验与参与团队](#) 以获取适合您的特定环境的具体指导和建议。

- [集群大小并扩展](#)
- [容量规划和大小](#)

#### A.1.1. 集群大小并扩展

[Red Hat Ceph Storage 硬件指南](#) 为优化、吞吐量优化、成本和容量优化的 Ceph 部署场景提供建议。按照最适合您的部署场景的建议，并添加支持计算工作负载所需的 NIC、CPU 和 RAM。

一个由 7 个节点组成的最佳的、小型配置。除非要求在您的环境中优化 IOPS 的性能，并且您使用的是所有闪存存储，否则应使用吞吐量优化部署场景。

可以有三个节点 Ceph Storage 集群配置。在这个配置中，您应该：

- 使用所有闪存存储。
- 在 `ceph.conf` 文件中，将 `replica_count` 参数设置为 3。
- 在 `ceph.conf` 文件中，将 `min_size` 参数设置为 2。

如果节点在这个配置中保留服务，则 IOPS 将继续。要保留 3 个数据副本，对第三个节点的复制会排队，直到它返回到服务。然后，数据被回填到第三个节点。



#### 注意

最多 64 节点的 HCI 配置已被测试。一些 HCI 环境示例已被记录为 128 个节点。大型集群（如这些集群）可通过支持例外和咨询服务参与来考虑。请联系红帽 [客户体验与参与团队](#) 以获得指导。

有两个 NUMA 节点的部署可以在一个 NUMA 节点上托管对一个 NUMA 节点和 Ceph OSD 服务的延迟敏感的 Compute 工作负载。如果两个节点上都存在网络接口，并且磁盘控制器位于节点 0 上，请将节点 0 上的网络接口用于 Storage 网络，并在节点 0 上托管 Ceph OSD 工作负载。在节点 1 上托管计算工作负载，并将其配置为在节点 1 上使用网络接口。当为您的部署获取硬件时，请注意，哪些 NIC 将使用哪些节点并尝试在存储和工作负载之间分割它们。

#### A.1.2. 容量规划和大小

[Red Hat Ceph Storage 硬件指南中定义的优化 Ceph](#) 解决方案为大多数不需要优化 IOPS 的部署提供均衡解决方案。除了解决方案提供的配置指南外，在创建环境时请注意以下几点：

- 每个 OSD 分配的 5 GB RAM 可确保 OSD 有足够的操作内存。确保您的硬件可以支持此要求。
- CPU 速度应当与正在使用的存储介质匹配。更快的存储介质的优点（如 SSD）可能会由于支持它们的 CPU 太慢。同样，快速 CPU 可以被更快的存储介质更有效地使用。平衡 CPU 和存储介质速度，以便不再成为其他的瓶颈。



## A.2. 配置超融合基础架构环境的指南和资源

以下指南包含额外的信息和流程，有助于配置超融合基础架构环境。

- [将 Red Hat Ceph 和 OpenStack 与 director 一起部署](#)
  - 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 Ceph 集群的说明。
- [使用 director 安装和管理 Red Hat OpenStack Platform](#)
  - 本指南提供有关 Red Hat OpenStack Platform 环境的端到端部署。其中包括安装 director、规划您的环境以及使用 director 创建 OpenStack 环境。
- [配置 Red Hat OpenStack Platform 网络](#)
  - 本指南提供有关 Red Hat OpenStack Platform 网络任务的详细信息。
- [配置持久性存储](#)
  - 本指南详细介绍了在 Red Hat OpenStack Platform 环境中使用和管理持久性存储的不同程序。它还包含配置和管理每种持久性存储类型的相应 OpenStack 服务的步骤。
- [配置裸机置备服务](#)
  - 本指南详细介绍了 Red Hat OpenStack Platform 环境的 overcloud 中的裸机置备服务安装和配置，以便为云用户置备和管理物理机器。
- [强化 Red Hat OpenStack Platform](#)
  - 本指南提供有关强化 Red Hat OpenStack Platform 环境安全性的良好实践建议和概念性信息。
- [发行注记](#)
  - 本文档概述此 Red Hat OpenStack Platform 发行版本的主要功能、功能增强和已知问题。