



Red Hat Ceph Storage 3

Ceph ブロックデバイスから OpenStack へのガイド

OpenStack のバックエンドとして Ceph を使用するように、Ceph、QEMU、libvirt、OpenStack を設定

Red Hat Ceph Storage 3 Ceph ブロックデバイスから OpenStack へのガイド

OpenStack のバックエンドとして Ceph を使用するように、Ceph、QEMU、libvirt、OpenStack を設定

Enter your first name here. Enter your surname here.

Enter your organisation's name here. Enter your organisational division here.

Enter your email address here.

法律上の通知

Copyright © 2022 | You need to change the HOLDER entity in the en-US/Ceph_Block_Device_to_OpenStack_Guide.ent file |.

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.

概要

本ガイドでは、Glance、Cinder、Cinder Backup、および Nova のバックエンドに Ceph を使用するように OpenStack および Ceph を設定する方法を説明します。

目次

前書き	3
第1章 CEPH プールの作成	5
第2章 CEPH クライアントのインストールおよび設定	6
2.1. CEPH 設定ファイルの OPENSTACK ノードへのコピー	6
2.2. CEPH クライアント認証の設定	6
第3章 OPENSTACK が CEPH を使用するように設定する手順	8
3.1. CINDER の設定	8
3.2. CINDER バックアップの設定	9
3.3. GLANCE の設定	11
3.4. NOVA の設定	12
3.5. OPENSTACK サービスの再起動	12

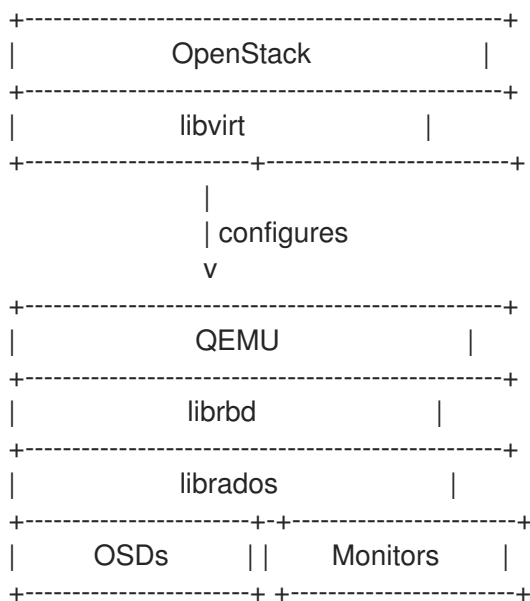
前書き

Red Hat OpenStack Platform director は、Glance、Cinder、Cinder Backup、および Nova のバックエンドとして Ceph を使用するための 2 つの方法を提供します。

1. **OpenStack は Ceph クラスターの作成**：OpenStack director は Ceph クラスターを起動することができます。これには、Ceph OSD 用のテンプレートを設定する必要があります。OpenStack は Ceph ノードのインストールと設定を処理します。このシナリオでは、OpenStack は OpenStack コントローラーノードで Ceph モニターをインストールします。
2. **OpenStack director から既存の Ceph クラスターへの接続**：OpenStack director (Red Hat OpenStack Platform director バージョン 9 以降) は、Ceph モニターに接続して、OpenStack のバックエンドとして使用するように Ceph クラスターを設定することができます。

前述の方法は、インストールと設定の多くを自動的に処理してくれるため、OpenStack のバックエンドとして Ceph を設定するための好ましい方法です。詳細は、[Red Hat OpenStack Platform](#) のドキュメントを参照してください。

本書では、Ceph をバックエンドとして使用するように Ceph、QEMU、libvirt、および OpenStack を設定する手順を詳しく説明します。本書は、Red Hat OpenStack Platform director を使用しないユーザーを対象としています。



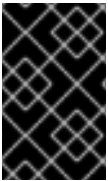
注記

OpenStack のバックエンドとして Ceph ブロックデバイスを使用するには、稼働中の Ceph ストレージクラスターと少なくとも 1 つの OpenStack ノードが必要です。

OpenStack の 3 つの部分は、Ceph のブロックデバイスと統合されます。

- **イメージ**: OpenStack Glance は仮想マシンのイメージを管理します。イメージを変更することはできません。OpenStack はイメージをバイナリーオブジェクトとして扱い、それに応じてイメージをダウンロードします。
- **ボリューム**: ボリュームはブロックデバイスです。OpenStack では、仮想マシンを起動したり、実行中の仮想マシンにボリュームを接続したりするためにボリュームを使用します。OpenStack は Cinder サービスを使用してボリュームを管理します。Ceph は、OpenStack Cinder および Cinder Backup のブラックリストとしての役割を果たすことができます。

- **ゲストディスク:** ゲストディスクは、ゲストオペレーティングシステムのディスクです。デフォルトでは、仮想マシンのブート時に、ディスクはハイパーバイザーのファイルシステム上のファイルとして表示されます（通常 `/var/lib/nova/instances/<uuid>/` 内）。OpenStack Glance は、Ceph ブロックデバイスにイメージを格納することができ、イメージの Copy-on-Write クローンを使用して Cinder を使用して仮想マシンをブートすることができます。



重要

Ceph は、仮想マシンのディスクをホストする QCOW2 をサポートしません。Ceph で仮想マシンを起動するには（一時バックエンドまたはボリュームからブート）、Glance イメージの形式を RAW にする必要があります。

OpenStack では、イメージ、ボリューム、またはゲストディスクの仮想マシンに Ceph を使用できません。3 つとも使用する必要はありません。

第1章 CEPH プールの作成

デフォルトでは、Ceph ブロックデバイスは **rd** プールを使用します。利用可能なプールを使用できません。以下の例では、Cinder、Cinder バックアップ、Glance、および Nova にそれぞれプールを作成します。Ceph クラスターが稼働中であることを確認してから、プールを作成します。

Red Hat は、[Ceph Placement Group's per Pool Calculator](#) を使用して、プールに適した配置グループ数を計算することを推奨します。プールの作成に関する詳細は、Red Hat Ceph **Storage 3** の『ストレージ戦略』ガイドの「プール」の章を参照してください。以下の例では、**128** が配置グループの数になります。

```
# ceph osd pool create volumes 128
# ceph osd pool create backups 128
# ceph osd pool create images 128
# ceph osd pool create vms 128
```

第2章 CEPH クライアントのインストールおよび設定

nova-compute、**cinder-backup**、および **cinder-volume** ノードには、Python バインディングとクライアントコマンドラインツールの両方が必要です。

```
# yum install python-rbd
# yum install ceph-common
```

glance-api ノードには、**librbd** 用の Python バインディングが必要です。

```
# yum install python-rbd
```

2.1. CEPH 設定ファイルの OPENSTACK ノードへのコピー

glance-api、**cinder-volume**、**nova-compute**、**cinder-backup** を実行するノードは Ceph クライアントとして機能します。それぞれに Ceph 設定ファイルが必要です。Ceph 設定ファイルを監視ノードから OSP ノードにコピーします。

```
# scp /etc/ceph/ceph.conf osp:/etc/ceph
```

2.2. CEPH クライアント認証の設定

Ceph monitor ノードから、Cinder、Cinder Backup、および Glance の新しいユーザーを作成します。

```
# ceph auth get-or-create client.cinder mon 'allow r' osd 'allow class-read object_prefix rbd_children,
allow rwx pool=volumes, allow rwx pool=vms, allow rx pool=images'

# ceph auth get-or-create client.cinder-backup mon 'allow r' osd 'allow class-read object_prefix
rbd_children, allow rwx pool=backups'

# ceph auth get-or-create client.glance mon 'allow r' osd 'allow class-read object_prefix rbd_children,
allow rwx pool=images'
```

client.cinder、**client.cinder-backup**、および **client.glance** のキーリングを適切なノードに追加し、それらの所有権を変更します。

```
# ceph auth get-or-create client.cinder | ssh {your-volume-server} sudo tee
/etc/ceph/ceph.client.cinder.keyring
# ssh {your-cinder-volume-server} chown cinder:cinder /etc/ceph/ceph.client.cinder.keyring

# ceph auth get-or-create client.cinder-backup | ssh {your-cinder-backup-server} tee
/etc/ceph/ceph.client.cinder-backup.keyring
# ssh {your-cinder-backup-server} chown cinder:cinder /etc/ceph/ceph.client.cinder-backup.keyring

# ceph auth get-or-create client.glance | ssh {your-glance-api-server} sudo tee
/etc/ceph/ceph.client.glance.keyring
# ssh {your-glance-api-server} chown glance:glance /etc/ceph/ceph.client.glance.keyring
```

nova-compute を実行しているノードには、**nova-compute** プロセスのキーリングファイルが必要です。

```
# ceph auth get-or-create client.cinder | ssh {your-nova-compute-server} tee
/etc/ceph/ceph.client.cinder.keyring
```

nova-compute を実行しているノードは、**libvirt** に **client.cinder** ユーザーの秘密鍵を保存する必要もあります。**libvirt** プロセスでは、Cinder からブロックデバイスを接続中にクラスターにアクセスする必要があります。**nova-compute** を実行しているノードに秘密鍵の一時コピーを作成します。

```
# ceph auth get-key client.cinder | ssh {your-compute-node} tee client.cinder.key
```

exclusive-lock 機能を使用する Ceph Block Device イメージがストレージクラスターに含まれている場合には、すべての Ceph Block Device ユーザーにクライアントをブラックリストに登録するパーミッションがあるようにしてください。

```
# ceph auth caps client.{ID} mon 'allow r, allow command "osd blacklist"' osd '{existing-OSD-user-
capabilities}'
```

コンピュータノードに戻ります。

```
# ssh {your-compute-node}
```

シークレットの UUID を生成し、後で **nova-compute** を設定するためにシークレットの UUID を保存します。

```
# uuidgen > uuid-secret.txt
```



注記

すべてのコンピュータノードで UUID は必要ありません。ただし、プラットフォームの整合性からは、同じ UUID を維持する方が適切です。

次に、コンピュータノードで秘密鍵を **libvirt** に追加し、キーの一時コピーを削除します。

```
cat > secret.xml <<EOF
<secret ephemeral='no' private='no'>
  <uuid>`cat uuid-secret.txt`</uuid>
  <usage type='ceph'>
    <name>client.cinder secret</name>
  </usage>
</secret>
EOF
```

```
# virsh secret-define --file secret.xml
# virsh secret-set-value --secret $(cat uuid-secret.txt) --base64 $(cat client.cinder.key) && rm
client.cinder.key secret.xml
```

第3章 OPENSTACK が CEPH を使用するように設定する手順

3.1. CINDER の設定

cinder-volume ノードには、Ceph ブロックデバイスドライバー、**ボリューム** プール、ユーザー、およびシークレットの UUID が Ceph ブロックデバイスと対話する必要があります。Cinder を設定するには、以下の手順を実行します。

1. Cinder 設定ファイルを開きます。

```
# vim /etc/cinder/cinder.conf
```

2. **[DEFAULT]** セクションで、Cinder のバックエンドとして Ceph を有効にします。

```
enabled_backends = ceph
```

3. Glance API のバージョンが 2 に設定されていることを確認します。複数の cinder バックエンドを **enabled_backends** に設定する場合には、**glance_api_version = 2** 設定は **[DEFAULT]** セクションではなく、**[ceph]** セクションになければなりません。

```
glance_api_version = 2
```

4. **cinder.conf** ファイルに **[ceph]** セクションを作成します。以下のステップで **[ceph]** セクションに Ceph 設定を追加します。

5. **volume_driver** の設定を指定し、Ceph ブロックデバイスドライバーを使用するように設定します。以下に例を示します。

```
volume_driver = cinder.volume.drivers.rbd.RBDDriver
```

6. クラスタ名と Ceph の設定ファイルの場所を指定します。典型的なデプロイメントでは、**ceph** クラスタのクラスタ名および Ceph 設定ファイルは **/etc/ceph/ceph.conf** にあります。Ceph クラスタ名が **ceph** でない場合には、クラスタ名と設定ファイルパスを適切に指定します。以下に例を示します。

```
rbd_cluster_name = us-west  
rbd_ceph_conf = /etc/ceph/us-west.conf
```

7. デフォルトでは、OSP は Ceph ボリュームを **rbd** プールに保存します。以前に作成した **volumes** プールを使用するには、**rbd_pool** 設定を指定し、**volumes** プールを設定します。以下に例を示します。

```
rbd_pool = volumes
```

8. OSP には、ボリュームのデフォルトのユーザー名やシークレットの UUID がありません。**rbd_user** を指定して、**cinder** ユーザーに設定します。次に **rbd_secret_uuid** 設定を指定し、**uuid-secret.txt** ファイルに保存されている生成された UUID に設定します。以下に例を示します。

```
rbd_user = cinder  
rbd_secret_uuid = 4b5fd580-360c-4f8c-abb5-c83bb9a3f964
```

9. 以下の設定を指定します。

```

rdm_flatten_volume_from_snapshot = false
rdm_max_clone_depth = 5
rdm_store_chunk_size = 4
rados_connect_timeout = -1

```

その結果、設定は以下のようになります。

```

[DEFAULT]
enabled_backends = ceph
glance_api_version = 2
...

[ceph]
volume_driver = cinder.volume.drivers.rbd.RBDDriver
rbd_cluster_name = ceph
rbd_pool = volumes
rbd_user = cinder
rbd_ceph_conf = /etc/ceph/ceph.conf
rbd_flatten_volume_from_snapshot = false
rbd_secret_uuid = 4b5fd580-360c-4f8c-abb5-c83bb9a3f964
rbd_max_clone_depth = 5
rbd_store_chunk_size = 4
rados_connect_timeout = -1

```



注記

デフォルトの **[lvm]** セクションと、その設定を削除することを検討してください。

3.2. CINDER バックアップの設定

cinder-backup ノードには特定のデーモンが必要です。Cinder バックアップを設定するには、以下の手順を実行します。

1. Cinder 設定ファイルを開きます。

```
# vim /etc/cinder/cinder.conf
```

2. 設定ファイルの **[ceph]** セクションに移動します。
3. **backup_driver** の設定を指定し、Ceph ドライバーに設定します。

```
backup_driver = cinder.backup.drivers.ceph
```

4. **backup_ceph_conf** の設定を指定し、Ceph 設定ファイルへのパスを指定します。

```
backup_ceph_conf = /etc/ceph/ceph.conf
```

**注記**

Cinder バックアップ Ceph 設定ファイルは、Cinder に使用する Ceph 設定ファイルとは異なる場合があります。たとえば、別の Ceph クラスタを指す場合があります。

5. バックアップ用の Ceph プールを指定します。

```
backup_ceph_pool = backups
```

**注記**

Cinder で使用する Cinder バックアップに同じプールを使用することもできますが、推奨されません。異なる CRUSH 階層でプールを使用することを検討してください。

6. **backup_ceph_user** 設定を指定し、ユーザーを **cinder-backup** として指定します。

```
backup_ceph_user = cinder-backup
```

7. 以下の設定を指定します。

```
backup_ceph_chunk_size = 134217728
backup_ceph_stripe_unit = 0
backup_ceph_stripe_count = 0
restore_discard_excess_bytes = true
```

Cinder 設定が含まれているので、**cinder.conf** ファイルの **[ceph]** セクションは以下のようになります。

```
[ceph]
volume_driver = cinder.volume.drivers.rbd.RBDDriver
rbd_cluster_name = ceph
rbd_pool = volumes
rbd_user = cinder
rbd_ceph_conf = /etc/ceph/ceph.conf
rbd_flatten_volume_from_snapshot = false
rbd_secret_uuid = 4b5fd580-360c-4f8c-abb5-c83bb9a3f964
rbd_max_clone_depth = 5
rbd_store_chunk_size = 4
rados_connect_timeout = -1

backup_driver = cinder.backup.drivers.ceph
backup_ceph_user = cinder-backup
backup_ceph_conf = /etc/ceph/ceph.conf
backup_ceph_chunk_size = 134217728
backup_ceph_pool = backups
backup_ceph_stripe_unit = 0
backup_ceph_stripe_count = 0
restore_discard_excess_bytes = true
```

`/etc/openstack-dashboard/` で Cinder バックアップが有効になっているかどうかを確認します。この設定は `local_settings` または `local_settings.py` という名前のファイルにある必要があります。以下に例を示します。

```
cat /etc/openstack-dashboard/local_settings | grep enable_backup
```

`enable_backup` を **False** に設定すると、これを **True** に設定します。以下に例を示します。

```
OPENSTACK_CINDER_FEATURES = {
    'enable_backup': True,
}
```

3.3. GLANCE の設定

デフォルトで Ceph ブロックデバイスを使用するには、`/etc/glance/glance-api.conf` ファイルを編集します。必要に応じて以下の設定のコメントを解除し、値を変更します。別のプールを使用した場合は、ユーザーまたは Ceph の設定ファイルの適切な値が適用されます。

```
# vim /etc/glance/glance-api.conf

stores = rbd
default_store = rbd
rbd_store_chunk_size = 8
rbd_store_pool = images
rbd_store_user = glance
rbd_store_ceph_conf = /etc/ceph/ceph.conf
```

CoW (copy-on-write) クローン作成を有効にするには、`show_image_direct_url` を **True** に設定します。

```
show_image_direct_url = True
```



重要

CoW を有効にすると、Glance の API 経由でバックエンドの場所を公開するため、エンドポイントは一般にアクセスできないはずです。

必要に応じてキャッシュ管理を無効にします。`flavor` は、`keystone+cachemanagement` ではなく、`keystone` にのみ設定する必要があります。

```
flavor = keystone
```

Red Hat では、イメージに以下のプロパティを推奨します。

```
hw_scsi_model=virtio-scsi
hw_disk_bus=scsi
hw_qemu_guest_agent=yes
os_require_quiesce=yes
```

`virtio-scsi` コントローラーのパフォーマンスが向上し、破棄操作に対応します。SCSI/SAS ドライブを使用するシステムの場合は、`cinder` ブロックデバイスをすべてそのコントローラーに接続します。ま

た、QEMU ゲストエージェントを有効にし、QEMU ゲストエージェントを介して **fs-freeze/thaw** 呼び出しを送信します。

3.4. NOVA の設定

すべての **nova-compute** ノードで、Ceph 設定ファイルを編集して Nova の一時バックエンドを設定し、すべての仮想マシンを Ceph に直接ブートします。

1. Ceph 設定ファイルを開きます。

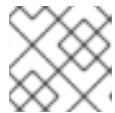
```
# vim /etc/ceph/ceph.conf
```

2. 以下のセクションを Ceph 設定ファイルの **[client]** セクションに追加します。

```
[client]
rd cache = true
rd cache writethrough until flush = true
rd concurrent management ops = 20
admin socket = /var/run/ceph/guests/$cluster-$type.$id.$pid.$cctid.asok
log file = /var/log/ceph/qemu-guest-$pid.log
```

3. 管理ソケットおよびログファイルのディレクトリを作成し、**qemu** ユーザーおよび **libvirt** グループを使用するようにパーミッションを変更します。

```
mkdir -p /var/run/ceph/guests/ /var/log/ceph/
chown qemu:libvirt /var/run/ceph/guests /var/log/ceph/
```



注記

ディレクトリは SELinux または AppArmor が許可している必要があります。

すべての **nova-compute** ノードで、**[libvirt]** セクションの下にある **/etc/nova/nova.conf** ファイルを編集し、以下の設定を行います。

```
[libvirt]
images_type = rbd
images_rbd_pool = vms
images_rbd_ceph_conf = /etc/ceph/ceph.conf
rbd_user = cinder
rbd_secret_uuid = 4b5fd580-360c-4f8c-abb5-c83bb9a3f964
disk_cachemodes="network=writeback"
inject_password = false
inject_key = false
inject_partition = -2
live_migration_flag="VIR_MIGRATE_UNDEFINE_SOURCE,VIR_MIGRATE_PEER2PEER,VIR_MIGRATE_LIVE,VIR_MIGRATE_PERSIST_DEST,VIR_MIGRATE_TUNNELLED"
hw_disk_discard = unmap
```

Ceph 設定ファイルが **/etc/ceph/ceph.conf** ではない場合は、正しいパスを指定します。**rbd_user_secret** の UUID は、**uuid-secret.txt** ファイルの UUID に置き換えます。

3.5. OPENSTACK サービスの再起動

Ceph ブロックデバイスドライバーをアクティブにするには、ブロックデバイスプール名と Ceph ユーザー名を設定に読み込み、対応する設定ファイルを変更した後に適切な OpenStack サービスを再起動します。

```
# systemctl restart openstack-cinder-volume  
# systemctl restart openstack-cinder-backup  
# systemctl restart openstack-glance-api  
# systemctl restart openstack-nova-compute
```