



# Red Hat OpenStack Platform 16.2

## 大規模デプロイメントにおける推奨事項

大規模な OpenStack Platform をデプロイする際のハードウェア要件および設定



# Red Hat OpenStack Platform 16.2 大規模デプロイメントにおける推奨事項

---

大規模な OpenStack Platform をデプロイする際のハードウェア要件および設定

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/Recommendations\_for\_Large\_Deployments.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<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.

## 概要

本ガイドでは、大規模な Red Hat OpenStack Platform をデプロイする際のさまざまな推奨事項を記載します。これらの推奨事項には、ハードウェアの推奨事項、アンダークラウドのチューニング、およびオーバークラウドの設定が含まれます。

---

## 目次

前書き .....	3
多様性を受け入れるオープンソースの強化 .....	4
RED HAT ドキュメントへのフィードバック (英語のみ) .....	5
第1章 大規模デプロイメントにおける推奨事項 .....	6
第2章 大規模な RED HAT OPENSTACK デプロイメントで推奨される仕様 .....	7
2.1. アンダークラウドのシステム要件 .....	7
2.2. オーバークラウドコントローラーノードのシステム要件 .....	7
2.3. オーバークラウドコンピュートノードのシステム要件 .....	10
2.4. RED HAT CEPH STORAGE ノードのシステム要件 .....	11
第3章 RED HAT OPENSTACK デプロイメントのベストプラクティス .....	13
3.1. RED HAT OPENSTACK デプロイメントの準備 .....	13
3.2. RED HAT OPENSTACK のデプロイメント設定 .....	15
3.3. アンダークラウドのチューニング .....	16
3.4. オーバークラウドのチューニング .....	17
第4章 デバッグの推奨問題および既知の問題 .....	19
4.1. 既知の問題 .....	19
4.2. イントロスペクションのデバッグ .....	19
4.3. デプロイメントのデバッグ .....	19



# 前書き

## 多様性を受け入れるオープンソースの強化

Red Hat では、コード、ドキュメント、Web プロパティにおける配慮に欠ける用語の置き換えに取り組んでいます。まずは、マスター (master)、スレーブ (slave)、ブラックリスト (blacklist)、ホワイトリスト (whitelist) の 4 つの用語の置き換えから始めます。この取り組みは膨大な作業を要するため、今後の複数のリリースで段階的に用語の置き換えを実施して参ります。詳細は、[弊社](#) の CTO、Chris Wright の [メッセージ](#) を参照してください。



## RED HAT ドキュメントへのフィードバック (英語のみ)

弊社ドキュメントに対するご意見をお聞かせください。ドキュメントの改善点があればお知らせください。

### ドキュメントへのダイレクトフィードバック (DDF) 機能の使用 (英語版のみ)

特定の文章、段落、またはコードブロックに対して直接コメントを送付するには、DDF の **Add Feedback** 機能を使用してください。なお、この機能は英語版のドキュメントでのみご利用いただけます。

1. **Multi-page HTML** 形式でドキュメントを表示します。
2. ドキュメントの右上隅に **Feedback** ボタンが表示されていることを確認してください。
3. コメントするテキスト部分をハイライト表示します。
4. **Add Feedback** をクリックします。
5. **Add Feedback** フィールドにコメントを入力します。
6. (オプション) ドキュメントチームが連絡を取り問題についてお伺いできるように、ご自分のメールアドレスを追加します。
7. **Submit** をクリックします。

## 第1章 大規模デプロイメントにおける推奨事項

大規模な Red Hat OpenStack Platform (RHOSP) 環境をデプロイする際には、以下のアンダークラウドおよびオーバークラウドの推奨事項、仕様、および設定を使用します。100 を超えるオーバークラウドノードが含まれる RHOSP 16.2 デプロイメントは、大規模な環境とみなされます。Red Hat は、ミニオンを使用しない RHOSP 16.2 を使用して、700 のオーバークラウドノードが含まれる大規模な環境で、最大限のパフォーマンスをテストおよび検証しています。

DCN ベースのデプロイメントの場合には、中央サイトおよびエッジサイトからのノードの数が非常に大きい場合があります。DCN デプロイメントに関する推奨事項は、Red Hat Global Support Services にお問い合わせください。

## 第2章 大規模な RED HAT OPENSTACK デプロイメントで推奨される仕様

提供される推奨事項を使用して、大規模なクラスターデプロイメントをスケーリングできます。

以下の手順の値は、Red Hat OpenStack Platform Performance & Scale Team が実施したテストに基づいており、個々の環境によって異なる可能性があります。詳細は、「[Scaling Red Hat OpenStack Platform 16.1 to more than 700 nodes](#)」を参照してください。

### 2.1. アンダークラウドのシステム要件

最適なパフォーマンスを得るには、物理サーバーにアンダークラウドノードをインストールします。ただし、仮想アンダークラウドノードを使用する場合、仮想マシンには、以下の表で説明されている物理マシンと同様の十分なリソースを確保するようにしてください。

表2.1 推奨されるアンダークラウドノードの仕様

システム要件	説明
ノード数	1
CPU の数	32 コア、64 スレッド
ディスク	500 GB のルートディスク (1x SSD または 2x 7200 RPM のハードドライブ (RAID 1))  Object Storage (swift) 用 500 GB のディスク (1x SSD または 2x 7200 RPM のハードドライブ (RAID 1))
メモリー容量	256 GB
ネットワーク	25 Gbps のネットワークインターフェースまたは 10 Gbps のネットワークインターフェース

### 2.2. オーバークラウドコントローラーノードのシステム要件

すべてのコントロールプレーンサービスは、3つのノードで稼働する必要があります。通常、すべてのコントロールプレーンサービスは3つのコントローラーノードに分散してデプロイされます。

#### コントローラーサービスのスケーリング

コントローラーサービスで利用可能なリソースを増やすには、これらのサービスを追加のノードにスケーリングできます。たとえば、**db** または **messaging** コントローラーサービスを専用のノードにデプロイして、コントローラーノードの負荷を軽減することができます。

コントローラーサービスをスケーリングするには、コンポーザブルロールを使用してスケーリングするサービスのセットを定義します。コンポーザブルロールを使用する場合には、各サービスは3つの追加の専用ノードで実行される必要があります、Pacemaker のクォーラムを維持するために、コントロールプレーン内のノードの合計数を追加する必要があります。

この例のコントロールプレーンは、以下の9ノードで構成されます。

- コントローラーノード 3 台
- データベースノード 3 台
- メッセージングノード 3 台

詳しい情報は、『[Advanced Overcloud Customization](#)』の「[Composable services and custom roles](#)」を参照してください。

コンポーザブルロールを使用したコントローラーサービスのスケーリングに関する質問は、Red Hat Global Professional Services にお問い合わせください。

### ストレージに関する考慮事項

オーバークラウドデプロイメントのコントローラーノードを計画する場合には、十分なストレージを追加します。OpenStack Telemetry Metrics (gnocchi) および OpenStack Image (glance) は、I/O 負荷の高いサービスです。オーバークラウドは I/O の負荷を Ceph OSD サーバーに移すので、Image サービスおよびテレメトリー用に Ceph Storage を使用します。

デプロイメントに Ceph ストレージが含まれていない場合には、Telemetry Metrics (gnocchi) および Image (glance) サービスが利用できる Object Storage (swift) 用に専用のディスクまたはノードを使用します。コントローラーノードで Object Storage を使用する場合は、ルートディスクとは別に NVMe デバイスを使用し、オブジェクトデータ保存時のディスク使用率を削減します。

### CPU に関する考慮事項

コントローラーノードが受け取る API 呼び出し、AMQP メッセージ、およびデータベースクエリーの数が、コントローラーノードでの CPU メモリー消費に影響を与えます。各 Red Hat OpenStack Platform (RHOSP) コンポーネントがタスクを同時に処理および実行する能力は、個々の RHOSP コンポーネントに設定されるワーカースレッドの数によっても制限されます。RHOSP director がコントローラー上で設定するコンポーネントのワーカースレッドの数は、CPU 数によって制限されます。

デプロイメントで Ceph Storage ノードを使用する場合、ノード数が 700 を超える大規模な環境では以下の仕様が推奨されます。

表2.2 Ceph Storage ノードを使用する場合に推奨されるコントローラーノードの仕様

システム要件	説明
ノード数	Controller ロールに含まれるコントローラーサービスを持つ 3 台のコントローラーノード  オプションとして、専用ノードでコントローラーサービスをスケーリングするには、コンポーザブルサービスを使用します。詳しい情報は、『 <a href="#">Advanced Overcloud Customization</a> 』の「 <a href="#">Composable services and custom roles</a> 」を参照してください。
CPU の数	2 ソケット (それぞれ 32 コア、64 スレッド)
ディスク	500 GB のルートディスク (1x SSD または 2x 7200 RPM のハードドライブ (RAID 1))
メモリー容量	384 GB

システム要件	説明
ネットワーク	<p>25 Gbps のネットワークインターフェースまたは 10 Gbps のネットワークインターフェース。10 Gbps ネットワークインターフェースを使用する場合には、ネットワークボンディングを使用して 2 つのボンディングを作成します。</p> <ul style="list-style-type: none"> <li>● プロビジョニング (bond0、mode4)、内部 API (bond0、mode4)、プロジェクト (bond0、mode4)</li> <li>● ストレージ (bond1、mode4)、ストレージ管理 (bond1、mode4)</li> </ul>

デプロイメントで Ceph Storage ノードを使用しない場合、ノード数が 700 を超える大規模な環境では以下の仕様が推奨されます。

表2.3 Ceph Storage ノードを使用しない場合に推奨されるコントローラーノードの仕様

システム要件	説明
ノード数	<p>Controller ロールに含まれるコントローラーサービスを持つ 3 台のコントローラーノード</p> <p>オプションとして、専用ノードでコントローラーサービスをスケールアップするには、コンポーザブルサービスを使用します。詳しい情報は、『<a href="#">Advanced Overcloud Customization</a>』の「<a href="#">Composable services and custom roles</a>」を参照してください。</p>
CPU の数	2 ソケット (それぞれ 32 コア、64 スレッド)
ディスク	<p>500 GB のルートディスク (1xSSD または 2x 7200 RPM のハードドライブ (RAID 1))</p> <p>Swift 用 500 GB のディスク (1xSSD または 2x 7200 RPM のハードドライブ (RAID 1))</p>
メモリー容量	384 GB

システム要件	説明
ネットワーク	<p>25 Gbps のネットワークインターフェースまたは 10 Gbps のネットワークインターフェース。10 Gbps ネットワークインターフェースを使用する場合には、ネットワークボンディングを使用して 2 つのボンディングを作成します。</p> <ul style="list-style-type: none"> <li>● プロビジョニング (bond0、mode4)、内部 API (bond0、mode4)、プロジェクト (bond0、mode4)</li> <li>● ストレージ (bond1、mode4)、ストレージ管理 (bond1、mode4)</li> </ul>

### 2.3. オーバークラウドコンピュートノードのシステム要件

オーバークラウドのデプロイメントを計画する際には、コンピュートノードに推奨されるシステム要件を確認します。

表2.4 コンピュートノードに推奨される仕様

システム要件	説明
ノード数	Red Hat は、さまざまなコンポーザブル Compute ロールで 700 ノードのスケールについてテストを実施しています。
CPU の数	2 ソケット (それぞれ 12 コア、24 スレッド)
ディスク	<p>500 GB のルートディスク (1x SSD または 2x 7200 RPM のハードドライブ (RAID 1))</p> <p>Image サービス (glance) のイメージキャッシュ用 500 GB のディスク (1x SSD または 2x 7200 RPM のハードドライブ (RAID 1))</p>
メモリー容量	<p>128 GB (NUMA ノードあたり 64 GB)。デフォルトでは 2 GB がホスト用に確保されます。</p> <p>分散仮想ルーターでは、確保されるメモリーを 5 GB に増やします。</p>

システム要件	説明
ネットワーク	<p>25 Gbps のネットワークインターフェースまたは 10 Gbps のネットワークインターフェース。10 Gbps ネットワークインターフェースを使用する場合には、ネットワークボンディングを使用して 2 つのボンディングを作成します。</p> <ul style="list-style-type: none"> <li>● プロビジョニング (bond0、mode4)、内部 API (bond0、mode4)、プロジェクト (bond0、mode4)</li> <li>● ストレージ (bond1、mode4)</li> </ul>

## 2.4. RED HAT CEPH STORAGE ノードのシステム要件

オーバークラウドのデプロイメントを計画する際には、Ceph ストレージノードに推奨される以下のシステム要件を確認します。

表2.5 Ceph Storage ノードに推奨される仕様

システム要件	説明
ノード数	<p>3 方向レプリケーションのノードが少なくとも 5 台必要です。オールフラッシュ設定を使用する場合は、双方向レプリケーションのノードが少なくとも 3 台必要です。</p>
CPU の数	<p>1 OSD あたり 1 つの Intel Broadwell CPU コア (ストレージ I/O の要件に対応するため)。I/O 負荷が軽い場合は、Ceph をブロックデバイスの速度で実行する必要がない場合があります。たとえば、一部の NFV アプリケーションの場合、Ceph はデータの持続性、高可用性、および低レイテンシーを提供しますが、スループットはターゲットではないため、CPU 処理能力を下げるのが許容されます。</p>
メモリー容量	<p>1 OSD あたり 5 GB のメモリーを確保するようにしてください。これは、OSD プロセスメモリー用だけでなく、パフォーマンスを最適化するために OSD データおよびメタデータをキャッシュするために必要です。ハイパーコンバージドインフラストラクチャー (HCI) 環境では、コンピュータノードの仕様と共に必要なメモリーを計算します。</p>

システム要件	説明
ネットワーク	ネットワーク容量 (MB/s 単位) を Ceph デバイスの合計 MB/s 容量よりも大きくし、大規模な I/O 転送サイズを使用するワークロードをサポートするようにしてください。OSD 間のトラフィックを別の物理ネットワークポートセットに移動して、クラスターネットワークを使用して書き込みレイテンシーを軽減します。Red Hat OpenStack Platform でこれを行うには、ネットワーク用に別の VLAN を設定し、別の物理ネットワークインターフェースに VLAN を割り当てます。
ディスク	bluestore <b>block.db</b> パーティションにソリッドステートドライブ (SSD) ディスクを使用して、ハードディスクドライブ (HDD) の I/O 競合を減らします。これにより、書き込み IOPS の速度が向上しますが、SSD は 1 秒あたりの読み取りの入出力操作には影響を及ぼしません。SATA/SAS SSD ジャーナルを使用する場合には、通常 SSD:HDD の比率を 1:5 にする必要があります。NVM SSD ジャーナルを使用する場合は、通常 SSD:HDD の比率を 1:10 (負荷のほとんどが読み取りの場合には 1:15) にすることができます。ただし、この比率が高すぎると、SSD ジャーナルデバイスの障害により OSD に影響が及ぶ可能性があります。

Ceph ノードのハードウェア要件に関する詳細は、Red Hat Ceph Storage 4 『[ハードウェアガイド](#)』の「[ハードウェアを選択する一般的な原則](#)」を参照してください。

Ceph ノードのデプロイメント設定についての詳しい情報は、『[Deploying an overcloud with containerized Red Hat Ceph](#)』を参照してください。

ストレージのレプリケーション数の変更に関する詳細は、『[Red Hat Ceph Storage 設定ガイド](#)』の「[プール、配置グループ、および CRUSH 設定オプション](#)」を参照してください。



## 第3章 RED HAT OPENSTACK デプロイメントのベストプラクティス

OpenStack のデプロイを計画して準備する場合には、以下のベストプラクティスを確認してください。お使いの環境で、これらのプラクティスの1つまたは複数を実用することができます。

### 3.1. RED HAT OPENSTACK デプロイメントの準備

Red Hat OpenStack Platform (RHOSP) をデプロイする前に、以下のデプロイメント準備タスクの一覧を確認してください。お使いの環境で、デプロイメント準備タスクの1つまたは複数を実用することができます。

イントロスペクションのサブネット範囲を設定して、一度にイントロスペクションを実施する最大オーバークラウドノードに対応する。

director を使用して RHOSP をデプロイおよび設定する場合には、コントロールプレーンネットワークに CIDR 表記を使用して、現在または今後追加するすべてのオーバークラウドノードに対応します。

オーバークラウドイメージの root パスワードを設定して、オーバークラウドイメージへのコンソールアクセスを許可する。

ネットワークが正しく設定されていない場合に、コンソールを使用して失敗したデプロイメントのトラブルシューティングを行います。詳細は、『[Partner Integration Guide](#)』の「[Installing virt-customize to the director](#)」および「[Setting the Root Password](#)」を参照してください。この推奨事項を実施する場合、パスワード管理に関する組織の情報セキュリティポリシーに従います。

あるいは、`userdata_root_password.yaml` テンプレートを使用して、`NodeUserData` パラメーターを使用して root パスワードを設定することができます。テンプレートは `/usr/share/openstack-tripleo-heat-templates/firstboot/userdata_root_password.yaml` にあります。

以下の例では、テンプレートを使用して `NodeUserData` パラメーターを設定します。

```
resource_registry:
  OS::TripleO::NodeUserData: /usr/share/openstack-tripleo-heat-
    templates/firstboot/userdata_root_password.yaml
parameter_defaults:
  NodeRootPassword: '<password>'
```

スケジューラーヒントを使用して、ハードウェアをロールに割り当てる。

- スケジューラーヒントを使用して、**Controller**、**Compute**、**CephStorage** などのロールにハードウェアを割り当てます。スケジューラーヒントにより、特定のハードウェアのみに影響するデプロイメントの問題をより簡単に特定できます。
- 単一プロセスである **nova-scheduler** は、多数のノードをスケジューリングする際に酷使される可能性があります。スケジューラーヒントがタグの照合を実装する際に、スケジューラーヒントは **nova-scheduler** への負荷を軽減します。その結果、**nova-scheduler** はデプロイメント時のスケジューリングエラーが少なくなり、スケジューラーヒントを使用する場合、デプロイメントの所要時間が短縮されました。
- スケジューラーヒントを使用する場合は、プロファイルのタグ付けを使用しないでください。
- パフォーマンステストでは、特定のロールに同じハードウェアを使用して、テストおよびパフォーマンスの結果の差異を軽減します。

- 詳細は、『Advanced Overcloud Customization』の「Assigning Specific Node IDs」を参照してください。

各ノードのルートディスクヒントとして World Wide Name (WWN) を設定し、デプロイメントおよび起動時にノードが誤ったディスクを使用しないようにする。

ノードに複数のディスクが含まれる場合は、イントロスペクションデータを使用して、各ノードのルートディスクヒントとして WWN を設定します。これにより、デプロイメントおよび起動時にノードが誤ったディスクを使用しないようになります。詳しい情報は、『Director Installation and Usage』の「Defining the Root Disk for Nodes」を参照してください。

複数のディスクを持つノードで Bare Metal サービス (ironic) の自動クリーニングを有効にする。

Bare Metal サービスの自動クリーニングを使用して、複数のディスクを持ち複数のブートローダーがある可能性があるノードでメタデータを消去します。ディスクに複数のブートローダーが存在するため、ノードはブートディスクとの一貫性を失い、その結果、誤った URL を使用するメタデータをプルしようとするノードのデプロイメントに失敗します。

Bare Metal サービスの自動クリーニングを有効にするには、アンダークラウドノードで **undercloud.conf** ファイルを編集し、以下の行を追加します。

```
clean_nodes = true
```

Bare Metal (ironic) イントロスペクションのノード数を制限する。

すべてのノードで同時にイントロスペクションを実行する場合には、ネットワークの制約により失敗する可能性があります。一度に最大 50 のノードでイントロスペクションを実施します。

**undercloud.conf** ファイルの **dhcp\_start** および **dhcp\_end** の範囲が、環境にあるノードの数に対して十分な大きさになるようにしてください。

利用可能な IP が十分でない場合は、範囲のサイズを超えては発行しないでください。これにより、同時に実行するイントロスペクション操作の数が制限されます。イントロスペクションの DHCP リースが期限切れになるのを許可するには、イントロスペクションが完了してから数分間は IP アドレスをさらに発行しないでください。

異なるタイプの設定向けに Ceph を準備する。

以下の一覧は、異なるタイプの設定の推奨事項のセットです。

- **オールフラッシュ OSD 設定**

それぞれの OSD には、デバイス種別の IOPS 能力に応じて追加の CPU が必要になるため、Ceph IOPS は少数の OSD で CPU に制限されます。これは、従来の HDD よりも 2 桁大きい IOPS 能力を持つことのできる NVM SSD の場合に言えます。SATA/SAS SSD の場合、HDD よりも 1 桁大きいランダム IOPS/OSD が予想されますが、シーケンシャル IOPS は 2-4 倍しか増えません。Ceph が OSD デバイス用に必要とする CPU リソースよりも少ないリソースしか Ceph に提供できません。

- **ハイパーコンバージドインフラストラクチャー (HCI)**

OpenStack Compute (nova) ゲスト用に、CPU パワー、メモリー、およびネットワークの半分以上を確保することが推奨されます。OpenStack Compute (nova) ゲストと Ceph Storage の両方をサポートするのに十分な CPU パワーとメモリーがあることを確認します。Ceph Storage のメモリー消費は弾力的ではないため、メモリー消費を確認します。マルチ CPU ソケットシステムでは、NUMA ピニングされた Ceph で Ceph の CPU 消費を 1 つのソケットに制限します。たとえば、**numactl -N 0 -p 0** コマンドを使用します。Ceph のメモリー消費を 1 つのソケットにハードピニングしないでください。

- **NFV 等のレイテンシーに敏感なアプリケーション**

異なる NUMA ソケットおよびネットワークカード上で動作するネットワークアプリケーションにより、可能であれば Ceph を Ceph が使用するネットワークカードと同じ CPU ソケットに配置し、ネットワークカードの割り込みをその CPU ソケットに制限します。

デュアルブートローダーを使用する場合は、OSD のマッピングに `disk-by-path` を使用します。これにより、デバイス名を使用するのとは異なり、ユーザーは一貫性のあるデプロイメントを行うことができます。以下のスニペットは、`disk-by-path` マッピングの `CephAnsibleDisksConfig` の例です。

```
CephAnsibleDisksConfig:
  osd_scenario: non-collocated
  devices:
    - /dev/disk/by-path/pci-0000:03:00.0-scsi-0:2:0:0
    - /dev/disk/by-path/pci-0000:03:00.0-scsi-0:2:1:0
  dedicated_devices:
    - /dev/nvme0n1
    - /dev/nvme0n1
  journal_size: 512
```

## 3.2. RED HAT OPENSTACK のデプロイメント設定

Red Hat OpenStack Platform (RHOSP) デプロイメント設定の推奨事項に関する以下の一覧を確認してください。

### 小規模なデプロイメントにより heat テンプレートを検証する

3つ以上のコントローラーノード、1つのコンピュータノード、および3つの Ceph Storage ノードで構成される、小規模な環境をデプロイします。この設定を使用して、すべての heat テンプレートが正しいことを確認することができます。

### アンダークラウドでテレメトリ通知を無効にする

以下の OpenStack サービスについて、アンダークラウドでテレメトリの通知を無効にして、RabbitMQ キューを削減することができます。

- Compute (nova)
- Networking (neutron)
- Orchestration (heat)
- Identity (keystone)

通知を無効にするには、`/usr/share/openstack-tripleo-heat-templates/environments/disable-telemetry.yaml` テンプレートで、通知ドライバーの設定を `noop` に設定します。

### 同時にプロビジョニングされるノードの数を制限する

平均的なエンタープライズレベルのラックユニット内に収まる典型的なサーバーの数は 50 台であるため、1つのラック分の平均的なノードをデプロイできます。

デプロイメントでの問題を診断するために必要なデバッグを最小限にするには、一度に 50 を超えるノードをデプロイしないでください。ただし、より大きな数のノードをデプロイする場合には、Red Hat では最大 100 ノードの同時テストに成功しています。

コンピュータノードをバッチでスケールアップするには、`--limit` オプションを指定して `openstack overcloud deploy` コマンドを使用します。これにより、時間が節約され、アンダークラウドでのリソース消費が削減されます。

### 未使用の NIC を無効にする

デプロイ中にオーバークラウドに未使用の NIC がある場合には、NIC 設定テンプレートで未使用のインターフェースを定義して、インターフェースを **use\_dhcp: false** および **defroute: false** に設定する必要があります。

未使用のインターフェースを定義しない場合、イントロスペクションおよびスケーリング操作中に、ルーティングの問題や IP 割り当ての問題が発生する可能性があります。デフォルトでは、NIC は **BOOTPROTO=dhcp** を設定します。つまり、未使用のオーバークラウド NIC は、PXE プロビジョニングに必要な IP アドレスを消費します。これにより、ノードで利用可能な IP アドレスのプールが減少する場合があります。

### 未使用の Bare Metal Provisioning (ironic) ノードの電源をオフにする

メンテナンスモードにある未使用の Bare Metal Provisioning (ironic) ノードの電源をオフにしてください。Red Hat は、以前のデプロイメントからのノードが電源オンの状態でメンテナンスモードのままになるケースを特定しています。これは、Bare Metal の自動クリーニングで発生する可能性があります。この場合、クリーニングに失敗したノードがメンテナンスモードに設定されます。Bare Metal Provisioning は、メンテナンスモードのノードの電源状態を追跡せず、電源状態を誤ってオフとして報告します。これにより、進行中のデプロイメントで問題が発生する可能性があります。デプロイメントの失敗後に再デプロイする場合には、ノードの電源管理デバイスを使用するすべての未使用のノードの電源をオフにしてください。

## 3.3. アンダークラウドのチューニング

RHOSP のデプロイメントをスケーリングする予定で、デフォルトのアンダークラウド設定にチューニングを適用する場合には、本セクションを確認してください。

### Telemetry サービス (ceilometer) を使用する場合は、サービスのパフォーマンスが向上します

Telemetry は CPU 負荷の高いサービスであるため、RHOSP 16.2 では Telemetry はデフォルトで有効にされません。Telemetry を使用する場合は、サービスのパフォーマンスを向上させることができます。

詳しくは、『[Deployment Recommendations for Specific Red Hat OpenStack Platform Services](#)』の「[Telemetry](#)」を参照してください。

### プロビジョニングプロセスと設定プロセスを分離します

- スタックおよび関連する RHOSP リソースのみを作成するには、**--stack-only** オプションを指定してデプロイメントコマンドを実行できます。オーバークラウドに必要なすべての環境ファイルを追加します。

```
$ openstack overcloud deploy \
--templates \
-e <environment-file1.yaml> \
-e <environment-file2.yaml> \
...
--stack-only
```

- スタックのプロビジョニングが完了したら、**tripleo-admin** ユーザーのアンダークラウドからオーバークラウドへの SSH アクセスを有効にすることができます。**config-download** プロセスでは、**tripleo-admin** ユーザーを使用して Ansible ベースの設定を実施します。

```
$ openstack overcloud admin authorize
```

- オーバークラウドスタックの作成を無効にして、ソフトウェア設定を適用する **config-download** ワークフローだけを実行するには、**--config-download-only option** を指定して

デプロイメントコマンドを実行することができます。オーバークラウドに必要なすべての環境ファイルを追加します。

```
$ openstack overcloud deploy \
--templates \
-e <environment-file1.yaml> \
-e <environment-file2.yaml> \
...
--config-download-only
```

- **config-download** Playbook の実行を特定のノードまたはノードセットに制限するには、**--limit** オプションを使用します。スケールアップ操作の場合、新しいノードにのみソフトウェア設定を適用するには、**--config-download-only** オプションと共に **--limit** オプションを使用します。

```
$ openstack overcloud deploy \
--templates \
-e <environment-file1.yaml> \
-e <environment-file2.yaml> \
...
--config-download-only --config-download-timeout --limit <Undercloud>,<Controller>,
<Compute-1>,<Compute-2>
```

### 3.4. オーバークラウドのチューニング

Red Hat OpenStack Platform (RHOSP) のデプロイメントをスケーリングする予定で、デフォルトのオーバークラウド設定にチューニングを適用する場合には、以下のセクションを確認してください。

#### ファイルオーバーを防ぐために OVN OVSDDB クライアント確認間隔を増やします

大規模な RHOSP デプロイメントの OVSDDB クライアント確認間隔を増やします。Pacemaker は、設定されたタイムアウト内に OVN からの応答を取得しない場合に、**ovn-dbs-bundle** のフェイルオーバーをトリガーします。OVN OVSDDB クライアント確認間隔を 360 秒に増やすには、heat テンプレートの **OVNDBSPacemakerTimeout** パラメーターを編集します。

```
OVNDBSPacemakerTimeout: 360
```

それぞれのコンピューターノードおよびコントローラーノードで、OVN コントローラーは OVN SBDB を定期的に確認し、これらの要求がタイムアウトした場合には、OVN コントローラーが再同期します。リソースの作成要求が複数のコンピューターノードおよびコントローラーノードに読み込まれる場合、デフォルトの 60 秒のタイムアウト値は十分ではありません。OVN SBDB クライアント確認間隔を 180 秒に増やすには、heat テンプレートの **OVNOpenflowProbeInterval** パラメーターを編集します。

```
ControllerParameters:
OVNRemoteProbeInterval: 180000
OVNOpenflowProbeInterval: 180
```



## 注記

RHOSP ユーザーおよびサービスによってトリガーされた操作時に、リソースの制約 (CPU または メモリーリソースの制約) により、複数のコンポーネントが設定したタイムアウト値に到達する場合があります。これにより、haproxy フロントエンドまたはバックエンドへのタイムアウトリクエストの失敗、メッセージングタイムアウト、db クエリー関連の失敗、クラスターの不安定性などが生じる可能性があります。初回のデプロイメント後にオーバークラウド環境のベンチマークテストを実施し、タイムアウト関連のボトルネックを特定するのに役立てます。

## 第4章 デバッグの推奨問題および既知の問題

デプロイメントのトラブルシューティングに役立つ可能性のあるデバッグの提案については、以下のセクションを参照してください。

### 4.1. 既知の問題

既存の現在の制限の概要を以下に示します。

#### BZ#1857451: Ansible フォークの値には上限があり、現在の計算を変更する必要がある

デフォルトでは、mistral の Ansible Playbook は、`ansible.cfg` ファイルの `10*CPU_COUNT` フォークを使用するように設定されています。Ansible の実行を特定ノードまたはノードセットに制限する `--limit` オプションを使用せず、Ansible の実行をすべての既存ノードで実行するよう設定されている場合、Ansible はメモリーをほぼ 100% 消費します。

### 4.2. イントロスペクションのデバッグ

イントロスペクションをデバッグする際には、以下の一覧に示す推奨事項を確認してください。

#### undercloud.conf ファイルでイントロスペクション用 DHCP 範囲と NIC を確認する

これらの値のいずれかが誤りである場合は修正し、`openstack undercloud install` コマンドを再度実行します。

#### DHCP 範囲のノードが許可できるよりも多くのノードでイントロスペクションを試行しないようにする

各ノードの DHCP リースは、イントロスペクションの終了後約 2 分間アクティブな状態を維持します。

#### ターゲットノードが応答している状態にする

全ノードでイントロスペクションに異常が発生した場合には、設定済みの NIC を使用してネイティブ VLAN 経由でターゲットノードを ping できること、および帯域外インターフェースの認証情報およびアドレスが正しいことを確認します。

#### コンソールでイントロスペクションコマンドを確認する

特定のノードをデバッグする際には、ノードのブート時にコンソールを監視し、ノードのイントロスペクションコマンドを確認します。PXE プロセスの完了前にノードが停止した場合は、接続、IP の割り当て、およびネットワーク負荷を確認します。ノードが BIOS を終了し、イントロスペクションイメージでブートする場合は、障害はまれで、ほぼ接続性の問題に関連します。イントロスペクションイメージからのハートビートが、アンダークラウドへの伝送中に中断されないようにします。

### 4.3. デプロイメントのデバッグ

デプロイメントをデバッグする際には、以下の推奨事項を使用してください。

#### プロビジョニングネットワーク上のアドレスを提供する DHCP サーバーを検証する

プロビジョニングネットワーク上のアドレスを提供する追加の DHCP サーバーにより、Red Hat OpenStack Platform director はマシンを検査およびプロビジョニングできなくなります。

- DHCP または PXE のイントロスペクションの問題については、以下のコマンドを入力します。

```
$ sudo tcpdump -i any port 67 or port 68 or port 69
```

- DHCP または PXE のデプロイメントの問題については、以下のコマンドを入力します。

```
$ sudo ip netns exec qdhcp tcpdump -i <interface> port 67 or port 68 or port 69
```

### 障害の発生したディスクまたは外部ディスクの状態を確認する

障害の発生したディスクまたは外部ディスクの場合、ディスクの状態をチェックし、マシンの帯域外管理に応じて、失敗したディスクまたは外部ディスクの状態が **Up** に設定されていることを確認します。ディスクは、デプロイメントサイクル中に **Up** の状態を終了し、ディスクがベースオペレーティングシステムに表示される順番を変更する可能性があります。

以下のコマンドを使用して、失敗したオーバークラウドのデプロイメントをデバッグする

- **openstack stack failures list overcloud**
- **heat resource-list -n5 overcloud | grep -i fail**
- **less /var/lib/mistral/config-download-latest/ansible.log**

コマンドの出力を確認するには、障害が発生するノードにログインして、**/var/log/** および **/var/log/containers/** のログファイルを確認します。