



# Red Hat Virtualization 4.0

## テクニカルリファレンス

Red Hat Virtualization 環境のテクニカルアーキテクチャー



# Red Hat Virtualization 4.0 テクニカルリファレンス

---

Red Hat Virtualization 環境のテクニカルアーキテクチャー

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/Technical\_Reference.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

このリファレンスでは、Red Hat Virtualization 環境で使用される概念、コンポーネント、およびテクノロジーについて説明します。

## 目次

<b>第1章 はじめに</b> .....	<b>4</b>
1.1. RED HAT VIRTUALIZATION MANAGER	4
1.2. RED HAT VIRTUALIZATION HOST	4
1.3. MANAGER にアクセスするためのインターフェイス	6
1.4. MANAGER をサポートするコンポーネント	7
1.5. ストレージ	8
1.6. ネットワーク	9
1.7. データセンター	12
<b>第2章 ストレージ</b> .....	<b>13</b>
2.1. ストレージドメインの概要	13
2.2. ストレージバックイングストレージドメインの種類	13
2.3. ストレージドメインタイプ	14
2.4. 仮想ディスクイメージのストレージ形式	14
2.5. 仮想ディスクイメージストレージ割り当てポリシー	15
2.6. RED HAT VIRTUALIZATION のストレージメタデータバージョン	15
2.7. RED HAT VIRTUALIZATION でのストレージドメインの自動リカバリー	16
2.8. STORAGE POOL MANAGER	16
2.9. STORAGE POOL MANAGER の選択プロセス	18
2.10. RED HAT VIRTUALIZATION における排他的リソースと SANLOCK	18
2.11. シンプロビジョニングとストレージのオーバーコミットメント	19
2.12. 論理ボリューム拡張	20
<b>第3章 ネットワーク</b> .....	<b>22</b>
3.1. ネットワークアーキテクチャー	22
3.2. 概要：基本的なネットワーク用語	22
3.3. ネットワークインターフェイスコントローラー	22
3.4. ブリッジ	22
3.5. ボンド	23
ボンディングモード	23
3.6. ボンディング用のスイッチ設定	24
3.7. 仮想ネットワークインターフェイスカード	24
3.8. 仮想 LAN (VLAN)	26
3.9. ネットワークラベル	27
3.10. クラスタネットワーク	28
3.11. 論理ネットワーク	29
3.12. 必須ネットワーク、任意のネットワーク、および仮想マシンネットワーク	32
3.13. 仮想マシンの接続性	32
3.14. PORT MIRRORING	32
3.15. ホストネットワーク設定	33
3.16. BRIDGE CONFIGURATION	34
3.17. VLAN 設定	34
3.18. ブリッジおよびボンディングの設定	35
3.19. 複数のブリッジ、複数 VLAN、および NIC の設定	36
3.20. 複数のブリッジ、複数の VLAN、およびボンディングの設定	36
<b>第4章 電源管理</b> .....	<b>38</b>
4.1. 電力管理とフェンシングの概要	38
4.2. RED HAT VIRTUALIZATION の PROXY による電源管理	38
4.3. 電源管理	38
4.4. フェンシング	40
4.5. ソフトフェンシングホスト	41

4.6. 複数の電力管理フェンシングエージェントの使用	42
<b>第5章 負荷分散、スケジューリング、および移行</b>	<b>44</b>
5.1. 負荷分散、スケジューリング、および移行	44
5.2. 負荷分散ポリシー	44
5.3. 負荷分散ポリシー: VM_EVENLY_DISTRIBUTED	45
5.4. 負荷分散ポリシー: EVENLY_DISTRIBUTED	45
5.5. 負荷分散ポリシー: POWER_SAVING	46
5.6. 負荷分散ポリシー: なし	46
5.7. 負荷分散ポリシー: INCLUSTERUPGRADE	47
5.8. 高可用性仮想マシンの予約	47
5.9. スケジューリング	47
5.10. マイグレーション	48
<b>第6章 ディレクトリーサービス</b>	<b>49</b>
6.1. ディレクトリーサービス	49
6.2. ローカル認証: 内部ドメイン	49
6.3. GSSAPI を使用したリモート認証	50
<b>第7章 テンプレートおよびプール</b>	<b>52</b>
7.1. テンプレートおよびプール	52
7.2. テンプレート	52
7.3. POOLS	53
<b>第8章 仮想マシンのスナップショット</b>	<b>54</b>
8.1. スナップショット	54
8.2. RED HAT VIRTUALIZATION のライブスナップショット	55
8.3. スナップショットの作成	55
8.4. スナップショットプレビュー	57
8.5. スナップショットの削除	58
<b>第9章 ハードウェアドライバーとデバイス</b>	<b>60</b>
9.1. 仮想化されたハードウェア	60
9.2. RED HAT VIRTUALIZATION における安定したデバイスアドレス	60
9.3. 中央処理装置 (CPU)	61
9.4. システムデバイス	61
9.5. ネットワークデバイス	62
9.6. グラフィックデバイス	62
9.7. ストレージデバイス	63
9.8. サウンドデバイス	63
9.9. シリアルドライバー	63
9.10. バルーンドライバー	63
<b>第10章 最小要件および技術制限</b>	<b>65</b>
10.1. 最小要件およびサポートされる制限	65
10.2. リソースの制限	65
10.3. クラスターの制限	65
10.4. ストレージドメインの制限	66
10.5. RED HAT VIRTUALIZATION MANAGER の制限	67
10.6. ハイパーバイザーの要件	68
10.7. ゲストの要件とサポートの制限	71
10.8. SPICE の制限	72
10.9. 追加の参考資料	72

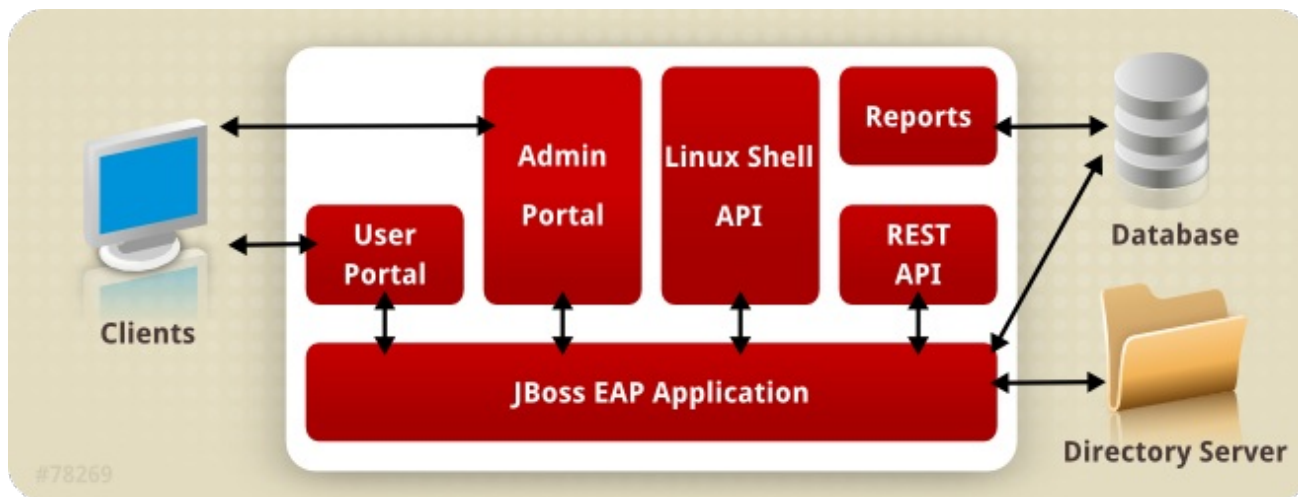


## 第1章 はじめに

### 1.1. RED HAT VIRTUALIZATION MANAGER

Red Hat Virtualization Manager は、仮想化環境に一元管理を提供します。Red Hat Virtualization Manager にアクセスする際に、さまざまなインターフェイスを使用できます。各インターフェイスは、異なる方法で仮想化環境へのアクセスを容易にします。

図1.1 Red Hat Virtualization Manager のアーキテクチャー



Red Hat Virtualization Manager は、グラフィカルインターフェイスとアプリケーション プログラミングインターフェイス(API)を提供します。各インターフェイスは、Red Hat JBoss Enterprise Application Platform の組み込みインスタンスによって提供されるアプリケーションである Manager に接続します。Red Hat JBoss Enterprise Application Platform に加えて、Red Hat Virtualization をサポートする他の多くのコンポーネントがあります。

### 1.2. RED HAT VIRTUALIZATION HOST

Red Hat Virtualization 環境には、1つ以上のホストが接続されています。ホストは、仮想マシンが利用する物理ハードウェアを提供するサーバーです。

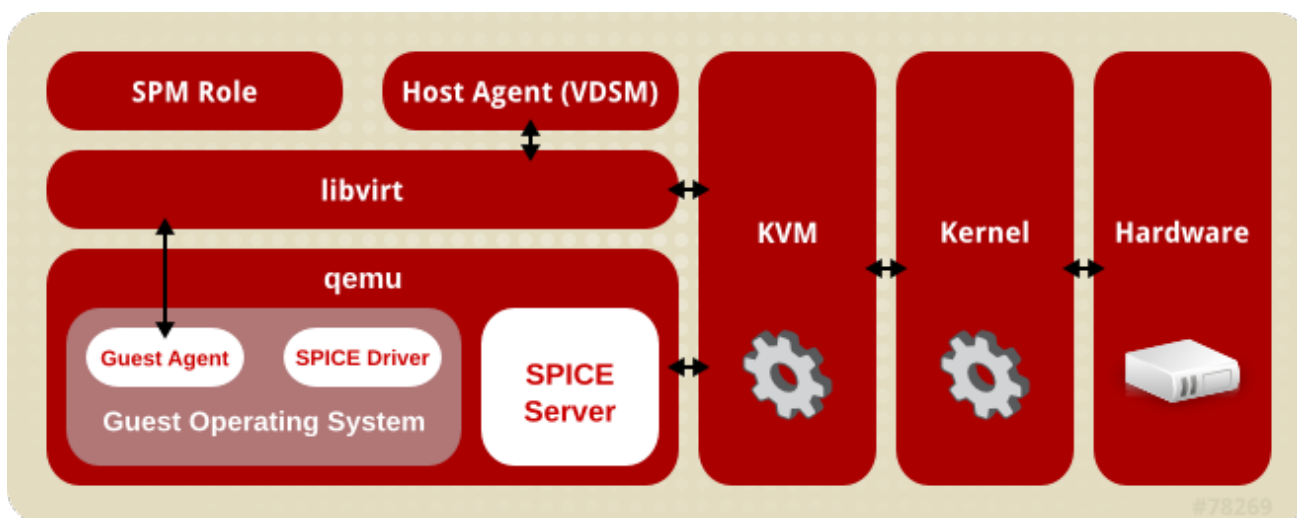
Red Hat Virtualization ホスト (RHVH) は、仮想化ホストを作成するために特別にカスタマイズされたインストールメディアを使用してインストールされた最適化されているオペレーティングシステムを実行します。

Red Hat Enterprise Linux ホストは、インストール後にホストとして使用できるように設定された標準の Red Hat Enterprise Linux オペレーティングシステムを実行しているサーバーです。

ホストのいずれの方法でも、同じ方法で残りの仮想化環境と対話するホストが作成されます。つまり、どちらも **ホスト** と呼ばれます。



図1.2 ホストアーキテクチャー



### カーネルベースの仮想マシン(KVM)

カーネルベースの仮想マシン (KVM) は、Intel VT または AMD-V ハードウェア拡張機能を使用して完全仮想化を提供するロード可能なカーネルモジュールです。KVM 自体はカーネルスペースで実行されますが、KVM で実行されるゲストはユーザースペースで個別の QEMU プロセスとして実行されます。KVM を使用すると、ホストはその物理ハードウェアを仮想マシンで使用できるようになります。

### QEMU

QEMU は、完全なシステムエミュレーションを提供するために使用されるマルチプラットフォームエミュレーターです。QEMU は、1つ以上のプロセッサと周辺機器を含む PC などの完全なシステムをエミュレートします。QEMU を使用して、さまざまなオペレーティングシステムを起動したり、システムコードをデバッグしたりできます。QEMU は、KVM および適切な仮想化拡張機能を備えたプロセッサと連携して動作し、完全なハードウェア支援による仮想化を提供します。

### Red Hat Virtualization Manager ホストエージェント (VDSM)

Red Hat Virtualization では、**VDSM** は仮想マシンおよびストレージでアクションを開始します。また、ホスト間の通信も容易になります。VDSM は、メモリー、ストレージ、ネットワークなどのホストリソースを監視します。さらに、VDSM は、仮想マシンの作成、統計の蓄積、ログの収集などのタスクを管理します。VDSM インスタンスは各ホストで実行し、設定可能なポート **54321** を使用して Red Hat Virtualization Manager から管理コマンドを受け取ります。

### VDSM-REG

**VDSM** は **VDSM-REG** を使用して各ホストを Red Hat Virtualization Manager に登録します。**VDSM-REG** は、ポート **80** またはポート **443** を使用して、ホストとそのホストに関する情報を提供します。

### libvirt

Libvirt は、仮想マシンとそれに関連する仮想デバイスの管理を容易にします。Red Hat Virtualization Manager が仮想マシンのライフサイクルコマンド (開始、停止、再起動) を開始すると、VDSM は関連するホストマシンで libvirt を呼び出してそれらを実行します。

### Storage Pool Manager (SPM)

Storage Pool Manager (SPM) は、データセンター内の1つのホストに割り当てられたロールです。SPM ホストは、データセンターのすべてのストレージドメイン構造メタデータを変更する唯一の権限を持っています。これには、仮想ディスクイメージ、スナップショット、およびテンプレートの作成、削除、および操作が含まれます。また、SAN (Storage Area Network) 上のスパースブロック

デバイス用のストレージの割り当ても含まれます。SPM のロールは、データセンター内の任意のホストに移行できます。その結果、データセンター内のすべてのホストは、データセンターで定義されているすべてのストレージドメインにアクセスする必要があります。

Red Hat Virtualization Manager は、SPM が常に利用できることを確認します。ストレージ接続エラーが発生すると、Manager は SPM のロールを別のホストに再割り当てします。

## ゲストのオペレーティングシステム

ゲストオペレーティングシステムは、Red Hat Virtualization 環境の仮想マシンに変更せずにインストールできます。ゲストオペレーティングシステム、およびゲスト上のアプリケーションは、仮想化環境を認識せず、正常に実行されます。

Red Hat は、仮想化デバイスへのより高速で効率的なアクセスを可能にする拡張デバイスドライバーを提供します。Red Hat Virtualization Guest Agent をゲストにインストールすることもできます。これにより、管理コンソールに拡張ゲスト情報が提供されます。

## 1.3. MANAGER にアクセスするためのインターフェイス

### ユーザーポータル

デスクトップ仮想化は、個人コンピューターのデスクトップ環境に類似したデスクトップ環境をユーザーに提供します。ユーザーポータルは、*Virtual Desktop Infrastructure* をユーザーに配信します。ユーザーは、Web ブラウザーからユーザーポータルにアクセスして、割り当てられた仮想デスクトップを表示し、アクセスします。ユーザーポータルでユーザーが使用できるアクションは、システム管理者が設定します。標準ユーザーは、システム管理者によって割り当てられたデスクトップを起動、停止、および使用できます。パワーユーザーは、いくつかの管理アクションを実行できます。どちらのタイプのユーザーは同じ URL からユーザーポータルにアクセスし、ログイン時にパーミッションレベルに適したオプションと共に表示されます。

- **標準のユーザーアクセス**

標準ユーザーは、仮想デスクトップの電源をオンまたはオフにし、ユーザーポータルを介して接続できます。仮想マシンへの直接接続は、*Simple Protocol for Independent Computing Environments (SPICE)* または *Virtual Network Computing (VNC)* クライアントにより容易になります。どちらのプロトコルも、ローカルにインストールされたデスクトップ環境と同様の環境をユーザーに提供します。管理者は、仮想マシンの作成時に仮想マシンへの接続に使用するプロトコルを指定します。

ユーザーポータルおよびサポートされるブラウザーおよびクライアントの詳細は、[ユーザーポータルの概要](#) を参照してください。

- **Power User Access**

Red Hat Virtualization User Portal は、仮想リソースの作成、使用、監視を行うグラフィカルユーザーインターフェイスをパワーユーザーに提供します。システム管理者は、ユーザーにパワーユーザーアクセスを付与することで、一部の管理タスクを委譲できます。標準ユーザーが実行できるタスクに加えて、パワーユーザーは次のことができます。

- 仮想マシンの作成、編集、および削除
- 仮想ディスクおよびネットワークインターフェイスを管理します。
- 仮想マシンにユーザーパーミッションを割り当てます。
- テンプレートを作成して使用し、仮想マシンを迅速にデプロイします。

- リソースの使用状況と重大度の高いイベントを監視します。
- スナップショットを作成して使用し、仮想マシンを以前の状態に復元します。

パワーユーザーは、委譲する仮想マシン管理タスクを実行できます。データセンターおよびクラスターレベルの管理タスクは、環境管理者用に保存されます。

## Administration Portal

管理ポータルは、Red Hat Virtualization Manager サーバーのグラフィカル管理インターフェイスです。管理者は、Web ブラウザーからを使用して、仮想化環境のすべての要素を監視、作成、および維持できます。管理ポータルから実行できるタスクには、以下が含まれます。

- 仮想インフラストラクチャー（ネットワーク、ストレージドメイン）の作成および管理。
- ホストのインストールおよび管理。
- 論理エンティティ（データセンター、クラスター）の作成および管理。
- 仮想マシンの作成および管理。
- Red Hat Virtualization のユーザーおよびパーミッション管理。

管理ポータルは JavaScript を使用して表示されます。

管理ポータルの機能については、[Red Hat Virtualization Administration Guide](#) で詳しく説明しています。管理ポータルがサポートするブラウザーおよびプラットフォームに関する情報は、[Red Hat Virtualization Installation Guide](#) を参照してください。

## Representational State Transfer (REST) API

Red Hat Virtualization REST API は、Red Hat Virtualization 環境の交換および制御のためのソフトウェアインターフェイスを提供します。REST API は、HTTP アクションをサポートする任意のプログラミング言語で使用することができます。

REST API 開発者および管理者は以下を行うことができます。

- エンタープライズ IT システムとの統合
- サードパーティーの仮想化ソフトウェアとの統合
- 自動メンテナンスおよびエラーチェックタスクを実行します。
- スクリプトを使用して、Red Hat Virtualization 環境で反復タスクを自動化します。

API 仕様と使用例については、[Red Hat Virtualization REST API Guide](#) を参照してください。

## 1.4. MANAGER をサポートするコンポーネント

### Red Hat JBoss Enterprise Application Platform

Red Hat JBoss Enterprise Application Platform は Java アプリケーションサーバーです。クロスプラットフォームの Java アプリケーションの効率的な開発と配信をサポートするフレームワークを提供します。Red Hat Virtualization Manager は、Red Hat JBoss Enterprise Application Platform を使用して提供されます。



## 重要

Red Hat Virtualization Manager にバンドルされている Red Hat JBoss Enterprise Application Platform のバージョンは、他のアプリケーションを提供するために使用されません。これは、Red Hat Virtualization Manager にサービスを提供するという特定の目的のためにカスタマイズされています。Manager に含まれている Red Hat JBoss Enterprise Application Platform を追加の目的で使用すると、Red Hat Virtualization 環境にサービスを提供する機能に悪影響を及ぼします。

## レポートおよび履歴データの収集

Red Hat Virtualization Manager には、ホスト、仮想マシン、およびストレージに関するモニタリングデータを収集する Data Warehouse が含まれています。多数の事前定義されたレポートが利用可能です。お客様は、SQL をサポートするクエリツールを使用して、環境を分析し、レポートを作成できます。

Red Hat Virtualization Manager のインストールプロセスにより、2つのデータベースが作成されます。これらのデータベースは、インストール時に選択された Postgres インスタンスで作成されます。

- engine データベースは、Red Hat Virtualization Manager が使用するメインのデータストアです。仮想化環境に関する情報 (状態、設定、およびパフォーマンスなど) が、このデータベースに保管されます。
- ovirt\_engine\_history データベースには、engine 運用データベースから時間の経過とともに照合される設定情報および統計メトリクスが含まれます。engine データベースの設定データは1分ごとに検査され、変更は ovirt\_engine\_history データベースに複製されます。データベースへの変更を追跡することで、データベース内のオブジェクトに関する情報が提供されます。これにより、Red Hat Virtualization 環境のパフォーマンスを分析して向上させ、問題を解決することができます。

ovirt\_engine\_history データベースに基づいてレポートを生成する方法は、『Red Hat Virtualization Data Warehouse Guide』の [History Database](#) を参照してください。



## 重要

ovirt\_engine\_history データベースのデータレプリケーションは、**RHEVM History Service** (ovirt-engine-dwhd)によって実行されます。

## ディレクトリーサービス

ディレクトリーサービスは、ユーザーおよび組織の情報をネットワークベースのストレージで一元的に保存します。保存される情報の種類には、アプリケーション設定、ユーザープロファイル、グループデータ、ポリシー、およびアクセス制御が含まれます。Red Hat Virtualization Manager は、Active Directory、Identity Management (IdM)、OpenLDAP、および Red Hat Directory Server 9 をサポートします。管理目的のみのローカル内部ドメインもあります。この内部ドメインには、admin ユーザーという1人のユーザーしかいません。

## 1.5. ストレージ

Red Hat Virtualization は、仮想ディスクイメージ、テンプレート、スナップショット、および ISO ファイルに集中ストレージシステムを使用します。ストレージは、ストレージドメインで設定されるストレージプールに論理的にグループ化されます。ストレージドメインは、ストレージ容量と、ストレージ

の内部構造を説明するメタデータの組み合わせです。ストレージドメインには、data、export、および ISO の3つのタイプがあります。

データストレージドメインは、各データセンターに必要な唯一のストレージドメインです。データストレージドメインは、単一のデータセンター専用です。エクスポートドメインと ISO ドメインは任意です。ストレージドメインは共有リソースであり、データセンター内のすべてのホストがアクセスできる必要があります。

ストレージネットワークは、ネットワークファイルシステム(NFS)、Internet Small Computer System Interface (iSCSI)、GlusterFS、Fibre Channel Protocol (FCP)、または POSIX 準拠のネットワークファイルシステムを使用して実装できます。

NFS（およびその他の POSIX 準拠のファイルシステム）ドメインでは、仮想ディスク、テンプレート、およびスナップショットはすべて単純なファイルです。

SAN (iSCSI/FCP) ドメインでは、ブロックデバイスは論理ボリュームマネージャー(LVM)によりボリュームグループ(VG)に集約されます。各仮想ディスク、テンプレート、およびスナップショットは、VG 上の論理ボリューム（論理ボリューム）です。LVMの詳細は、[Red Hat Enterprise Linux Logical Volume Manager Administration Guide](#) を参照してください。

### データストレージドメイン

データドメインは、環境で実行しているすべての仮想マシンの仮想ハードディスクイメージを保持します。仮想マシンのテンプレートやスナップショットもデータドメインに保存されます。データセンター間でデータドメインを共有することはできません。

### ストレージドメインのエクスポート

エクスポートドメインは、データセンターと Red Hat Virtualization 環境間でイメージをコピーして移動するために使用される一時ストレージリポジトリです。エクスポートドメインを使用して、仮想マシンとテンプレートをバックアップできます。エクスポートドメインはデータセンター間で移動できますが、同時に1つのデータセンターでしか有効にできません。

### ISO ストレージドメイン

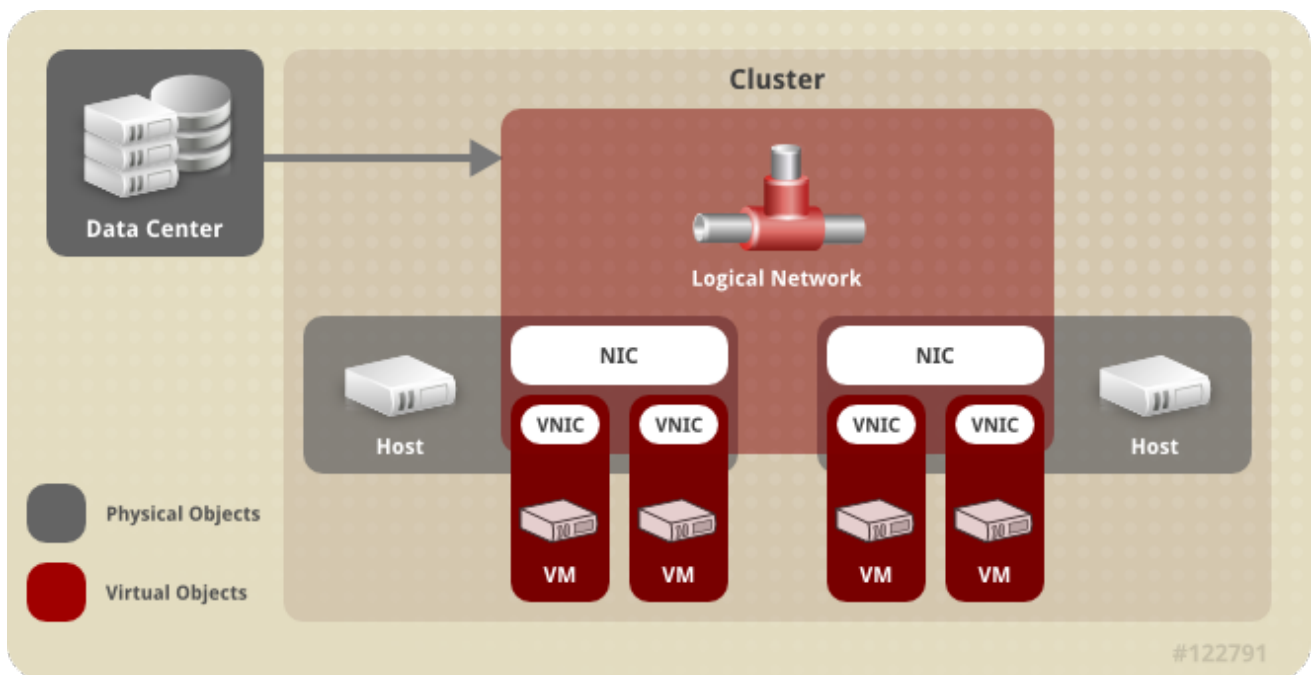
ISO ドメインは、仮想マシンのオペレーティングシステムおよびアプリケーションをインストールするために使用される論理 CD-ROM である ISO ファイルを保存します。物理 CD-ROM または DVD のライブラリーを置き換える論理エンティティとして、ISO ドメインはデータセンターの物理メディアの必要性を削除します。ISO ドメインは、異なるデータセンターで共有することができます。

## 1.6. ネットワーク

Red Hat Virtualization ネットワークアーキテクチャーは、Red Hat Virtualization 環境のさまざまな要素間の接続を容易にします。ネットワークアーキテクチャーは、ネットワーク接続をサポートするだけでなく、ネットワークの分離も可能にします。



図1.3 ネットワークアーキテクチャー



ネットワーキングは、Red Hat Virtualization でいくつかのレイヤーで定義されています。基盤となる物理ネットワークインフラストラクチャーは、ハードウェアと Red Hat Virtualization 環境の論理コンポーネント間の接続を可能にするように配置および設定されている必要があります。

### ネットワークインフラストラクチャー層

Red Hat Virtualization ネットワークアーキテクチャーは、いくつかの一般的なハードウェアおよびソフトウェアデバイスに依存しています。

- ネットワークインターフェイスコントローラー(NIC)は、ホストをネットワークに接続する物理ネットワークインターフェイスデバイスです。
- 仮想NIC (vNIC)は、ホストの物理NICを使用して動作する論理NICです。これらは、仮想マシンへのネットワーク接続を提供します。
- ボンドは、複数のNICを単一のインターフェイスにバインドします。
- ブリッジは、パケット交換ネットワークのパケット転送技術です。これらは、仮想マシンの論理ネットワークのベースを形成します。

### 論理ネットワーク

論理ネットワークでは、環境要件に基づいてネットワークトラフィックを分離できます。論理ネットワークの種類は次のとおりです。

- 仮想マシンネットワークトラフィックを伝送する論理ネットワーク、
- 仮想マシンネットワークトラフィックを伝送しない論理ネットワーク、
- オプションの論理ネットワーク
- 必要なネットワーク。

すべての論理ネットワークは、必須または任意のいずれかです。

仮想マシンネットワークトラフィックを伝送する論理ネットワークは、ソフトウェアブリッジデバイスとしてホストレベルで実装されます。デフォルトでは、Red Hat Virtualization Manager のインストール中に1つの論理ネットワーク(**ovirtmgmt** 管理ネットワーク)が定義されます。

管理者が追加できるその他の論理ネットワークは、専用のストレージ論理ネットワークと専用のディスプレイ論理ネットワークです。仮想マシントラフィックを伝送しない論理ネットワークには、ホスト上に関連付けられたブリッジデバイスがありません。これらは、ホストネットワークインターフェイスに直接関連付けられています。

Red Hat Virtualization は、管理関連のネットワークトラフィックを移行関連のネットワークトラフィックから分離します。これにより、ライブマイグレーションに専用ネットワーク (ルーティングなし) を使用できるようになり、移行中に管理ネットワーク (ovirtmgmt) がハイパーバイザーへの接続を失うことがなくなります。

## 異なる層の論理ネットワークの説明

論理ネットワークは、仮想化環境の各レイヤーに異なる影響を及ぼします。

### データセンター層

論理ネットワークはデータセンターレベルで定義されます。各データセンターには、デフォルトで **ovirtmgmt** 管理ネットワークがあります。それ以上の論理ネットワークはオプションですが、推奨されます。**仮想マシンネットワーク** およびカスタム MTU の指定はデータセンターレベルで設定できます。データセンター用に定義された論理ネットワークも、論理ネットワークを使用するクラスターに追加する必要があります。

### クラスター層

論理ネットワークはデータセンターから利用可能になり、それらを使用するクラスターに追加する必要があります。デフォルトでは、各クラスターは管理ネットワークに接続されています。オプションで、クラスターの親データセンター用に定義されたクラスター論理ネットワークに追加できます。必要な論理ネットワークがクラスターに追加されたら、クラスター内のホストごとに実装する必要があります。任意の論理ネットワークは、必要に応じてホストに追加できます。

### ホスト層

仮想マシンの論理ネットワークは、特定のネットワークインターフェイスに関連付けられたソフトウェアブリッジデバイスとして、クラスター内のホストごとに実装されます。非仮想マシンの論理ネットワークにはブリッジが関連付けられておらず、ホストネットワークインターフェイスに直接関連付けられています。各ホストには、Red Hat Virtualization 環境に含まれている結果として、そのネットワークデバイスの1つを使用するブリッジとして実装された管理ネットワークがあります。クラスターに追加されたさらに必要な論理ネットワークは、クラスターで動作可能になるために、各ホストのネットワークインターフェイスに関連付けられている必要があります。

### 仮想マシン層

論理ネットワークは、ネットワークを物理マシンで利用できるようにするのと同じ方法で、仮想マシンで利用できるようにすることができます。仮想マシンは、その仮想 NIC を、それを実行するホストに実装されている任意の仮想マシン論理ネットワークに接続できます。次に、仮想マシンは、接続されている論理ネットワーク上で使用可能な他のデバイスまたは宛先への接続を取得します。

#### 例1.1 Management Network

**ovirtmgmt** という名前の管理論理ネットワークは、Red Hat Virtualization Manager のインストール時に自動的に作成されます。**ovirtmgmt** ネットワークは、Red Hat Virtualization Manager とホスト間のトラフィックの管理専用です。特に目的のブリッジが設定されていない場合、**ovirtmgmt** はすべてのトラフィックのデフォルトのブリッジになります。

## 1.7. データセンター

データセンターは、Red Hat Virtualization における最高レベルの抽象化です。データセンターは、3 種類のサブコンテナで設定されるコンテナです。

- **ストレージコンテナ**は、ストレージドメインの接続情報など、ストレージタイプおよびストレージドメインに関する情報を保持します。ストレージはデータセンター用に定義されており、データセンター内のすべてのクラスターで使用できます。データセンター内のすべてのホストクラスターは、同じストレージドメインにアクセスできます。
- **network コンテナ**は、データセンターの論理ネットワークに関する情報を保持します。これには、ネットワークアドレス、VLAN タグ、STP サポートなどの詳細が含まれます。論理ネットワークはデータセンターに対して定義され、オプションでクラスターレベルで実装されません。
- **クラスター コンテナ**はクラスターを保持します。クラスターは、AMD または Intel プロセッサのいずれかの互換性のあるプロセッサコアを備えたホストのグループです。クラスターは移行ドメインです。仮想マシンは、他のクラスターではなく、クラスター内の任意のホストにライブマイグレーションできます。1つのデータセンターに複数のクラスターを保持でき、各クラスターに複数のホストを含めることができます。



## 第2章 ストレージ

### 2.1. ストレージドメインの概要

ストレージドメインは、共通のストレージインターフェイスを持つイメージのコレクションです。ストレージドメインには、テンプレートと仮想マシンの完全なイメージ (スナップショットを含む)、ISO ファイル、およびそれら自体に関するメタデータが含まれています。ストレージドメインは、ブロックデバイス (SAN - iSCSI または FCP) またはファイルシステム (NAS - NFS、GlusterFS、またはその他の POSIX 準拠のファイルシステム) のいずれかで設定できます。

NAS では、すべての仮想ディスク、テンプレート、およびスナップショットはファイルです。

SAN (iSCSI/FCP) では、各仮想ディスク、テンプレート、またはスナップショットは論理ボリュームです。ブロックデバイスは、ボリュームグループと呼ばれる論理エンティティに集約され、LVM (Logical Volume Manager) によって論理ボリュームに分割されて仮想ハードディスクとして使用されます。LVM の詳細は、[Red Hat Enterprise Linux Logical Volume Manager Administration Guide](#) を参照してください。

仮想ディスクには、QCOW2 または RAW のいずれかの形式を使用できます。ストレージのタイプは、スパースまたは事前割り当てのいずれかです。スナップショットは常にスパースですが、RAW または sparse として作成されたディスクに対して取得できます。

同じストレージドメインを共有する仮想マシンは、同じクラスターに属するホスト間で移行することができます。

### 2.2. ストレージバックングストレージドメインの種類

ストレージドメインは、ブロックベースおよびファイルベースのストレージを使用して実装できます。

#### ファイルベースのストレージ

Red Hat Virtualization でサポートされているファイルベースのストレージタイプは、NFS、GlusterFS、その他の POSIX 準拠のファイルシステム、およびホストにローカルなストレージです。

ファイルベースのストレージは、Red Hat Virtualization 環境の外部で管理されます。

NFS ストレージは、Red Hat Enterprise Linux NFS サーバーまたはその他のサードパーティーのネットワーク接続ストレージサーバーによって管理されます。

ホストは、独自のローカルストレージファイルシステムを管理できます。

#### ブロックベースのストレージ

ブロックストレージは、フォーマットされていないブロックデバイスを使用します。ブロックデバイスは、論理ボリュームマネージャー(LVM)によってボリュームグループに集約されます。LVM のインスタンスは、他のホストで実行されているインスタンスを認識せずに、すべてのホストで実行されます。VDSM は、ボリュームグループの変更をスキャンすることにより、LVM の上にクラスターリングロジックを追加します。変更が検出されると、VDSM は、ボリュームグループ情報を更新するように指示することにより、個々のホストを更新します。ホストはボリュームグループを論理ボリュームに分割し、論理ボリュームのメタデータをディスクに書き込みます。既存のストレージドメインにさらにストレージ容量が追加されると、Red Hat Virtualization Manager により、各ホストの VDSM がボリュームグループ情報を更新します。

論理ユニット番号 (LUN) は、個々のブロックデバイスです。サポートされているブロックストレージプロトコルの1つである iSCSI、FCoE、または SAS は、LUN への接続に使用されます。Red Hat

Virtualization Manager は、LUN へのソフトウェア iSCSI 接続を管理します。他のすべてのブロックストレージ接続は、Red Hat Virtualization 環境の外部で管理されます。論理ボリュームの作成、論理ボリュームの拡張または削除、新しい LUN の追加など、ブロックベースのストレージ環境での変更は、Storage Pool Manager と呼ばれる特別に選択されたホスト上の LVM によって処理されます。次に、変更は VDSM によって同期され、ストレージメタデータはクラスター内のすべてのホスト間で更新されます。

## 2.3. ストレージドメインタイプ

Red Hat Virtualization は、3 種類のストレージドメインをサポートします。これには、各ストレージドメインがサポートするストレージタイプが含まれます。

- *Data Storage Domain* は、Red Hat Virtualization 環境内のすべての仮想マシンのハードディスクイメージを保存します。ディスクイメージには、インストールされているオペレーティングシステム、または仮想マシンによって保存または生成されたデータが含まれている場合があります。データストレージドメインは、NFS、iSCSI、FCP、GlusterFS、および POSIX 準拠のストレージをサポートします。データドメインを複数のデータセンター間で共有することはできません。
- *Export Storage Domain* は、ハードディスクイメージと、データセンター間で転送される仮想マシンテンプレートの一時的なストレージを提供します。さらに、エクスポートストレージドメインは、バックアップされた仮想マシンのコピーを保存します。エクスポートストレージドメインは NFS ストレージをサポートします。複数のデータセンターが単一のエクスポートストレージドメインにアクセスできますが、一度に使用できるのは1つのデータセンターのみです。
- *ISO ストレージドメイン* は、イメージとも呼ばれる ISO ファイルを保存します。ISO ファイルは、物理 CD または DVD を表したものです。Red Hat Virtualization 環境では、一般的なタイプの ISO ファイルは、オペレーティングシステムのインストールディスク、アプリケーションのインストールディスク、およびゲストエージェントのインストールディスクです。これらのイメージは、物理ディスクをディスクドライブに挿入して起動するのと同じ方法で、仮想マシンに接続して起動できます。ISO ストレージドメインにより、データセンター内のすべてのホストが ISO を共有できるようになり、物理的な光メディアが不要になります。

## 2.4. 仮想ディスクイメージのストレージ形式

### QCOW2 形式の仮想マシンストレージ

QCOW2 は、仮想ディスクイメージのストレージ形式です。QCOW は、書き込み時の QEMU コピーを表します。QCOW2 形式は、論理ブロックと物理ブロックの間にマッピングを追加することにより、物理ストレージ層を仮想層から切り離します。各論理ブロックは物理オフセットにマッピングされ、ストレージのオーバーコミットメントと仮想マシンのスナップショットを有効にします。各 QCOW ボリュームは、基礎となるディスクイメージに加えられた変更のみを表します。

初期マッピングは、すべての論理ブロックをバックアップファイルまたはボリュームのオフセットにポイントします。スナップショットの後に仮想マシンが QCOW2 ボリュームにデータを書き込むと、関連するブロックがバックアップボリュームから読み取られ、新しい情報で変更されて、新しいスナップショット QCOW2 ボリュームに書き込まれます。次に、新しい場所を指すように地図が更新されます。

### RAW

RAW ストレージ形式には、RAW 形式で保存された仮想ディスクイメージにフォーマットが適用されないことにおいて、QCOW2 よりもパフォーマンス上の利点があります。RAW 形式で保存されたディスクイメージでの仮想マシンのデータ操作では、ホストへの追加作業は必要ありません。仮想

マシンが仮想ディスクの特定のオフセットにデータを書き込む場合、I/O はバッキングファイルまたは論理ボリュームの同じオフセットに書き込まれます。

Raw 形式では、ストレージレイから外部で管理されているシンプロビジョニングされた LUN を使用しない限り、定義されたイメージのスペース全体を事前に割り当てる必要があります。

## 2.5. 仮想ディスクイメージストレージ割り当てポリシー

### 事前に割り当てられたストレージ

仮想ディスクイメージに必要なすべてのストレージは、仮想マシンの作成前に割り当てられます。仮想マシン用に 20GB のディスクイメージが作成される場合、ディスクイメージは 20GB のストレージドメイン容量を使用します。事前に割り当てられたディスクイメージは拡大できません。ストレージの事前割り当ては、実行時にストレージの割り当てが行われなため、書き込み時間が短縮されることを意味しますが、柔軟性が犠牲になります。この方法でストレージを割り振ると、ストレージをオーバーコミットする Red Hat Virtualization Manager の容量が減少します。ストレージの遅延に対する許容度が低い、高強度の I/O タスクに使用される仮想マシンには、事前に割り当てられたストレージをお勧めします。一般に、サーバー仮想マシンはこの説明に適合します。

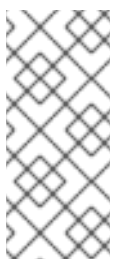


#### 注記

ストレージバックエンドによって提供されるシンプロビジョニング機能を使用している場合でも、仮想マシンのストレージをプロビジョニングするときに、管理ポータルから事前に割り当てられたストレージを選択する必要があります。

### 部分的に割り当てられたストレージ

仮想ディスクイメージのサイズの上限は、仮想マシンの作成時に設定されます。最初は、ディスクイメージはストレージドメインの容量を使用していません。上限に達するまで、仮想マシンがデータをディスクに書き込むにつれて、使用量は増加します。ディスクイメージ内のデータが削除されても、容量はストレージドメインに戻されません。わずかに割り当てられたストレージは、ストレージの遅延にある程度の許容度がある低または中強度の I/O タスクを持つ仮想マシンに適しています。一般に、デスクトップ仮想マシンはこの説明に適合します。



#### 注記

シンプロビジョニング機能がストレージバックエンドによって提供される場合は、シンプロビジョニングの推奨される実装として使用する必要があります。ストレージは、事前に割り当てられたグラフィカルユーザーインターフェイスからプロビジョニングし、シンプロビジョニングをバックエンドソリューションに任せる必要があります。

## 2.6. RED HAT VIRTUALIZATION のストレージメタデータバージョン

Red Hat Virtualization は、ストレージドメインに関する情報をストレージドメイン自体のメタデータとして保存します。Red Hat Virtualization の各メジャーリリースでは、ストレージメタデータの実装が改善されています。

- V1 メタデータ (Red Hat Virtualization 2.x シリーズ)

各ストレージドメインには、独自の構造を記述するメタデータと、仮想ディスクイメージのバックアップに使用されるすべての物理ボリュームの名前が含まれています。

マスタートメインには、ストレージプール内のすべてのドメインと物理ボリューム名のメタデータが追加で含まれています。このメタデータの合計サイズは 2 KB に制限されており、プールに含めることができるストレージドメインの数が制限されます。

テンプレートと仮想マシンのベースイメージは読み取り専用です。

V1 メタデータは、NFS、iSCSI、および FC ストレージドメインに適用できます。

- **V2 メタデータ (Red Hat Enterprise Virtualization 3.0)**

すべてのストレージドメインとプールのメタデータは、論理ボリュームに書き込まれるのではなく、論理ボリュームタグとして保存されます。仮想ディスクボリュームに関するメタデータは、引き続きドメインの論理ボリュームに保存されます。

物理ボリューム名はメタデータに含まれなくなりました。

テンプレートと仮想マシンのベースイメージは読み取り専用です。

V2 メタデータは、iSCSI および FC ストレージドメインに適用できます。

- **V3 metadata (Red Hat Enterprise Virtualization 3.1+)**

すべてのストレージドメインとプールのメタデータは、論理ボリュームに書き込まれるのではなく、論理ボリュームタグとして保存されます。仮想ディスクボリュームに関するメタデータは、引き続きドメインの論理ボリュームに保存されます。

仮想マシンとテンプレートのベースイメージは読み取り専用ではなくなりました。この変更により、ライブスナップショット、ライブストレージの移行、およびスナップショットからのクローン作成が可能になります。

英語以外のボリューム名に対して、Unicode メタデータのサポートが追加されました。

V3 メタデータは、NFS、GlusterFS、POSIX、iSCSI、および FC ストレージドメインに適用できます。

## 2.7. RED HAT VIRTUALIZATION でのストレージドメインの自動リカバリー

Red Hat Virtualization 環境のホストは、各ドメインからメタデータを読み取ることにより、データセンターのストレージドメインを監視します。データセンター内のすべてのホストがストレージドメインにアクセスできないと報告すると、ストレージドメインは非アクティブになります。

Manager は、非アクティブなストレージドメインを切断するのではなく、たとえば一時的なネットワークの停止などにより、ストレージドメインが一時的に非アクティブになっていると見なします。5 分ごとに、Manager は非アクティブなストレージドメインの再アクティブ化を試みます。

ストレージ接続の中断の原因を修正するために管理者の介入が必要になる場合がありますが、接続が復元されると、マネージャーはストレージドメインの再アクティブ化を処理します。

## 2.8. STORAGE POOL MANAGER

Red Hat Virtualization は、メタデータを使用してストレージドメインの内部構造を記述します。構造メタデータは、各ストレージドメインのセグメントに書き込まれます。ホストは、単一のライターと複数のリーダーの設定に基づいて、ストレージドメインのメタデータを操作します。ストレージドメインの構造メタデータは、イメージとスナップショットの作成と削除、およびボリュームとドメインの拡張を追跡します。

データドメインの構造を変更できるホストは、Storage Pool Manager (SPM) と呼ばれます。SPM は、ディスクイメージの作成と削除、スナップショットの作成とマージ、ストレージドメイン間でのイメージのコピー、テンプレートの作成、ブロックデバイスのストレージ割り当てなど、データセンター内のすべてのメタデータの変更を調整します。データセンターごとに1つのSPMがあります。他のすべてのホストは、ストレージドメインの構造メタデータのみを読み取ることができます。

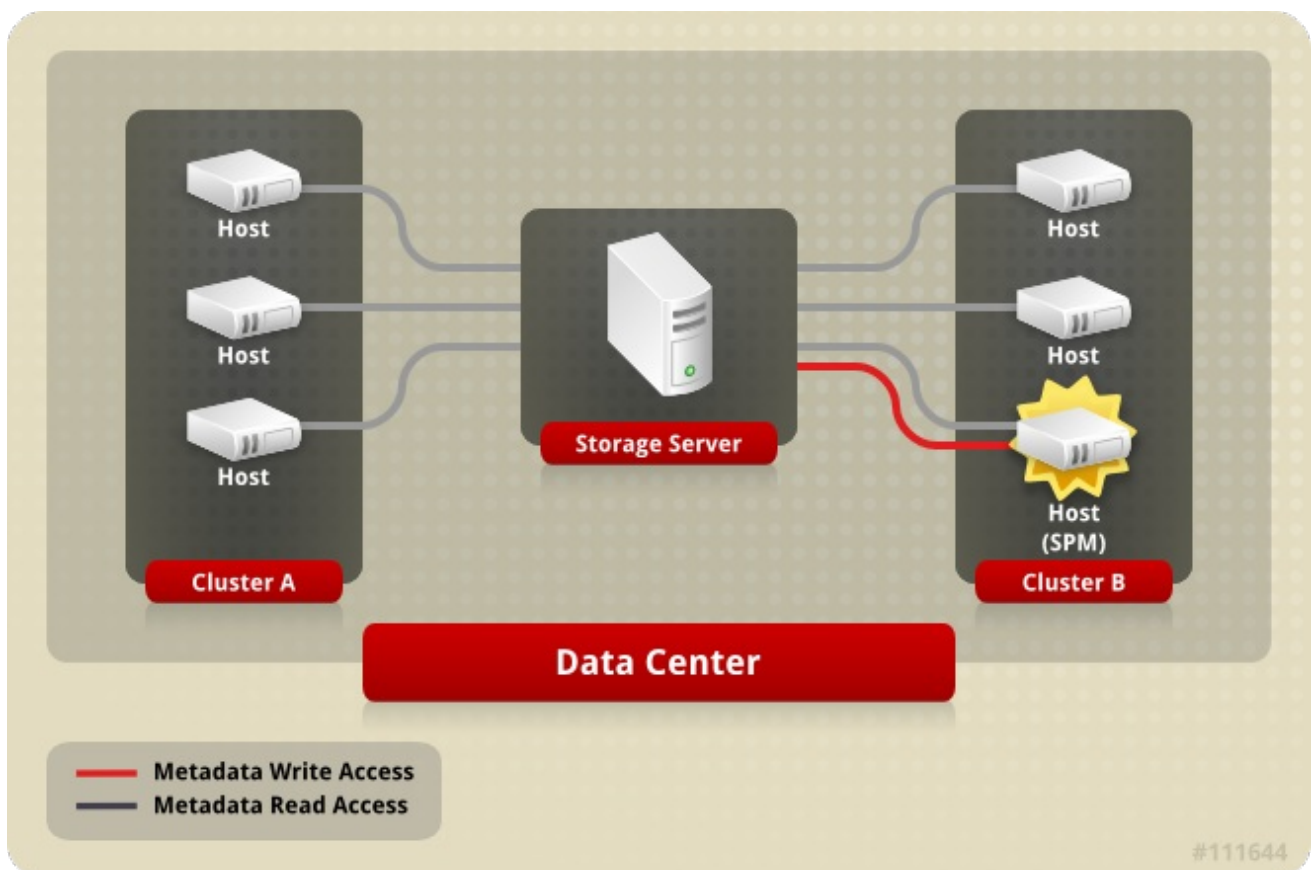
ホストはSPMとして手動で選択することも、Red Hat Virtualizationによって割り当てられることもできます。Managerは、潜在的なSPMホストがストレージ中心リースを想定してSPMのロールを割り当てます。リースにより、SPMホストはストレージメタデータを書き込むことができます。Managerまたはホストによって追跡されるのではなく、ストレージドメインに書き込まれるため、ストレージ中心です。ストレージ中心リースは、**leases** と呼ばれるマスターストレージドメイン内の特別な論理ボリュームに書き込まれます。データドメインの構造に関するメタデータは、**メタデータ** と呼ばれる特別な論理ボリュームに書き込まれます。**メタデータ** 論理ボリュームへの変更は、**leases** 論理ボリュームによって保護されます。

ManagerはVDSMを使用してホストに**spmStart** コマンドを実行し、そのホスト上のVDSMがストレージ中心リースを想定しようとします。ホストが成功すると、ホストはSPMになり、Red Hat Virtualization Managerが新しいホストにSPMのロールを引き受けるように要求するまで、ストレージ中心のリースを保持します。

次の場合、マネージャーはSPMのロールを別のホストに移動します。

- SPMホストはすべてのストレージドメインにアクセスできるわけではありませんが、マスターストレージドメインにアクセスできます。
- ストレージ接続が失われたか、リースボリュームがいっぱいで、書き込み操作を実行できないため、SPMホストはリースを更新できません。
- SPMホストがクラッシュしました。

図2.1 Storage Pool Managerは、構造メタデータを排他的に書き込みます。



## 2.9. STORAGE POOL MANAGER の選択プロセス

ホストに Storage Pool Manager (SPM) のロールが手動で割り当てられていない場合、SPM 選択プロセスは Red Hat Virtualization によって開始および管理されます。

まず、Red Hat Virtualization Manager は、VDSM がどのホストにストレージ中心のリースがあるかを確認するように要求します。

Red Hat Virtualization Manager は、ストレージドメインの最初の作成以降の SPM 割り当ての履歴を追跡します。SPM ロールの可用性は、次の 3 つの方法で確認されます。

- "getSPMstatus" コマンド: マネージャーは VDSM を使用して、最後に SPM ステータスがあったホストを確認し、"SPM"、"Contending"、または "Free" のいずれかを受け取ります。
- ストレージドメインのメタデータボリュームには、SPM ステータスの最後のホストが含まれています。
- ストレージドメインのメタデータボリュームには、SPM ステータスの最後のホストのバージョンが含まれています。

運用可能で応答性の高いホストがストレージ中心のリースを保持している場合、Red Hat Virtualization Manager は管理者ポータルでそのホスト SPM をマークします。それ以上のアクションは実行されません。

SPM ホストが応答しない場合は、到達不能と見なされます。ホストに電源管理が設定されている場合、ホストは自動的にフェンスされます。そうでない場合は、手動フェンシングが必要です。以前の Storage Pool Manager がフェンスで囲まれるまで、Storage Pool Manager のロールを新しいホストに割り当てることはできません。

SPM のロールとストレージ中心のリースが空いている場合、Red Hat Virtualization Manager はそれらをデータセンター内のランダムに選択された運用ホストに割り当てます。

新しいホストで SPM ロールの割り当てが失敗した場合、Red Hat Virtualization Manager は、操作が失敗したホストを含むリストにホストを追加し、これらのホストを SPM ロールの対象外としてマークします。このリストは、次の SPM 選択プロセスの開始時にクリアされるため、すべてのホストが再び適格になります。

Red Hat Virtualization Manager は、SPM の選択が成功するまで、失敗したホストのリストにないランダムに選択されたホストが、Storage Pool Manager ロールとストレージ中心のリースを引き継ぐように要求し続けます。

現在の SPM が応答しないか、その責任を果たせなくなるたびに、Red Hat Virtualization は Storage Pool Manager の選択プロセスを開始します。

## 2.10. RED HAT VIRTUALIZATION における排他的リソースと SANLOCK

Red Hat Virtualization 環境の特定のリソースには、排他的にアクセスする必要があります。

SPM のロールはそのようなリソースの 1 つです。複数のホストが SPM になると、同じデータが 2 つの場所から同時に変更される可能性があるため、データが破損するリスクがあります。

Red Hat Enterprise Virtualization 3.1 以前では、SPM の除外は、**セーフリース** と呼ばれる VDSM 機能を使用して維持および追跡されていました。リースは、データセンター内のすべてのストレージドメインの特別な領域に書き込まれました。環境内のすべてのホストは、ネットワークに依存しない方法で SPM ステータスを追跡できます。VDSM のセーフリースは、SPM ロールという 1 つのリソースの排他的のみを維持していました。

Sanlock は同じ機能を提供しますが、SPM ロールをロック可能なリソースの1つとして扱います。Sanlock は、追加のリソースをロックできるため、より柔軟性があります。

リソースのロックが必要なアプリケーションは、Sanlock に登録できます。登録されたアプリケーションは、Sanlock が自分に代わってリソースをロックするように要求できるため、他のアプリケーションはそのリソースにアクセスできません。たとえば、VDSM が SPM ステータスをロックする代わりに、VDSM は Sanlock にロックするように要求するようになりました。

ロックは、ロックスペースのディスクで追跡されます。ストレージドメインごとに1つのロックスペースがあります。SPM リソースのロックの場合、各ホストの liveness は、ストレージへの接続時に Manager から受け取った hostid を更新し、一定の間隔でタイムスタンプをロックスペースに書き込むホストの機能によってロックスペースで追跡されます。ids 論理ボリュームは、各ホストの一意の識別子を追跡し、ホストが hostid を更新するたびに更新されます。SPM リソースは、ライブホストのみが保持できます。

リソースは、leases 論理ボリュームのディスクで追跡されます。ディスク上の表現が、それを取得したプロセスの一意識別子で更新されたときに、リソースが取得されます。SPM ロールの場合、SPM リソースは、それを取得した hostid で更新されます。

各ホストの Sanlock プロセスは、リソースが取得されていることを確認するために1回だけリソースをチェックする必要があります。最初のチェックの後、Sanlock は、ロックされたリソースを持つホストのタイムスタンプが古くなるまで、ロックスペースを監視できます。

Sanlock は、リソースを使用するアプリケーションを監視します。たとえば、VDSM は SPM ステータスとホスト ID について監視されます。ホストが Manager から hostid を更新できない場合、ロックスペース内のすべてのリソースで多様性が失われます。Sanlock はリソースを更新して、リソースが使用されなくなったことを示します。

SPM ホストがストレージドメインのロックスペースにタイムスタンプを一定時間書き込むことができない場合、ホストの Sanlock のインスタンスは VDSM プロセスがそのリソースを解放することを要求します。VDSM プロセスが応答すると、そのリソースが解放され、ロックスペース内の SPM リソースを別のホストが取得できます。

SPM ホスト上の VDSM がリソースを解放する要求に応答しない場合、ホスト上の Sanlock は VDSM プロセスを強制終了します。kill コマンドが失敗した場合、Sanlock は sigkill を使用して VDSM を強制終了しようとしてエスカレーションします。sigkill が失敗した場合、Sanlock は **ウォッチドッグデーモン** に依存してホストを再起動します。

ホストの VDSM がホスト ID を更新し、ロックスペースにタイムスタンプを書き込むたびに、ウォッチドッグデーモンは **ペット** を受け取ります。VDSM がこれを実行できない場合、ウォッチドッグデーモンは pet されなくなります。ウォッチドッグデーモンが一定時間ペットを受信しないと、ホストを再起動します。この最終レベルのエスカレーションに達すると、SPM リソースが解放され、別のホストが取得できることが保証されます。

## 2.11. シンプルビジョニングとストレージのオーバーコミットメント

Red Hat Virtualization Manager は、仮想化環境内のストレージ使用を最適化するためのプロビジョニングポリシーを提供します。シンプルビジョニングポリシーを使用すると、ストレージリソースをオーバーコミットし、仮想化環境の実際のストレージ使用量に基づいてストレージをプロビジョニングできます。

ストレージのオーバーコミットメントとは、ストレージプールで物理的に利用できるよりも多くのストレージを仮想マシンに割り当てることです。一般に、仮想マシンは、割り当てられているストレージよりも少ないストレージを使用します。シンプルビジョニングにより、仮想マシンは、実際にはストレージのごく一部しか割り当てられていない場合でも、仮想マシンに定義されたストレージが完全に割り当てられているかのように動作できます。





## 注記

Red Hat Virtualization Manager は独自のシンプロビジョニング機能を提供しますが、ストレージバックエンドのシンプロビジョニング機能が提供されている場合はそれを使用する必要があります。

ストレージのオーバーコミットをサポートするため、しきい値が VDSM で定義され、論理ストレージの割り当てを実際のストレージ使用量と比較します。このしきい値は、ディスクイメージに書き込まれるデータが、バックアップする論理ボリュームよりも小さいことを確認するために使用されます。QEMU は、論理ボリュームに書き込まれる最大のオフセットを識別します。これは、ストレージの最大使用ポイントを示します。VDSM は、QEMU によってマークされた最大オフセットを監視して、使用量が定義されたしきい値を超えないようにします。VDSM が最大オフセットがしきい値を下回っていることを示し続ける限り、Red Hat Virtualization Manager は、問題の論理ボリュームに操作を続行するのに十分なストレージがあることを認識します。

QEMU が使用量がしきい値制限を超えることを示している場合、VDSM はディスクイメージがその論理ボリュームのサイズに近づくことを Manager と通信します。Red Hat Virtualization Manager は、SPM ホストが論理ボリュームを拡張することを要求します。このプロセスは、データセンターのデータストレージドメインに使用可能なスペースがある限り繰り返すことができます。データストレージドメインの空き容量が不足した場合は、手動でストレージ容量を追加して拡張する必要があります。

## 2.12. 論理ボリューム拡張

Red Hat Virtualization Manager は、シンプロビジョニングを使用して、ストレージプールで使用可能なストレージをオーバーコミットし、物理的に使用可能なストレージよりも多くのストレージを割り当てます。仮想マシンは、動作中にデータを書き込みます。シンプロビジョニングされたディスクイメージを備えた仮想マシンは、最終的に、ディスクイメージをサポートする論理ボリュームが保持できるよりも多くのデータを書き込みます。これが発生すると、論理ボリューム拡張が追加のストレージを提供し、仮想マシンの継続的な操作を容易にするために使用されます。

Red Hat Virtualization は、LVM 上にシンプロビジョニングメカニズムを提供します。QCOW2 形式のストレージを使用する場合、Red Hat Virtualization はホストシステムプロセス `qemu-kvm` に依存して、ディスク上のストレージブロックを論理ブロックに順次マッピングします。これにより、たとえば、1 GB の論理ボリュームでバックアップされた論理 100 GB ディスクの定義が可能になります。qemu-kvm が VDSM によって設定された使用量のしきい値を超えると、ローカル VDSM インスタンスは、論理ボリュームをさらに 1 ギガバイト拡張するように SPM に要求します。ボリューム拡張が必要な仮想マシンを実行しているホスト上の VDSM は、より多くのスペースが必要であることを SPM VDSM に通知します。SPM は論理ボリュームを拡張し、SPM VDSM インスタンスにより、ホスト VDSM はボリュームグループ情報をリフレッシュし、拡張操作が完了したことを認識します。ホストは操作を続行できます。

論理ボリューム拡張では、ホストが他のどのホストが SPM であるかを知っている必要はありません。それは SPM 自体でさえあり得ます。ストレージ拡張通信は、ストレージメールボックスを介して行われます。ストレージメールボックスは、データストレージドメイン上の専用の論理ボリュームです。論理ボリュームを拡張するために SPM を必要とするホストは、ストレージメールボックス内のその特定のホストに指定された領域にメッセージを書き込みます。SPM は、受信メールを定期的に読み取り、要求された論理ボリューム拡張を実行し、送信メールに応答を書き込みます。リクエストを送信した後、ホストは 2 秒ごとに受信メールの応答を監視します。ホストは、論理ボリューム拡張要求への正常な応答を受信すると、デバイスマッパーの論理ボリュームマップを更新して、新しく割り当てられたストレージを認識します。

ストレージプールで使用可能な物理ストレージがほぼ使い果たされると、リソースを補充する手段がなくても、複数のイメージで使用可能なストレージが不足する可能性があります。ストレージを使い切るストレージプールにより、QEMU は **enospc** エラー を返します。これは、デバイスに利用可能なスト



---

レージがなくなったことを示します。この時点で、実行中の仮想マシンは自動的に一時停止され、ボリュームグループに新しい LUN を追加するには手動による介入が必要になります。

新しい LUN がボリュームグループに追加されると、Storage Pool Manager は追加のストレージをそれを必要とする論理ボリュームに自動的に分散します。追加のリソースの自動割り当てにより、関連する仮想マシンは、中断されることなく自動的に操作を続行したり、停止した場合に操作を再開したりできます。

## 第3章 ネットワーク

### 3.1. ネットワークアーキテクチャー

Red Hat Virtualization ネットワークは、基本的なネットワーク、クラスター内のネットワーク、およびホストネットワーク設定の観点で説明できます。基本的なネットワーク用語は、ネットワークを容易にする基本的なハードウェアおよびソフトウェア要素を対象としています。クラスター内のネットワークには、ホスト、論理ネットワーク、仮想マシンなどのクラスターレベルのオブジェクト間のネットワーク対話が含まれます。ホストネットワーク設定では、ホスト内のネットワークでサポートされる設定について説明します。

適切に設計および構築されたネットワークにより、高帯域幅タスクが適切な帯域幅を受信し、ユーザーの対話がレイテンシーによって妨げられず、仮想マシンを移行ドメイン内で正常に移行できるようになります。ネットワークの構築が不適切に作成されると、許容できないレイテンシーや、ネットワークフラッキングによる移行およびクローン作成の失敗が発生する可能性があります。

### 3.2. 概要：基本的なネットワーク用語

Red Hat Virtualization は、以下を使用して、仮想マシン、仮想化ホスト、およびより広範なネットワーク間のネットワーク機能を提供します。

- ネットワークインターフェイスコントローラー(NIC)
- A Bridge
- ボンド
- 仮想 NIC
- 仮想 LAN (VLAN)

NIC、ブリッジ、および VNIC は、ホスト、仮想マシン、ローカルエリアネットワーク、およびインターネット間のネットワーク通信を可能にします。ボンドと VLAN は、セキュリティー、フォールトトレランス、およびネットワーク容量を強化するために任意で実装されます。

### 3.3. ネットワークインターフェイスコントローラー

NIC (ネットワークインターフェイスコントローラー) は、コンピューターをコンピューターネットワークに接続するネットワークアダプターまたは LAN アダプターです。NIC は、マシンの物理層とデータリンク層の両方で動作し、ネットワーク接続を可能にします。Red Hat Virtualization 環境のすべての仮想化ホストには少なくとも1つの NIC がありますが、ホストには2つ以上の NIC があるのが一般的です。

1つの物理 NIC には、複数の仮想 NIC (VNIC)を論理的に接続できます。仮想 NIC は、仮想マシンの物理ネットワークインターフェイスとして機能します。VNIC とそれをサポートする NIC を区別するために、Red Hat Virtualization Manager は各 VNIC に一意の MAC アドレスを割り当てます。

### 3.4. ブリッジ

ブリッジ は、パケット交換ネットワークでパケット転送を使用するソフトウェアデバイスです。ブリッジングにより、複数のネットワークインターフェイスデバイスが1つの NIC の接続を共有し、ネットワーク上で個別の物理デバイスとして表示されます。ブリッジはパケットのソースアドレスを調べて、

関連するターゲットアドレスを決定します。ターゲットアドレスが決定されると、ブリッジは将来の参照のために場所をテーブルに追加します。これにより、ホストはネットワークトラフィックをブリッジのメンバーである仮想マシンに関連付けられた VNIC にリダイレクトできます。

Red Hat Virtualization では、論理ネットワークはブリッジを使用して実装されます。これは、IP アドレスを受信するホスト上の物理インターフェイスではなく、ブリッジです。ブリッジに関連付けられた IP アドレスは、接続にブリッジを使用する仮想マシンと同じサブネット内にある必要はありません。ブリッジに、ブリッジを使用する仮想マシンと同じサブネット上の IP アドレスが割り当てられている場合、ホストは仮想マシンによって論理ネットワーク内でアドレス指定できます。原則として、仮想化ホストでネットワーク公開されたサービスを実行することは推奨されません。ゲストは VNIC によって論理ネットワークに接続され、ホストは NIC を使用して論理ネットワークのリモート要素に接続されます。各ゲストは、DHCP または静的により、VNIC の IP アドレスを独立して設定できます。ブリッジはホスト外のオブジェクトに接続できますが、このような接続は必須ではありません。

カスタムプロパティは、ブリッジとイーサネット接続の両方に定義できます。VDSM は、ネットワーク定義とカスタムプロパティをセットアップネットワークフックスクリプトに渡します。

### 3.5. ボンド

ボンディングは、複数のネットワークインターフェイスカードを1つのソフトウェア定義デバイスに集約したものです。ボンディングされたネットワークインターフェイスは、ボンディングに含まれるネットワークインターフェイスカードの伝送機能を組み合わせて単一のネットワークインターフェイスとして機能するため、単一のネットワークインターフェイスカードよりも高速な伝送速度を提供できます。また、ボンド自体が失敗するには、ボンド内のすべてのネットワークインターフェイスカードが失敗する必要があるため、ボンディングによってフォールトトレランスが向上します。ただし、1つの制限は、ボンディングされたネットワークインターフェイスを形成するネットワークインターフェイスカードは、ボンド内のすべてのネットワークインターフェイスカードが同じオプションとモードをサポートするように、同じメーカーとモデルである必要があることです。

結合の packets 分散アルゴリズムは、使用される結合モードによって決定されます。



#### 重要

モード 1、2、3、および 4 は、仮想マシン (ブリッジ) と非仮想マシン (ブリッジレス) の両方のネットワークタイプをサポートします。モード 0、5、および 6 は、非仮想マシン (ブリッジレス) ネットワークのみをサポートします。

#### ボンディングモード

Red Hat Virtualization はデフォルトでモード 4 を使用しますが、以下の一般的なボンディングモードをサポートします。

##### **Mode 0 (round-robin policy)**

ネットワークインターフェイスカードを介してパケットを順番に送信します。パケットは、ボンディングで最初に利用可能なネットワークインターフェイスカードで始まり、ボンドで最後に使用可能なネットワークインターフェイスカードで終わるループで送信されます。その後のすべてのループは、最初に使用可能なネットワークインターフェイスカードから始まります。モード 0 はフォールトトレランスを提供し、ボンド内のすべてのネットワークインターフェイスカード間で負荷を分散します。ただし、モード 0 はブリッジと組み合わせて使用できないため、仮想マシンの論理ネットワークとの互換性はありません。

##### **モード 1 (active-backup ポリシー)**

1 枚のネットワークインターフェイスカードがアクティブなまま、すべてのネットワークインターフェイスカードをバックアップ状態に設定します。アクティブなネットワークインターフェイスカードに障害が発生した場合、バックアップネットワークインターフェイスカードの 1 つが、ボンド

内の唯一のアクティブなネットワークインターフェイスカードとしてそのネットワークインターフェイスカードに置き換わります。モード1のボンディングのMACアドレスは、1つのポートにのみ表示され、ボンディングのMACアドレスがアクティブなネットワークインターフェイスカードを反映するように変更された場合に混乱を防ぐことができます。モード1はフォールトトレランスを提供し、Red Hat Virtualization でサポートされています。

### Mode 2 (XOR policy)

ネットワークインターフェイスカードのスレーブ数をモジュロとして、送信元および宛先 MAC アドレスでの XOR 操作の結果に基づいて、パケットの送信に使用するネットワークインターフェイスカードを選択します。この計算により、使用される宛先 MAC アドレスごとに同じネットワークインターフェイスカードが確実に選択されます。モード2は、フォールトトレランスと負荷分散を提供し、Red Hat Virtualization でサポートされています。

### Mode 3 (broadcast policy)

すべてのパケットをすべてのネットワークインターフェイスカードに送信します。モード3はフォールトトレランスを提供し、Red Hat Virtualization でサポートされています。

### Mode 4 (IEEE 802.3ad policy)

インターフェイスが同じ速度とデュプレックス設定を共有するアグリゲーショングループを作成します。モード4は、IEEE 802.3ad 仕様に従って、アクティブなアグリゲーショングループ内のすべてのネットワークインターフェイスカードを使用し、Red Hat Virtualization でサポートされます。

### Mode 5 (adaptive send load balancing policy)

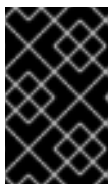
送信トラフィックの分散がボンド内の各ネットワークインターフェイスカードの負荷を考慮し、現在のネットワークインターフェイスカードがすべての受信トラフィックを受信するようにします。トラフィックの受信に割り当てられたネットワークインターフェイスカードに障害が発生した場合、別のネットワークインターフェイスカードが着信トラフィックの受信のロールに割り当てられます。モード5はブリッジと組み合わせて使用できないため、仮想マシンの論理ネットワークとは互換性がありません。

### Mode 6 (adaptive load balancing policy)

特別なスイッチ要件なしで、モード5 (適応型送信負荷分散ポリシー) と IPv4 トラフィックの受信負荷分散を組み合わせます。ARP ネゴシエーションは、受信負荷のバランスを取るために使用されます。モード6はブリッジと組み合わせて使用できないため、仮想マシンの論理ネットワークとは互換性がありません。

## 3.6. ボンディング用のスイッチ設定

スイッチの設定は、ハードウェアの要件によって異なります。オペレーティングシステムのデプロイメントおよびネットワーク設定ガイドを参照してください。



### 重要

すべてのタイプのスイッチに関しては、Cisco Port Aggregation Protocol (PAgP) プロトコルではなく、Link Aggregation Control Protocol (LACP) プロトコルでスイッチボンディングを設定することが重要です。

## 3.7. 仮想ネットワークインターフェイスカード

仮想ネットワークインターフェイスカードは、ホストの物理ネットワークインターフェイスカードに基

づく仮想ネットワークインターフェイスです。各ホストに複数のネットワークインターフェイスカードを含めることができ、各ネットワークインターフェイスカードは複数の仮想ネットワークインターフェイスカードのベースにすることができます。

仮想ネットワークインターフェイスカードを仮想マシンに割り当てると、Red Hat Virtualization Manager は、仮想ネットワークインターフェイスカードが接続されている仮想マシン、仮想ネットワークインターフェイスカード自体、および仮想ネットワークインターフェイスカードのベースとなる物理ホストネットワークインターフェイスカードの間に、いくつかの関連付けを作成します。具体的には、仮想ネットワークインターフェイスカードが仮想マシンに接続されていると、仮想ネットワークインターフェイスカードのベースとなる物理ホストネットワークインターフェイスカードに、新しい仮想ネットワークインターフェイスカードと MAC アドレスが作成されます。次に、仮想ネットワークインターフェイスカードがアタッチされた後に仮想マシンの初回起動時に、**libvirt** は仮想ネットワークインターフェイスカードを PCI アドレスに割り当てます。次に、MAC アドレスおよび PCI アドレスを使用して、仮想マシンの仮想ネットワークインターフェイスカードの名前(**eth0**など)を取得します。

テンプレートまたはスナップショットに基づいて仮想マシンを作成する場合、MAC アドレスを割り当て、それらの MAC アドレスを PCI アドレスに関連付けるプロセスは若干異なります。テンプレートまたはスナップショット用に PCI アドレスがすでに作成されている場合、そのテンプレートまたはスナップショットに基づいて作成された仮想マシンの仮想ネットワークインターフェイスカードは、その順番に割り当てられた PCI アドレスおよび MAC アドレスに従って順序付けられます。テンプレートに PCI アドレスがまだ作成されていない場合、そのテンプレートに基づいて作成された仮想マシンの仮想ネットワークインターフェイスカードは、仮想ネットワークインターフェイスカードの命名順に割り当てられます。スナップショット用に PCI アドレスがまだ作成されていない場合、Red Hat Virtualization Manager は、そのスナップショットに基づいて仮想マシンの仮想ネットワークインターフェイスカードに新しい MAC アドレスを割り当てます。

作成が完了すると、仮想ネットワークインターフェイスカードがネットワークブリッジデバイスに追加されます。ネットワークブリッジデバイスは、仮想マシンが仮想マシンの論理ネットワークに接続する方法です。

仮想化ホストで **ip addr show** コマンドを実行すると、そのホスト上の仮想マシンに関連付けられている仮想ネットワークインターフェイスカードがすべて表示されます。また、論理ネットワークをサポートするために作成されたネットワークブリッジと、ホストが使用するネットワークインターフェイスカードも表示されます。

```
[root@rhev-host-01 ~]# ip addr show
1: lo: <LOOPBACK,UP,LOWER_UP> mtu 16436 qdisc noqueue state UNKNOWN
    link/loopback 00:00:00:00:00:00 brd 00:00:00:00:00:00
    inet 127.0.0.1/8 scope host lo
    inet6 ::1/128 scope host
        valid_lft forever preferred_lft forever
2: eth0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc pfifo_fast state UP qlen 1000
    link/ether 00:21:86:a2:85:cd brd ff:ff:ff:ff:ff:ff
    inet6 fe80::221:86ff:fea2:85cd/64 scope link
        valid_lft forever preferred_lft forever
3: wlan0: <NO-CARRIER,BROADCAST,MULTICAST,UP> mtu 1500 qdisc mq state DOWN qlen
1000
    link/ether 00:21:6b:cc:14:6c brd ff:ff:ff:ff:ff:ff
5: ;vdsmdummy;; <BROADCAST,MULTICAST> mtu 1500 qdisc noop state DOWN
    link/ether 4a:d5:52:c2:7f:4b brd ff:ff:ff:ff:ff:ff
6: bond0: <BROADCAST,MULTICAST,MASTER> mtu 1500 qdisc noop state DOWN
    link/ether 00:00:00:00:00:00 brd ff:ff:ff:ff:ff:ff
7: bond4: <BROADCAST,MULTICAST,MASTER> mtu 1500 qdisc noop state DOWN
    link/ether 00:00:00:00:00:00 brd ff:ff:ff:ff:ff:ff
8: bond1: <BROADCAST,MULTICAST,MASTER> mtu 1500 qdisc noop state DOWN
    link/ether 00:00:00:00:00:00 brd ff:ff:ff:ff:ff:ff
```

```

9: bond2: <BROADCAST,MULTICAST,MASTER> mtu 1500 qdisc noop state DOWN
   link/ether 00:00:00:00:00:00 brd ff:ff:ff:ff:ff:ff
10: bond3: <BROADCAST,MULTICAST,MASTER> mtu 1500 qdisc noop state DOWN
   link/ether 00:00:00:00:00:00 brd ff:ff:ff:ff:ff:ff
11: ovirtmgmt: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc noqueue state
UNKNOWN
   link/ether 00:21:86:a2:85:cd brd ff:ff:ff:ff:ff:ff
   inet 10.64.32.134/23 brd 10.64.33.255 scope global ovirtmgmt
   inet6 fe80::221:86ff:fea2:85cd/64 scope link
   valid_lft forever preferred_lft forever

```

コマンドからのコンソール出力には、複数のデバイスが表示されます。1つのループバックデバイス (**lo**)、1つのイーサネットデバイス(**wlan0**)、1つの VDSM ダミーデバイス(;**vdsmdummy**;)、5つのボン  
ドデバイス(**bond0**、**bond4**、**bond1**、**bond2**、**bond3**)、および1つのネットワークブリッジ  
(**ovirtmgmt**)。

仮想ネットワークインターフェイスカードはすべて、ネットワークブリッジデバイスと論理ネット  
ワークのメンバーです。ブリッジメンバーシップは、**brctl show** コマンドを使用して表示できます。

```

[root@rhev-host-01 ~]# brctl show
bridge name bridge id STP enabled interfaces
ovirtmgmt 8000.e41f13b7fdd4 no vnet002
      vnet001
      vnet000
      eth0

```

**brctl show** コマンドの出力は、**virtio** 仮想ネットワークインターフェイスカードが **ovirtmgmt** ブ  
リッジのメンバーであることを示しています。仮想ネットワークインターフェイスカードが関連付けら  
れているすべての仮想マシンが **ovirtmgmt** 論理ネットワークに接続されている。**eth0** ネットワークイ  
ンターフェイスカードも **ovirtmgmt** ブリッジのメンバーです。**eth0** デバイスはスイッチに接続され、  
ホスト以外の接続を提供します。

### 3.8. 仮想 LAN (VLAN)

**VLAN (仮想 LAN)** は、ネットワークパケットに適用できる属性です。ネットワークパケットは、  
番号付き **VLAN** にタグ付けできます。**VLAN** は、スイッチレベルでネットワークトラフィックを完全に  
分離するために使用されるセキュリティー機能です。**VLAN** は完全に分離されており、相互に排他的で  
す。**Red Hat Virtualization Manager** は **VLAN** に対応しており、**VLAN** トラフィックのタグ付けとリダイ  
レクトが可能です。ただし、**VLAN** 実装には **VLAN** をサポートするスイッチが必要です。

スイッチレベルでは、ポートに VLAN 指定が割り当てられます。スイッチは、特定のポートから発信されたトラフィックに VLAN タグを適用し、トラフィックを VLAN の一部としてマークし、応答が同じ VLAN タグを確実に伝送するようにします。VLAN は、複数のスイッチにまたがって拡張できます。スイッチ上の VLAN タグ付きネットワークトラフィックは、正しい VLAN で指定されたポートに接続されたマシンを除いて、完全に検出できません。特定のポートを複数の VLAN にタグ付けすると、複数の VLAN からのトラフィックを単一のポートに送信し、トラフィックを受け取るマシンのソフトウェアを使用して解読することができます。

### 3.9. ネットワークラベル

ネットワークラベルを使用すると、論理ネットワークの作成および管理、およびそれらの論理ネットワークを物理ホストネットワークインターフェイスおよびボンディングに関連付けることに関連するいくつかの管理タスクを大幅に簡素化できます。

ネットワークラベルは、論理ネットワークまたは物理ホストネットワークインターフェイスに接続できるプレーンテキストの人間が読める形式のラベルです。ラベルの長さには厳密な制限はありませんが、小文字と大文字、アンダースコア、ハイフンの組み合わせを使用する必要があります。スペースや特殊文字は使用できません。

論理ネットワークまたは物理ホストネットワークインターフェイスにラベルを付けると、以下のように、同じラベルが接続されている他の論理ネットワークまたは物理ホストネットワークインターフェイスとの関連付けが作成されます。

#### ネットワークラベルの関連付け

- 論理ネットワークにラベルを付けると、その論理ネットワークは、指定されたラベルを持つ物理ホストネットワークインターフェイスに自動的に関連付けられます。
- 物理ホストネットワークインターフェイスにラベルを付けると、指定されたラベルの論理ネットワークはすべて、その物理ホストネットワークインターフェイスに自動的に関連付けられます。
- 論理ネットワークまたは物理ホストネットワークインターフェイスに接続されているラベルを変更すると、ラベルを削除して新しいラベルを追加するのと同じように機能します。関連する論理ネットワークまたは物理ホストネットワークインターフェイス間の関連付けが更新されます。

#### ネットワークラベルとクラスター

- ラベル付きの論理ネットワークがクラスターに追加され、そのクラスター内に同じラベルの物理ホストネットワークインターフェイスがある場合、論理ネットワークはその物理ホスト

ネットワークインターフェイスに自動的に追加されます。

- ラベル付き論理ネットワークがクラスターから切り離され、そのクラスター内に同じラベルの物理ホストネットワークインターフェイスがある場合、論理ネットワークはその物理ホストネットワークインターフェイスから自動的に切り離されます。

#### ネットワークラベルとロールを持つ論理ネットワーク

- ラベル付き論理ネットワークがディスプレイネットワークまたは移行ネットワークとして機能するように割り当てられると、その論理ネットワークは DHCP を使用して物理ホストネットワークインターフェイス上に設定され、論理ネットワークに IP アドレスを割り当てることができます。

ロールネットワーク (たとえば、移行ネットワークやディスプレイネットワーク) にラベルを設定すると、そのネットワークがすべてのホストに大量に展開されます。このようなネットワークの大量追加は、DHCP を使って実現しています。多くの静的 IP アドレスを入力するタスクのスケールでない性質のため、この大量展開の方法は、静的アドレスを入力する方法よりも選択されました。

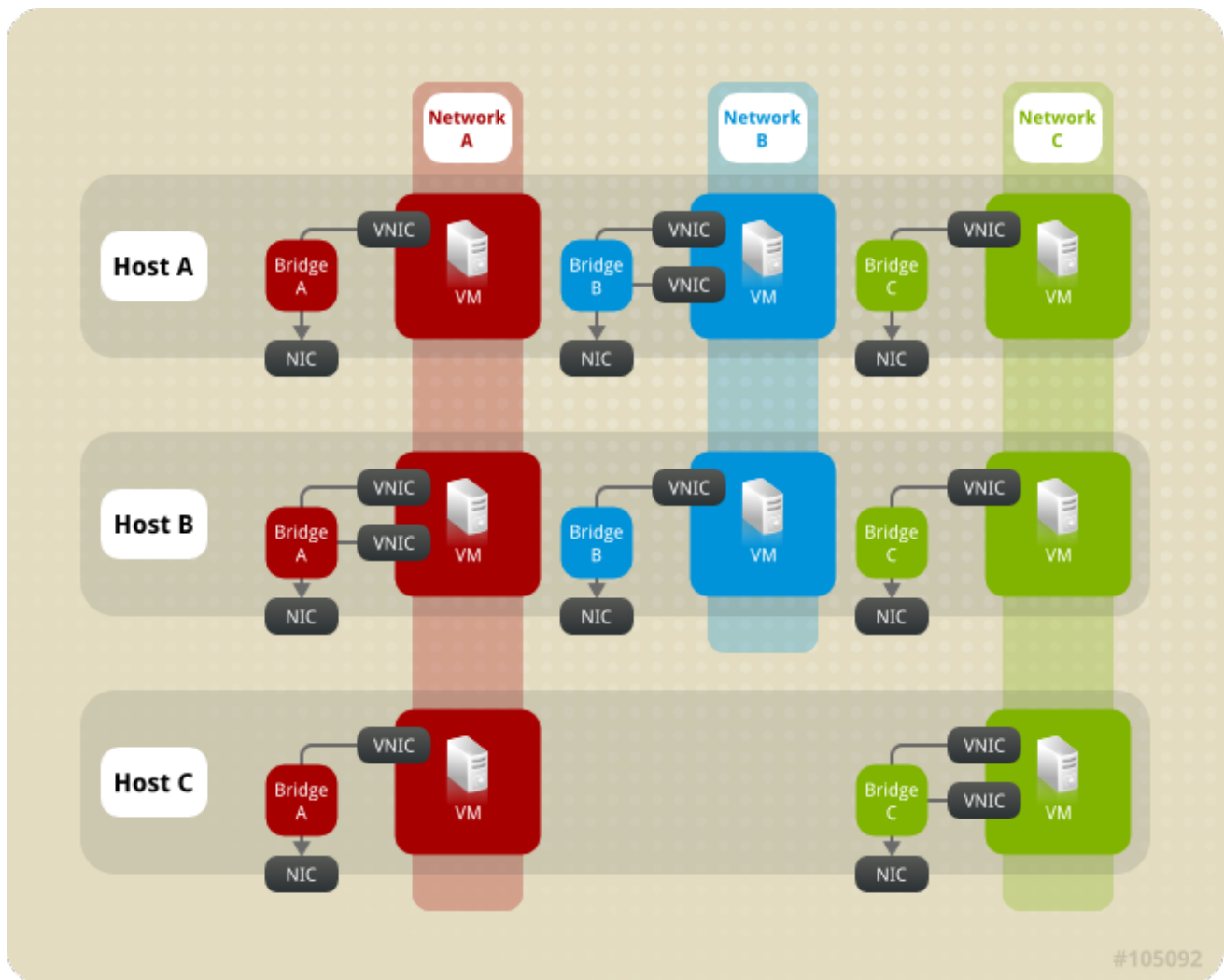
### 3.10. クラスターネットワーク

クラスターレベルのネットワークオブジェクトには、次のものがあります。

- クラスター
- 論理ネットワーク



図3.1 クラスタ内のネットワーキング



データセンターは複数のクラスタの論理グループであり、各クラスタは複数のホストの論理グループです。図3.1「クラスタ内のネットワーキング」単一クラスタの内容を示します。

クラスタ内のホストはすべて、同じストレージドメインにアクセスできます。クラスタ内のホストには、クラスタレベルでの論理ネットワークも適用されます。仮想マシンの論理ネットワークを仮想マシンで使用できるようにするには、Red Hat Virtualization Manager を使用して、クラスタ内のホストごとにネットワークを定義および実装する必要があります。他の論理ネットワークタイプは、それらを使用するホストにのみ実装できます。

マルチホストネットワーク設定では、更新されたネットワーク設定を、ネットワークが割り当てられたデータセンター内のすべてのホストに自動適用します。

### 3.11. 論理ネットワーク

論理ネットワークにより、Red Hat Virtualization 環境はネットワークトラフィックをタイプ別に分離できます。たとえば、ovirtmgmt ネットワークは、Manager とホスト間の管理通信に使用される Red Hat Virtualization のインストール時にデフォルトで作成されます。論理ネットワークの一般的な

使用法は、同様の要件と使用方法を持つネットワークトラフィックをグループ化することです。多くの場合、ストレージネットワークとディスプレイネットワークは、最適化とトラブルシューティングのためにそれぞれのタイプのトラフィックを分離するために管理者によって作成されます。

論理ネットワークの種類は次のとおりです。

- 仮想マシンネットワークトラフィックを伝送する論理ネットワーク、
- 仮想マシンネットワークトラフィックを伝送しない論理ネットワーク、
- オプションの論理ネットワーク
- 必要なネットワーク。

すべての論理ネットワークは、必須または任意のいずれかです。

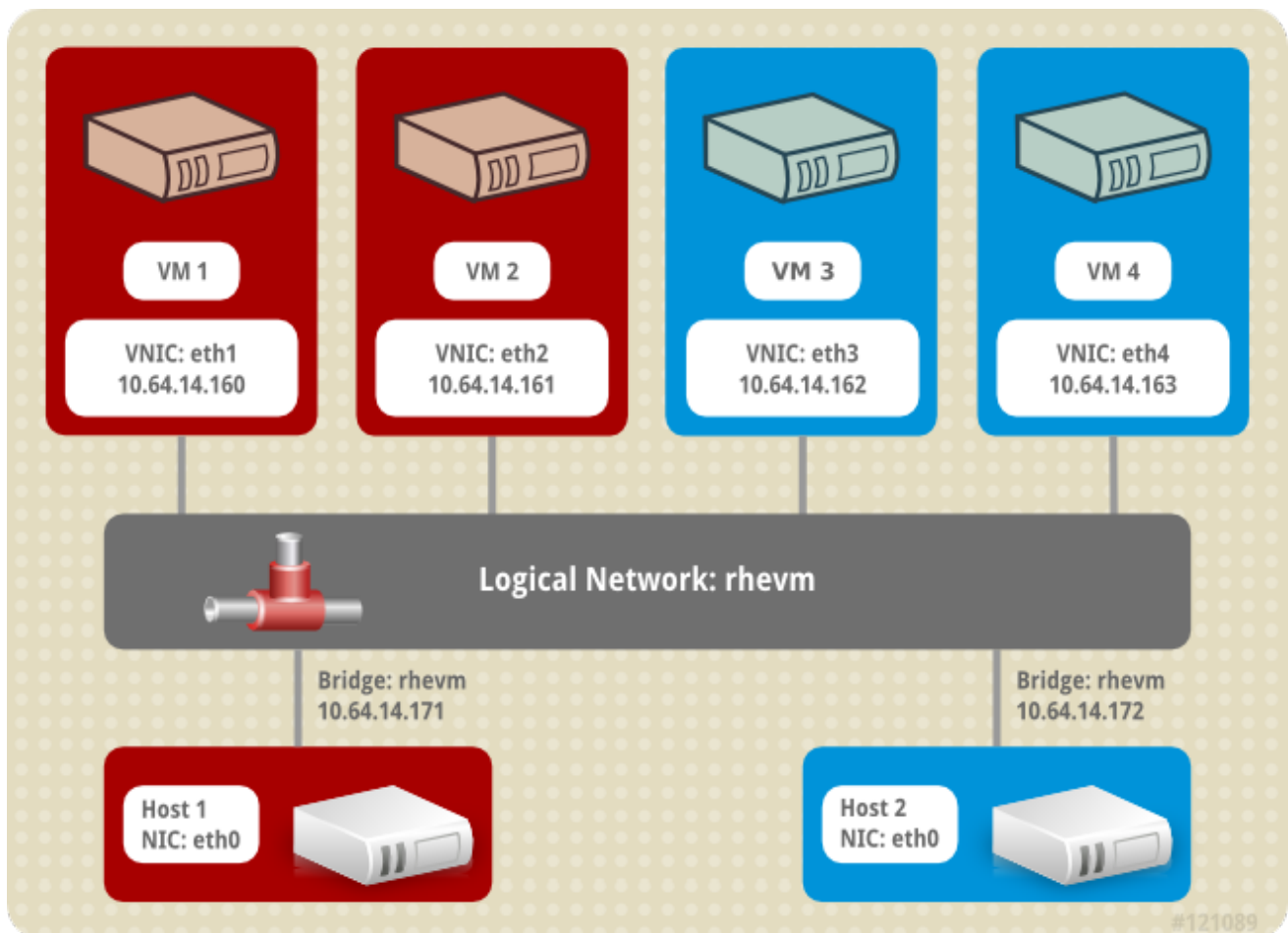
論理ネットワークはデータセンターレベルで定義され、ホストに追加されます。必要な論理ネットワークを動作させるには、特定のクラスター内の全ホストに実装する必要があります。

**Red Hat Virtualization** 環境の各仮想マシンの論理ネットワークは、ホスト上のネットワークブリッジデバイスによってサポートされます。したがって、クラスターに新しい仮想マシン論理ネットワークが定義されている場合は、論理ネットワークを仮想マシンで使用できるようにするには、クラスター内の各ホストに一致するブリッジデバイスを作成する必要があります。**Red Hat Virtualization Manager** は、仮想マシンの論理ネットワークに必要なブリッジを自動的に作成します。

仮想マシンの論理ネットワークをバックアップするために **Red Hat Virtualization Manager** によって作成されたブリッジデバイスは、ホストネットワークインターフェイスに関連付けられます。ブリッジの一部であるホストネットワークインターフェイスにネットワーク接続がある場合、ブリッジに含まれるネットワークインターフェイスはブリッジのネットワーク接続を共有します。仮想マシンを作成して特定の論理ネットワークに配置すると、その論理ネットワークのブリッジに仮想ネットワークカードが含まれます。その後、これらの仮想マシンは相互に、およびブリッジに接続されている他のオブジェクトと通信できます。

仮想マシンのネットワークトラフィックに使用されていない論理ネットワークは、ホストネットワークインターフェイスに直接関連付けられません。

図3.2 ovirtmgmt 論理ネットワーク。



### 例3.1 論理ネットワークの使用例。

Purple というデータセンターには、Pink と呼ばれるクラスターに Red と White という 2 つのホストがあります。Red と White の両方が、すべてのネットワーク機能にデフォルトの論理ネットワークである ovirtmgmt を使用しています。Pink を担当するシステム管理者は、Web サーバーと一部のクライアント仮想マシンを別の論理ネットワークに配置することで、Web サーバーのネットワークテストを分離することを決定します。新しい論理 ネットワーク `_testing` を呼び出すことを決定します。

まず、Purple データセンターの論理ネットワークを定義します。次に、これを Pink クラスターに適用します。論理ネットワークは、メンテナンスモードのホストに実装する必要があります。したがって、管理者はまず実行中の仮想マシンをすべて Red に移行し、White をメンテナンスモードにします。次に、ブリッジに含まれる物理ネットワークインターフェイスに関連付けられたネットワークを編集します。選択したネットワークインターフェイスの Link Status が Down から Non-Operational に変わります。稼働していないステータスは、Pink クラスターの各ホストに物理ネットワークインターフェイスを `network_testing` ネットワークに追加して、対応するブリッジをクラスター内のすべてのノードで設定する必要があります。次に、White をアクティブにし、実行中の仮想マシンをすべて移行して、Red のプロセスを繰り返します。

White と Red の両方に `network_testing` 論理ネットワークが物理ネットワークインターフェイスにブリッジされている場合は、`network_testing` 論理ネットワークが操作 になり、仮想マシンが

使用できるようになります。

### 3.12. 必須ネットワーク、任意のネットワーク、および仮想マシンネットワーク

必要なネットワークは、クラスター内のすべてのホストが使用できる必要がある論理ネットワークです。ホストに必要なネットワークが動作しなくなると、そのホストで実行されている仮想マシンは別のホストに移行されます。この移行の範囲は、選択したスケジューリングポリシーによって異なります。これは、ミッションクリティカルなワークロードを実行している仮想マシンがある場合に役立ちます。

オプションのネットワークは、明示的に **Required** として宣言されていない論理ネットワークです。オプションのネットワークは、それらを使用するホストにのみ実装できます。オプションのネットワークの有無に関わらず、ホストの操作ステータスには影響を及ぼしません。不要なネットワークが動作しなくなった場合、ネットワーク上で実行されている仮想マシンは別のホストに移行されません。これにより、大量の移行によって引き起こされる不要な I/O の過負荷を防ぎます。論理ネットワークが作成され、クラスターに追加されると、**Required** ボックスがデフォルトでチェックされることに注意してください。

ネットワークの **Required** の指定を変更するには、管理ポータルからネットワークを選択し、**Cluster** タブをクリックして、**Manage Networks** ボタンをクリックします。

仮想マシンネットワーク（ユーザーインターフェイスでは **VM ネットワーク** と呼ばれます）は、仮想マシンのネットワークトラフィックのみを伝送するように指定された論理ネットワークです。仮想マシンネットワークは、必須またはオプションにすることができます。オプションの仮想マシンネットワークを使用する仮想マシンは、そのネットワークを持つホストでのみ起動します。

### 3.13. 仮想マシンの接続性

**Red Hat Virtualization** では、仮想マシンの作成時に、仮想マシンの **NIC** が論理ネットワークに配置されます。この時点から、仮想マシンは同じネットワーク上の他の宛先と通信できます。

ホストの観点からは、仮想マシンが論理ネットワークに配置されると、仮想マシンの **NIC** をサポートする **VNIC** が論理ネットワークのブリッジデバイスに追加されます。たとえば、仮想マシンが **ovirtmgmt** 論理ネットワークにある場合、その **VNIC** は仮想マシンを実行するホストの **ovirtmgmt** ブリッジのメンバーとして追加されます。

### 3.14. PORT MIRRORING

ポートミラーリングは、特定の論理ネットワーク上のレイヤー 3 ネットワークトラフィックとホストを仮想マシン上の仮想インターフェイスにコピーします。この仮想マシンは、ネットワークのデバッグ

とチューニング、侵入検知、および同じホストと論理ネットワーク上の他の仮想マシンの動作の監視に使用できます。

コピーされる唯一のトラフィックは、1つのホスト上の1つの論理ネットワークの内部です。ホスト外のネットワークのトラフィックが増えることはありませんが、ポートミラーリングが有効になっている仮想マシンは、他の仮想マシンよりも多くのホスト CPU および RAM を使用します。

ポートミラーリングは、論理ネットワークの vNIC プロファイルで有効または無効にされており、次の制限があります。

- ポートミラーリングが有効になっているプロファイルの vNIC のホットプラグはサポートされていません。
- vNIC プロファイルが仮想マシンに接続されている場合は、ポートミラーリングを変更することができません。

上記の制限があるため、追加の専用 vNIC プロファイルでポートミラーリングを有効にすることが推奨されます。



#### 重要

ポートミラーリングを有効にすると、他のネットワークユーザーのプライバシーが低下します。

### 3.15. ホストネットワーク設定

仮想化ホストのネットワーク設定の一般的なタイプは次のとおりです。

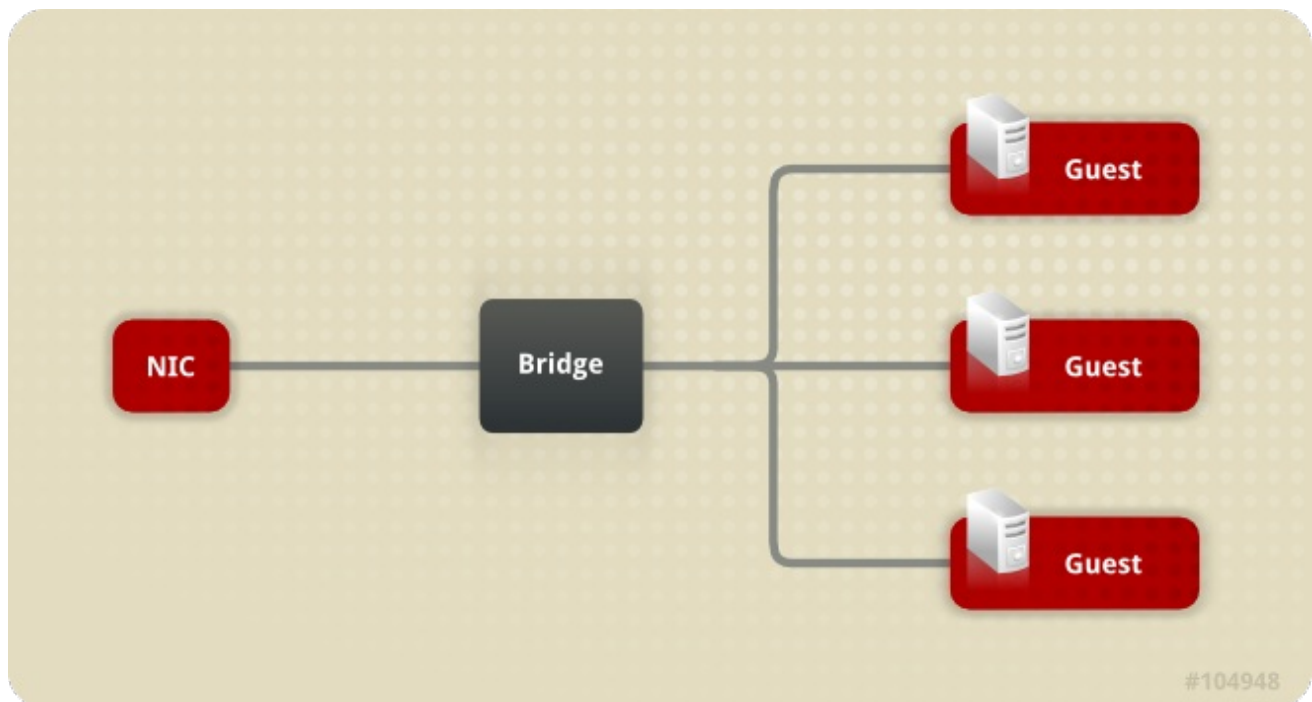
- ブリッジと NIC の設定
- ブリッジ、VLAN、および NIC の設定
- ブリッジ、ボンド、および VLAN の設定

- 複数のブリッジ、複数の VLAN、および NIC の設定

### 3.16. BRIDGE CONFIGURATION

Red Hat Virtualization の最も単純なホスト設定は Bridge および NIC 設定です。図3.3「ブリッジおよび NIC の設定」が示すように、この設定はブリッジを使用して1つ以上の仮想マシン（またはゲスト）をホストの NIC に接続します。

図3.3 ブリッジおよび NIC の設定

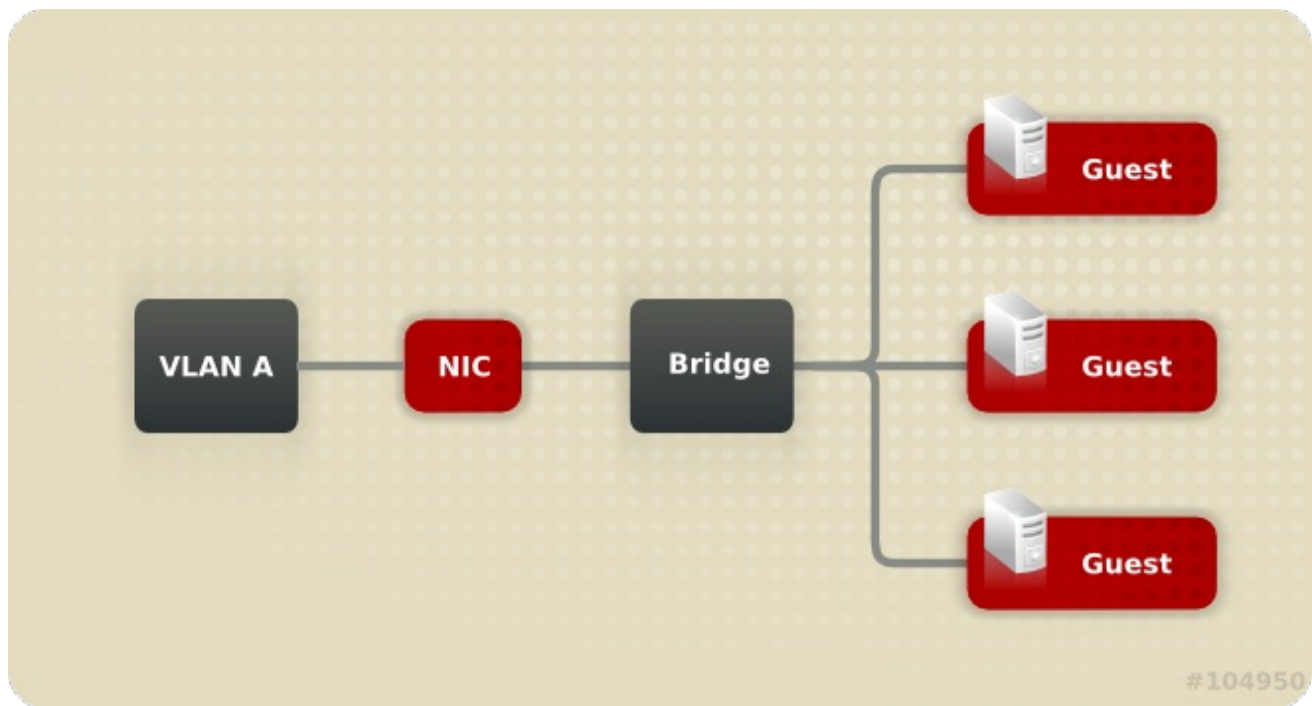


この設定の例は、Red Hat Virtualization Manager のインストール時にブリッジ `ovirtmgmt` の自動作成です。インストール時に、Red Hat Virtualization Manager はホストに VDSM をインストールします。VDSM インストールプロセスでは、ブリッジ `ovirtmgmt` を作成します。次に、`ovirtmgmt` ブリッジはホストの IP アドレスを取得して、ホストの管理通信を有効にします。

### 3.17. VLAN 設定

図3.4「ブリッジ、VLAN、および NIC の設定」は、ホスト NIC およびブリッジに接続するための仮想 LAN (VLAN) を含む代替設定を示しています。

図3.4 ブリッジ、VLAN、および NIC の設定

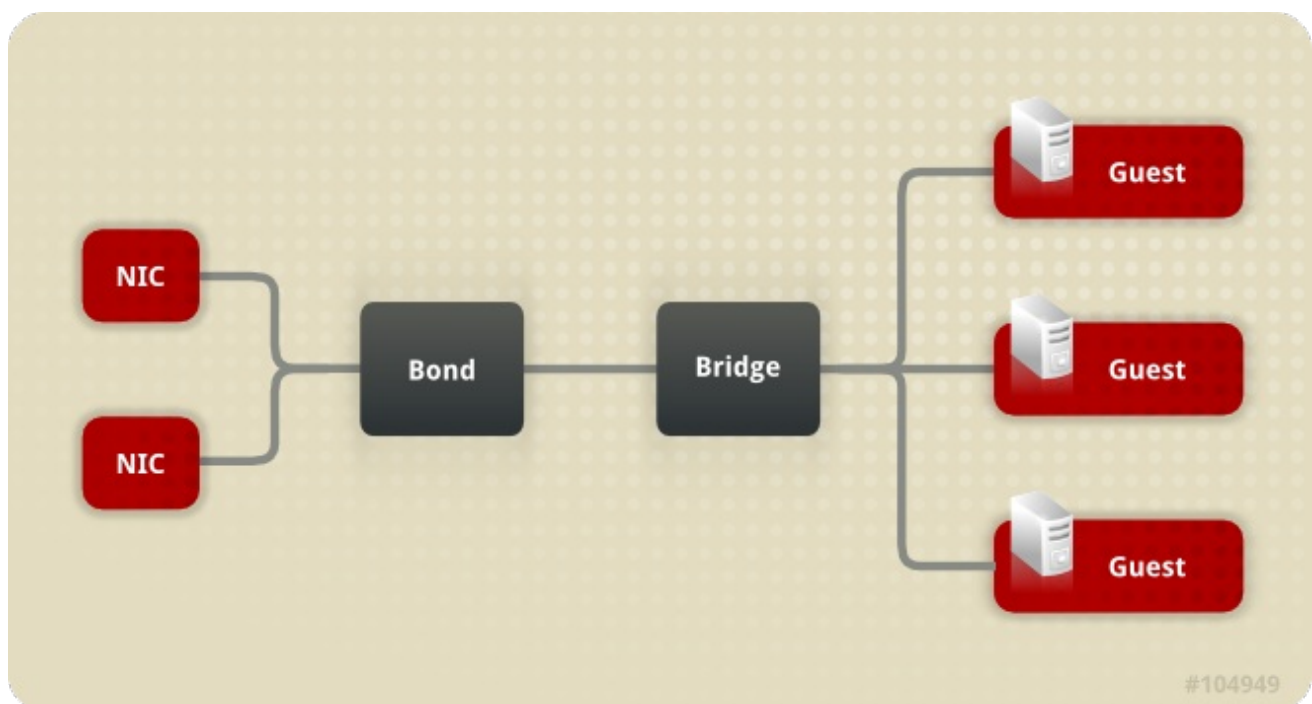


VLAN は、このネットワークを介したデータ転送のための安全なチャンネルと、複数の VLAN を使用して複数のブリッジを単一の NIC に接続するオプションをサポートするために含まれています。

### 3.18. ブリッジおよびボンディングの設定

[図3.5「ブリッジ、ボンド、および NIC の設定」](#) 複数のホスト NIC を同じブリッジおよびネットワークに接続するためのボンディングが含まれる設定を表示します。

図3.5 ブリッジ、ボンド、および NIC の設定



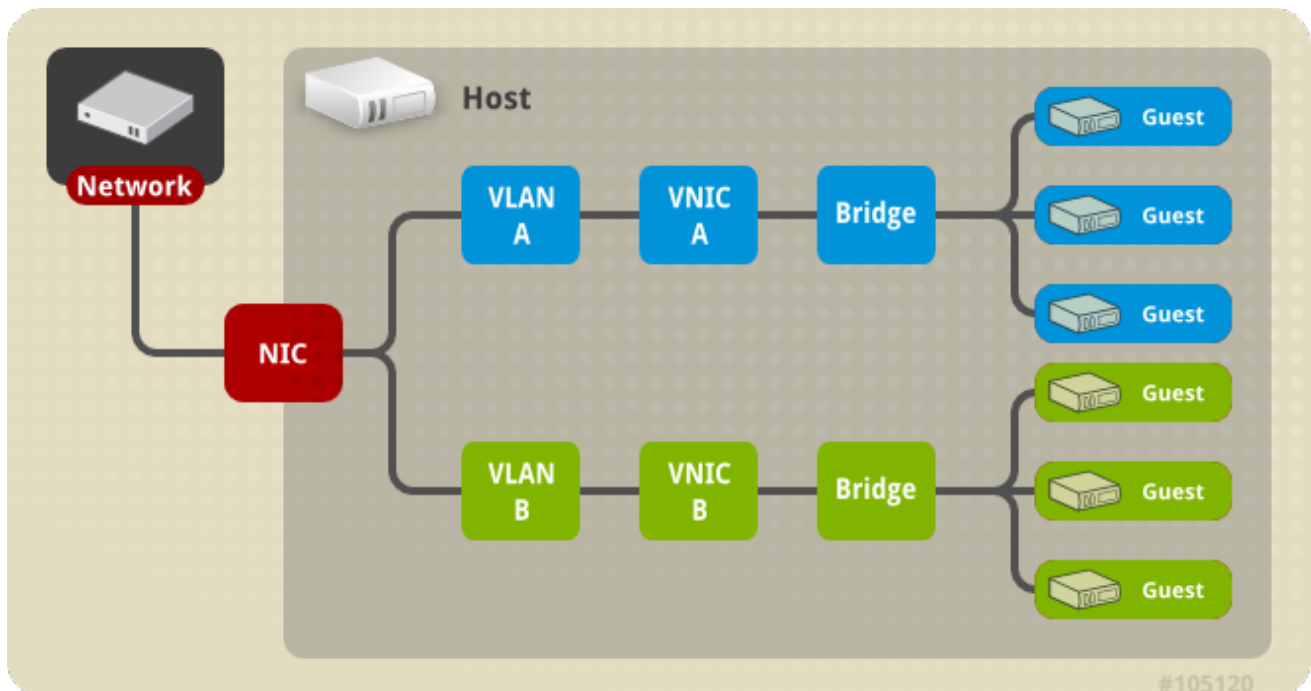


含まれるボンドは、2つ（またはそれ以上）の物理イーサネットリンクを組み合わせる論理リンクを作成します。結果として得られる利点には、ボンディングモードに応じて、NIC のフォールトトレランスと潜在的な帯域幅の拡張が含まれます。

### 3.19. 複数のブリッジ、複数 VLAN、および NIC の設定

図3.6「複数のブリッジ、複数 VLAN、および NIC の設定」は、1つの NIC を2つの VLAN に接続する設定を示しています。これは、2つの VLAN のいずれかにタグ付けされたネットワークトラフィックをホスト上の1つの NIC に渡すようにネットワークスイッチが設定されていることを前提としています。ホストは2つの VNIC を使用して、VLAN ごとに1つずつ VLAN トラフィックを分離します。次に、いずれかの VLAN にタグ付けされたトラフィックは、適切な VNIC をブリッジメンバーとして持つことで、別のブリッジに接続します。各ブリッジは、複数の仮想マシンによって接続されます。

図3.6 複数のブリッジ、複数 VLAN、および NIC の設定

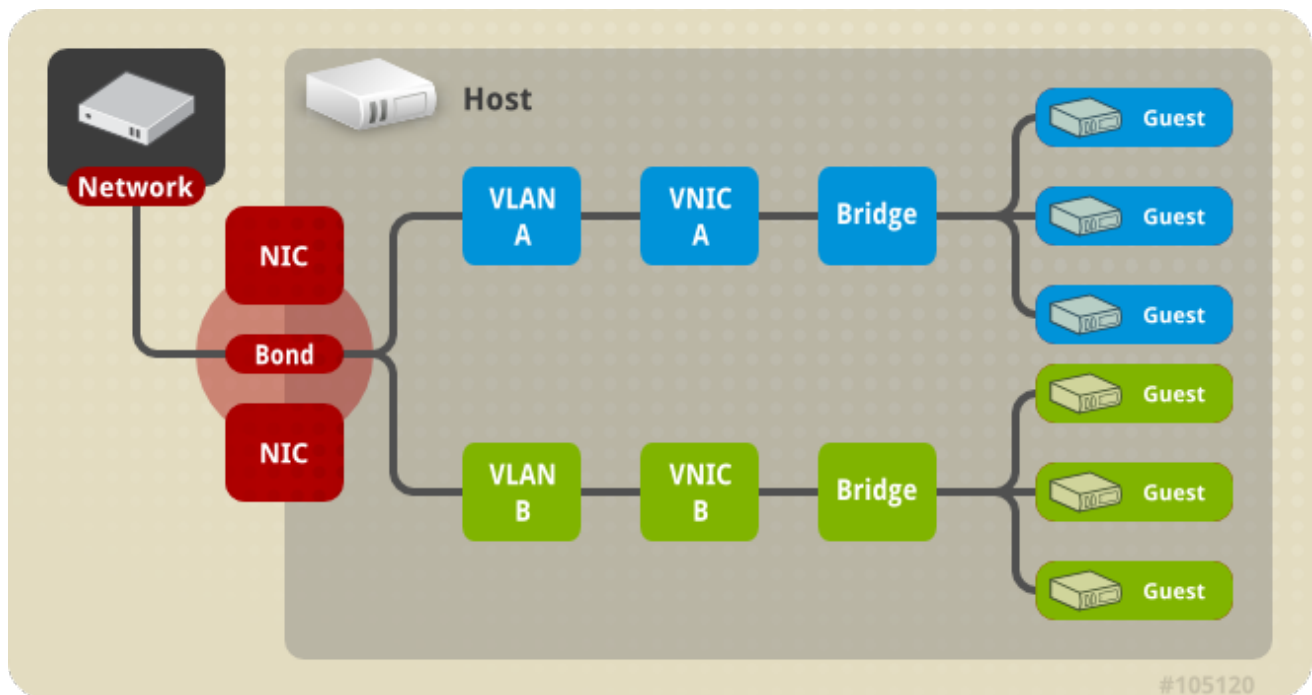


### 3.20. 複数のブリッジ、複数の VLAN、およびボンディングの設定

図3.7「ボンディング接続のある複数のブリッジ、複数 VLAN、および複数の NIC」複数の VLAN との接続を容易にするために、複数の NIC を結合する設定を表示します。



図3.7 ボンディング接続のある複数のブリッジ、複数 VLAN、および複数の NIC



この設定の各 VLAN は、NIC に接続するボンディングを介して定義されます。各 VLAN は個々のブリッジに接続し、各ブリッジは 1 つ以上のゲストに接続します。

## 第4章 電源管理

### 4.1. 電力管理とフェンシングの概要

Red Hat Virtualization 環境は、電力管理とフェンシングが設定されている場合に最も柔軟で回復力があります。電源管理により、Red Hat Virtualization Manager はホストのパワーサイクル操作を制御できます。最も重要なのは、問題が検出されたホストを再起動することです。フェンシングは、パフォーマンスの低下を防ぐために、問題のあるホストを再起動することにより、機能している Red Hat Virtualization 環境から分離するために使用されます。フェンスで囲まれたホストは、管理者の操作によって応答状態に戻り、その環境に再統合できます。

電源管理とフェンシングは、ホストのオペレーティングシステムとは独立してホストを再起動するために、特別な専用ハードウェアを利用します。Red Hat Virtualization Manager は、ネットワーク IP アドレスまたはホスト名を使用して電力管理デバイスに接続します。Red Hat Virtualization のコンテキストでは、電源管理デバイスとフェンシングデバイスは同じものです。

### 4.2. RED HAT VIRTUALIZATION の PROXY による電源管理

Red Hat Virtualization Manager は、フェンスエージェントと直接通信しません。その代わりに、Manager はプロキシを使用してホストの電源管理デバイスに電源管理コマンドを送信します。Manager は VDSM を使用してパワーマネージメントデバイスのアクションを実行するため、環境内の別のホストをフェンシングプロキシとして使用しています。

以下で選択することができます。

- フェンシングが必要なホストと同じクラスター内の任意のホスト。
- フェンシングが必要なホストと同じデータセンターにあるすべてのホスト。

実行可能なフェンシングプロキシホストのステータスは UP または Maintenance のいずれかです。

### 4.3. 電源管理

Red Hat Virtualization Manager は、稼働していないか、応答しない状態になったホストを再起動することができます。また使用率の低いホストの電源をオフにして電力を節約することができます。この機能は、適切に設定された電源管理デバイスによって異なります。Red Hat Virtualization 環境は、以下の電力管理デバイスをサポートします。

- *American Power Conversion (apc)*
- ありました。
- *Cisco Unified Computing System (cisco\_ucs)*
- *Dell Remote Access Card 5 (drac5)*
- *Dell Remote Access Card 7 (drac7)*
- 電子電源スイッチ (eps)
- *HPAppender System (hpblade)*
- *Integrated Lights Out (ilo,ilo2,ilo3,ilo4)*。
- *Intelligent Platform Management Interface (ipmilan)*
- *Remote Supervisor Adapter (rsa)*
- rsb.
- *Western Telematic, Inc (wti)*



#### 注記

APC 5.x 電源管理デバイスは、apc フェンスエージェントでは対応していません。代わりに apc\_snmp フェンスエージェントを使用してください。



一覧表示されている電源管理デバイスと通信するために、Red Hat Virtualization Manager は フェンスエージェントを使用します。Red Hat Virtualization Manager を使用すると、管理者は、デバイスが受け入れて応答するパラメーターを使用して、環境内の電力管理デバイスのフェンスエージェントを設定できます。基本的な設定オプションは、グラフィカルユーザーインターフェイスを使用して設定できます。特別な設定オプションも入力でき、解析されずにフェンスデバイスに渡されます。特別な設定オプションは特定のフェンスデバイスに固有ですが、基本的な設定オプションは、サポートされているすべての電源管理デバイスによって提供される機能用です。すべての電力管理デバイスによって提供される基本的な機能は次のとおりです。

- ステータス: ホストのステータスを確認します。
- 起動: ホストの電源をオンにします。
- 停止: ホストの電源を切ります。
- restart: ホストを再起動します。実際には、停止、待機、ステータス、開始、待機、ステータスとして実装されます。

ベストプラクティスは、電源管理設定を最初に設定するときに 1 回テストし、その後、機能が継続することを確認するために時々テストすることです。

復元力は、環境内のすべてのホストで適切に設定された電源管理デバイスによって提供されます。フェンシングエージェントを使用すると、Red Hat Virtualization Manager はホストの電源管理デバイスと通信して、問題のあるホストのオペレーティングシステムをバイパスし、ホストを再起動して他の環境からホストを分離できます。その後、Manager は、問題のあるホストによって保持されていた場合、SPM のロールを再割り当てし、他のホストで可用性の高い仮想マシンを安全に再起動できます。

#### 4.4. フェンシング

Red Hat Virtualization 環境のコンテキストでは、フェンシングは、Manager がフェンスエージェントを使用して開始し、電源管理デバイスによって実行されるホストの再起動です。フェンシングによ

り、クラスターは予期しないホスト障害に対応できるだけでなく、省電力、負荷分散、および仮想マシンの可用性ポリシーを適用できます。

フェンシングにより、Storage Pool Manager (SPM)のロールが常に機能ホストに割り当てられます。フェンスで囲まれたホストが SPM であった場合、SPM のロールは放棄され、応答可能なホストに再割り当てされます。SPM のロールを持つホストは、データドメイン構造のメタデータを書き込むことができる唯一のホストであるため、応答がなく、フェンスで囲まれていない SPM ホストにより、環境は仮想ディスクの作成と破棄、スナップショットの作成、論理ボリュームの拡張、およびデータドメイン構造メタデータへの変更を必要とする他のすべてのアクションを行うことができなくなります。

ホストが応答なくなると、そのホストで現在実行されているすべての仮想マシンも応答なくなる可能性があります。ただし、応答しないホストは、実行中の仮想マシンの仮想マシンハードディスクイメージのロックを保持します。2 番目のホストで仮想マシンを起動し、仮想マシンのハードディスクイメージに 2 番目のホストの書き込み権限を割り当てようとする、データが破損する可能性があります。

フェンシングにより、Red Hat Virtualization Manager は、仮想マシンのハードディスクイメージのロックが解除されたと見なすことができます。Manager は、フェンスエージェントを使用して、問題のあるホストが再起動されたことを確認できます。この確認を受け取ると、Red Hat Virtualization Manager は、データ破壊のリスクを冒すことなく、問題のあるホストから別のホスト上の仮想マシンを起動できます。フェンシングは、高可用性仮想マシンの基盤です。高可用性とマークされた仮想マシンは、データの破損を引き起こさないという確実性がなければ、代替ホストで安全に起動することはできません。

ホストが応答しなくなった場合、Red Hat Virtualization Manager は、アクションが実行される前に 30 秒の猶予期間が経過することを許可し、ホストが一時的なエラーから回復できるようにします。猶予期間が経過するまでにホストが応答しなくなった場合、Manager は応答しないホストからの悪影響を自動的に軽減し始めます。Manager は、ホストの電源管理カードのフェンシングエージェントを使用して、ホストを停止し、停止したことを確認し、ホストを開始し、ホストが開始されたことを確認します。ホストは起動を完了すると、フェンスで囲まれる前にその一部であったクラスターに再参加しようとしています。ホストが応答なくなる原因となっていた問題が解決された場合、ホストは自動的に Up ステータスに設定され、仮想マシンの起動とホストが再び可能となります。

#### 4.5. ソフトフェンシングホスト

ホストは予期せぬ問題で応答なくなることがありますが、VDSM は要求に応答できないものの、VDSM に依存している仮想マシンは生きており、アクセス可能です。このような場合は、VDSM を再起動することで VDSM が応答可能な状態に戻り、この問題が解決します。

"SSH Soft Fencing"とは、応答しないホストに対して Manager が SSH 経由で VDSM の再起動を試みるプロセスのことです。Manager が SSH 経由で VDSM の再起動に失敗した場合、外部フェンシングエージェントが設定されていれば、フェンシングの責任は外部フェンシングエージェントに移ります。

SSH でのソフトフェンシングは以下のように動作します。ホストでフェンシングを設定して有効にする必要があります。有効なプロキシーホスト (データセンター内の UP 状態の 2 番目のホスト) が存在する必要があります。Manager とホストの接続がタイムアウトすると、以下のようになります。

1. 最初のネットワーク障害では、ホストの状態が接続中に変わります。
2. その後、マネージャーは VDSM にステータスの問い合わせを 3 回試みるか、ホストの負荷に応じた間隔で待機します。間隔の長さを決定する式は、設定値 `TimeoutToResetVdsInSeconds` (デフォルトは 60 秒) + [`DelayResetPerVmInSeconds` (デフォルトは 0.5 秒)]\*(ホスト上で実行している仮想マシンの数)+ [`DelayResetForSpmlnSeconds` (デフォルトは 20 秒です)] \* 1 (ホストが SPM として実行している場合) または 0 (ホストが SPM として実行されていない場合)。VDSM に応答する最大時間を与えるために、Manager は上記の 2 つのオプションのうち長い方を選択します (VDSM のステータスまたは上記の式で決定された間隔を取得するための 3 回の試行)。
3. 間隔が経過したときにホストが応答しない場合は、`vdsml restart` を SSH 経由で実行します。
4. `vdsml restart` がホストと Manager 間の接続を再確立しない場合、ホストのステータスは `Non Responsive` に変更され、電源管理が設定されている場合、フェンシングは外部フェンシングエージェントに渡されます。



#### 注記

SSH を介したソフトフェンシングは、電源管理が設定されていないホストで実行できません。これはフェンシングとは異なります。フェンシングは、電源管理が設定されているホストでのみ実行できます。

#### 4.6. 複数の電力管理フェンシングエージェントの使用

シングルエージェントはプライマリーエージェントとして扱われます。セカンダリーエージェントは、フェンシングエージェントが 2 つある場合に有効です。たとえば、各電源スイッチに 2 つのエージェントが同じ電源スイッチに接続されているデュアル電源ホストの場合です。エージェントは、同じタイプでも異なるタイプでもかまいません。

ホストに複数のフェンシングエージェントを配置すると、フェンシング手順の信頼性が向上します。たとえば、ホスト上の唯一のフェンシングエージェントに障害が発生した場合、ホストは手動で再起動

されるまで非動作状態のままになります。以前にホストで実行していた仮想マシンは一時停止となり、元のホストが手動でフェンスされた後にのみ、クラスター内の別のホストにフェイルオーバーします。複数のエージェントがあり、最初のエージェントに障害が発生した場合は、次のエージェントを呼び出すことができます。

2つのフェンシングエージェントがホストで定義されている場合、それらは *同時* フローまたは *順次* フローを使用するように設定できます。

- **同時:** プライマリーエージェントとセカンダリーエージェントの両方が、ホストを停止するために **Stop** コマンドに応答する必要があります。1つのエージェントが開始コマンドに応答すると、ホストが起動します。
- **シーケンシャル:** ホストの停止または起動には、プライマリーエージェントが最初に使用され、失敗した場合はセカンダリーエージェントが使用されます。

## 第5章 負荷分散、スケジューリング、および移行

### 5.1. 負荷分散、スケジューリング、および移行

個々のホストには有限のハードウェアリソースがあり、障害が発生しやすくなっています。障害とリソースの枯渇を軽減するために、ホストはクラスターにグループ化されます。クラスターは、基本的に共有リソースのグループです。Red Hat Virtualization 環境は、負荷分散ポリシー、スケジューリング、および移行を使用して、ホストリソースの需要の変化に対応します。Manager は、クラスター内の単一のホストがそのクラスター内のすべての仮想マシンを担当しないようにすることができます。逆に、Manager は十分に活用されていないホストを認識し、そこからすべての仮想マシンを移行できるため、管理者はそのホストをシャットダウンして電力を節約できます。

次の3つのイベントの結果として、使用可能なリソースがチェックされます。

- 仮想マシンの起動 - リソースがチェックされ、仮想マシンが起動するホストが決定されません。
- 仮想マシンの移行 - 適切なターゲットホストを決定するために、リソースがチェックされます。
- 時間の経過 - リソースは定期的にチェックされ、個々のホストの負荷がクラスターの負荷分散ポリシーに準拠しているかどうか判断されます。

Manager は、クラスターの負荷分散ポリシーを使用して、クラスター内の1つのホストから別のホストへの仮想マシンの移行をスケジュールすることにより、使用可能なリソースの変更に応答します。次のセクションでは、負荷分散ポリシー、スケジューリング、および仮想マシンの移行の関係について説明します。

### 5.2. 負荷分散ポリシー

負荷分散ポリシーはクラスターに設定されます。クラスターには、それぞれ異なるハードウェアパラメーターと使用可能なメモリを持つ1つ以上のホストが含まれます。Red Hat Virtualization Manager は、負荷分散ポリシーを使用して、クラスター内のどのホストで仮想マシンを起動するかを決定します。負荷分散ポリシーにより、Manager は、仮想マシンを使用率の高いホストから使用率の低いホストにいつ移動するかを決定することもできます。



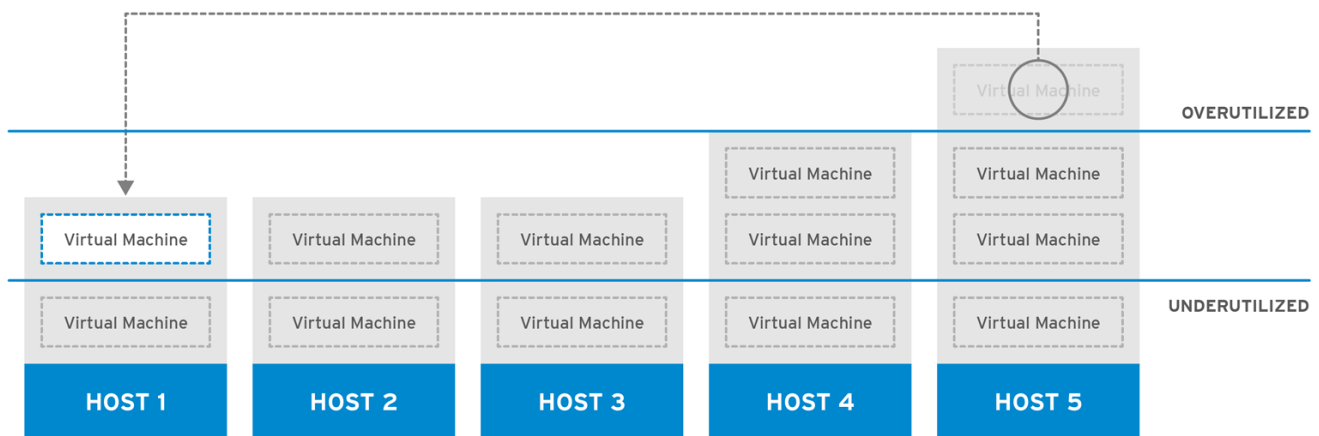
負荷分散プロセスは、データセンター内のクラスターごとに1分ごとに実行されます。これは、どのホストが過剰に使用されているか、どのホストが十分に使用されていないか、および仮想マシンの移行の有効なターゲットであるかを判別します。決定は、特定のクラスターの管理者によって設定された負荷分散ポリシーに基づいて行われます。負荷分散ポリシーのオプションは、VM\_Evenly\_Distributed、Evenly\_Distributed、Power\_Saving、およびNoneです。

### 5.3. 負荷分散ポリシー: VM\_EVENLY\_DISTRIBUTED

仮想マシンの均等分散ロードバランシングポリシーは、仮想マシンの数に基づいて、仮想マシンをホスト間で均等に分散します。高い仮想マシン数は、各ホストで実行できる仮想マシンの最大数であり、それを超えると、ホストの過負荷と見なされます。VM\_Evenly\_Distributed ポリシーを使用すると、管理者はホストに高い仮想マシン数を設定できます。最も使用率の高いホストと最も使用率の低いホスト間の仮想マシン数の最大包括的差異も、管理者によって設定されます。クラスターのすべてのホストが移行しきい値内に留まる仮想マシン数がある場合、クラスターが分散されます。管理者は、SPM ホストで予約する仮想マシンのスロット数も設定します。SPM ホストの負荷は他のホストよりも低いため、この変数は、実行できる仮想マシンの数を他のホストよりも少なく定義します。いずれかのホストが高い仮想マシン数よりも多くの仮想マシンを実行していて、少なくとも1つのホストの仮想マシン数が移行しきい値の範囲外である場合、仮想マシンはCPUが最も少ないクラスター内のホストに1つずつ移行されます。クラスター内のすべてのホストの仮想マシン数が移行しきい値内に収まるまで、一度に1台の仮想マシンが移行されます。

### 5.4. 負荷分散ポリシー: EVENLY\_DISTRIBUTED

図5.1 Evenly Distributed スケジューリングポリシー



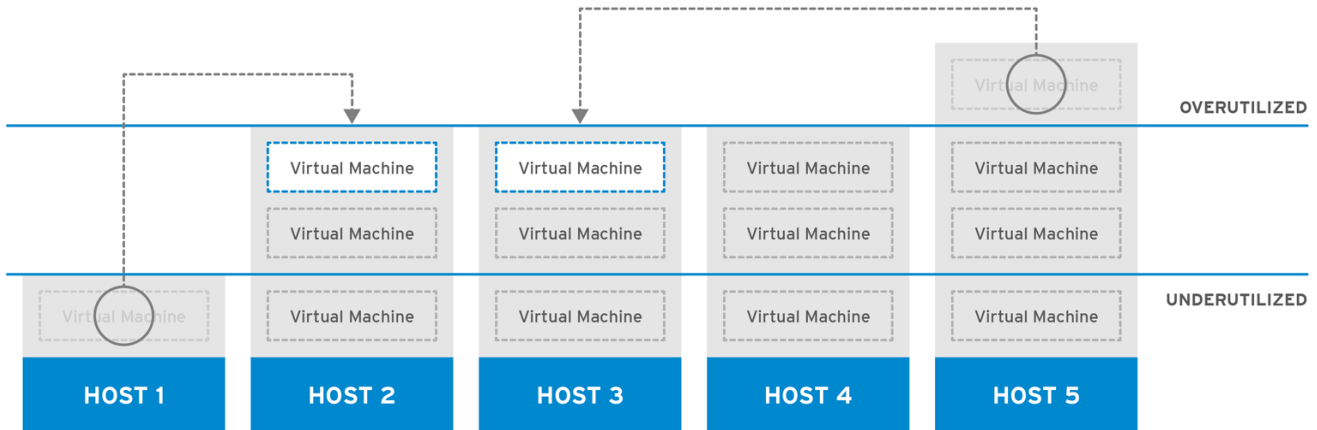
RHV\_444396\_0417

均等に分散された負荷分散ポリシーは、最小のCPU負荷または最大の使用可能なメモリに従って新しい仮想マシンのホストを選択します。クラスター内のホストに設定された時間許可される最大CPU負荷および最小使用可能メモリは、均等に分散されたスケジューリングポリシーのパラメーターによって定義されます。この制限を超えると、環境のパフォーマンスが低下します。均等に分散されたポリシーにより、管理者は仮想マシンを実行するためにこれらのレベルを設定できます。ホストが定義された最大CPU負荷または最小使用可能メモリに達し、ホストが設定された時間より長くそこにとどまる場合、そのホスト上の仮想マシンは、どのパラメーターが利用されているかによって、クラスター内の最小CPUまたは最大利用メモリを持つホストに1つずつ移行されます。ホストリソース

は 1 分に 1 回チェックされ、ホストの CPU 負荷が定義された制限を下回るか、ホストの使用可能なメモリーが定義された制限を超えるまで、一度に 1 つの仮想マシンが移行されます。

## 5.5. 負荷分散ポリシー: POWER\_SAVING

図5.2 Power Saving スケジューリングポリシー



RHV\_444396\_0417

省電力の負荷分散ポリシーは、最小の CPU または最大の使用可能なメモリーに従って、新しい仮想マシンのホストを選択します。クラスター内のホストに設定された時間許可される最大 CPU 負荷と最小使用可能メモリーは、省電力スケジューリングポリシーのパラメーターによって定義されます。この制限を超えると、環境のパフォーマンスが低下します。省電力パラメーターは、ホストの継続的な動作が非効率的な電力使用と見なされる前に、クラスター内のホストに設定された時間だけ許可される最小 CPU 負荷と最大使用可能メモリーも定義します。ホストが最大 CPU 負荷または最小使用可能メモリーに達し、設定された時間より長くそこにとどまる場合、そのホスト上の仮想マシンは、どのパラメーターが利用されているかによって、最小 CPU または最大利用メモリーを持つホストに 1 つずつ移行されます。ホストリソースは 1 分に 1 回チェックされ、ホストの CPU 負荷が定義された制限を下回るか、ホストの使用可能なメモリーが定義された制限を超えるまで、一度に 1 つの仮想マシンが移行されます。ホストの CPU 負荷が定義された最小レベルを下回るか、ホストの使用可能なメモリーが定義された最大レベルを超える場合、クラスター内の他のホストの仮想マシンが最大 CPU 負荷を超え、最小使用可能メモリーを下回る限り、そのホスト上の仮想マシンはクラスター内の他のホストに移行されます。十分に活用されていないホストから残りの仮想マシンが削除されると、Manager は自動的にホストマシンの電源を切り、負荷分散が必要な場合、またはクラスター内に十分な空きホストがない場合にホストを再起動します。

## 5.6. 負荷分散ポリシー: なし

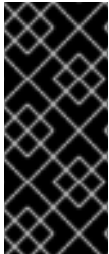
負荷分散ポリシーが選択されていない場合、仮想マシンは、CPU 使用率と使用可能なメモリーが最も少ないクラスター内のホストで起動します。CPU 使用率を決定するために、仮想 CPU カウントと CPU 使用率を考慮した複合メトリックが使用されます。ホストの選択ポイントは新しい仮想マシンの起動時のみであるため、このアプローチは最も動的ではありません。ホストに対する需要の増加を反映するために、仮想マシンは自動的に移行されません。

管理者は、どのホストが特定の仮想マシンの適切な移行ターゲットであるかを決定する必要があります。ピンングを使用して、仮想マシンを特定のホストに関連付けることもできます。固定すると、仮想

マシンが他のホストに自動的に移行されなくなります。リソースが大量に消費される環境では、手動移行が最善のアプローチです。

## 5.7. 負荷分散ポリシー：INCLUSTERUPGRADE

InClusterUpgrade スケジューリングポリシーは、ホストオペレーティングシステムのバージョンに基づいて仮想マシンを配布します。現在実行している仮想マシンよりも新しいオペレーティングシステムを持つホストは、同じオペレーティングシステムを持つホストよりも優先されます。新しいオペレーティングシステムを持つホストに移行する仮想マシンは、古いオペレーティングシステムに移行されません。仮想マシンは、クラスター内の任意のホストで再起動できます。このポリシーにより、クラスターでオペレーティングシステムのバージョンが混在するようにすることで、クラスター内のホストをアップグレードできます。ポリシーを有効にする前に、前提条件を満たす必要があります。Red Hat Enterprise 『Virtualization 3.6 Upgrade Guide の [Upgrading Hosts in a Cluster from Red Hat Enterprise Linux 6 to Red Hat Enterprise Linux 7](#)』を参照してください。



### 重要

InClusterUpgrade スケジューリングポリシーは、メジャーバージョン間のアップグレードのみに使用されます。たとえば、Red Hat Enterprise Linux 6 から Red Hat Enterprise Linux 7 へのアップグレードは以下ようになります。

## 5.8. 高可用性仮想マシンの予約

高可用性 (HA) 仮想マシン予約ポリシーにより、Red Hat Virtualization Manager は高可用性仮想マシンのクラスター容量を監視できます。Manager には、個々の仮想マシンに高可用性のフラグを立てる機能があります。つまり、ホストに障害が発生した場合、これらの仮想マシンは代替ホストで再起動されます。このポリシーは、クラスター内のホスト間で高可用性の仮想マシンのバランスを取ります。クラスター内のいずれかのホストに障害が発生した場合、残りのホストは、クラスターのパフォーマンスに影響を与えることなく、可用性の高い仮想マシンの移行負荷をサポートできます。高可用性仮想マシンの予約が有効になっている場合、Manager は、既存のホストに予期しない障害が発生した場合に HA 仮想マシンが移行するための適切な容量がクラスター内に存在することを確認します。

## 5.9. スケジューリング

Red Hat Virtualization では、スケジューリングとは、Red Hat Virtualization Manager がクラスター内のホストを新規または移行された仮想マシンのターゲットとして選択する方法を指します。

ホストが仮想マシンを起動したり、別のホストから移行された仮想マシンを受け入れたりする資格を得るには、ホストで起動または移行される仮想マシンの要件をサポートするのに十分な空きメモリと CPU が必要です。仮想マシンは、オーバーロードされた CPU を持つホストで起動しません。デフォルトでは、ホストの CPU が 5 分間 80% を超える負荷がかかった場合に過負荷であると見なされますが、この値はスケジューリングポリシーを使用して変更できます。複数のホストが適格なターゲットである場合は、クラスターの負荷分散ポリシーに基づいて 1 つが選択されます。たとえば、Evenly\_Distributed ポリシーが有効な場合、Manager は CPU 使用率が最も低いホストを選択しま

す。Power\_Saving ポリシーが有効な場合は、最大サービスレベルと最小サービスレベルの間で CPU 使用率が最も低いホストが選択されます。特定のホストの Storage Pool Manager (SPM)ステータスも、仮想マシンまたは仮想マシンの移行を開始するターゲットとしての適格性に影響します。非 SPM ホストが優先ターゲットホストです。たとえば、クラスターで開始された最初の仮想マシンは、SPM のロールがそのクラスター内のホストによって保持されている場合、SPM ホスト上で実行されません。

## 5.10. マイグレーション

Red Hat Virtualization Manager は、移行を使用してクラスターの負荷分散ポリシーを適用します。仮想マシンの移行は、クラスターの負荷分散ポリシーとクラスター内のホストに対する現在の要求に従って行われます。移行は、ホストがフェンスされているとき、またはメンテナンスモードに移行したときに自動的に発生するように設定することもできます。Red Hat Virtualization Manager は、最初に CPU 使用率が最も低い仮想マシンを移行します。これはパーセンテージとして計算され、I/O 操作が CPU 使用率に影響を与える場合を除いて、RAM 使用量または I/O 操作は考慮されません。同じ CPU 使用率の仮想マシンが複数ある場合、最初に移行されるのは、仮想マシンの CPU 使用率を決定するために Red Hat Virtualization Manager によって実行されるデータベースクエリーによって返される最初の仮想マシンです。

仮想マシンの移行には、デフォルトで次の制限があります。

- 各仮想マシンの移行には、52 MiBps (メガバイト/秒) の帯域幅制限が課せられます。
- 移行は、仮想マシンメモリーの GB あたり 64 秒後にタイムアウトになります。
- 進行が 240 秒間停止すると、移行は中止されます。
- 同時送信移行は、ホストごとの CPU コアごとに 1 つ、または 2 つのうち小さい方に制限されます。

移行設定のチューニングに関する詳細は、を参照 <https://access.redhat.com/solutions/744423> してください。

## 第6章 ディレクトリーサービス

### 6.1. ディレクトリーサービス

Red Hat Virtualization プラットフォームは、ユーザーの認証と承認をディレクトリーサービスに依存しています。ユーザーポータル、管理ポータル、REST API を含むすべての Manager インターフェイスとの対話は、認証され、承認されたユーザーに制限されます。Red Hat Virtualization 環境内の仮想マシンは、同じディレクトリーサービスを使用して認証と承認を提供できますが、そうするように設定する必要があります。Red Hat Virtualization Manager で使用するディレクトリーサービスの現在サポートされているプロバイダーは、*Identity Management (IdM)*、*Red Hat Directory Server 9 (RHDS)*、*Active Directory (AD)*、および *OpenLDAP* です。Red Hat Virtualization Manager は、以下のディレクトリーサーバーとインターフェイスします。

- ポータルログイン (ユーザー、パワーユーザー、管理者、REST API)。
- ユーザー情報を表示するためのクエリー。
- Manager をドメインに追加します。

認証とは、一部のデータを生成した当事者の検証と識別、および生成されたデータの整合性の検証です。プリンシパルは、身元が確認された当事者です。検証者は、プリンシパルのアイデンティティーの保証を要求する当事者です。Red Hat Virtualization の場合は、Manager がベリファイアであり、ユーザーがプリンシパルです。データの整合性とは、受信したデータがプリンシパルによって生成されたデータと同じであることを保証することです。

機密性と承認は認証と密接に関連しています。機密保持は、データを受け取することを意図していない人への開示からデータを保護します。強力な認証方法は、任意で機密性を提供できます。承認は、プリンシパルが操作を実行できるかどうかを決定します。Red Hat Virtualization は、ディレクトリーサービスを使用してユーザーをロールに関連付け、それに応じて認証を提供します。承認は通常、プリンシパルが認証された後に実行され、ベリファイアのローカルまたはリモートの情報に基づく場合があります。

インストール中に、ローカルの内部ドメインが Red Hat Virtualization 環境の管理用に自動的に設定されます。インストールが完了したら、さらにドメインを追加できます。

### 6.2. ローカル認証: 内部ドメイン

Red Hat Virtualization Manager は、インストール中に限定された内部管理ドメインを作成します。このドメインは、ディレクトリーサーバー上のディレクトリーサービスユーザーとしてではなく、Red Hat Virtualization PostgreSQL データベースのキーに基づいて存在するため、AD または IdM ドメインと同じではありません。内部ドメインには `admin@internal` ユーザーという 1 人のユーザーしかいないため、内部ドメインも外部ドメインとは異なります。このアプローチを初期認証に採用することで、完全に機能的なディレクトリーサーバーを必要とせずに Red Hat Virtualization を評価でき、外部ディレクトリーサービスの問題をトラブルシューティングするための管理者アカウントを利用できるようになります。

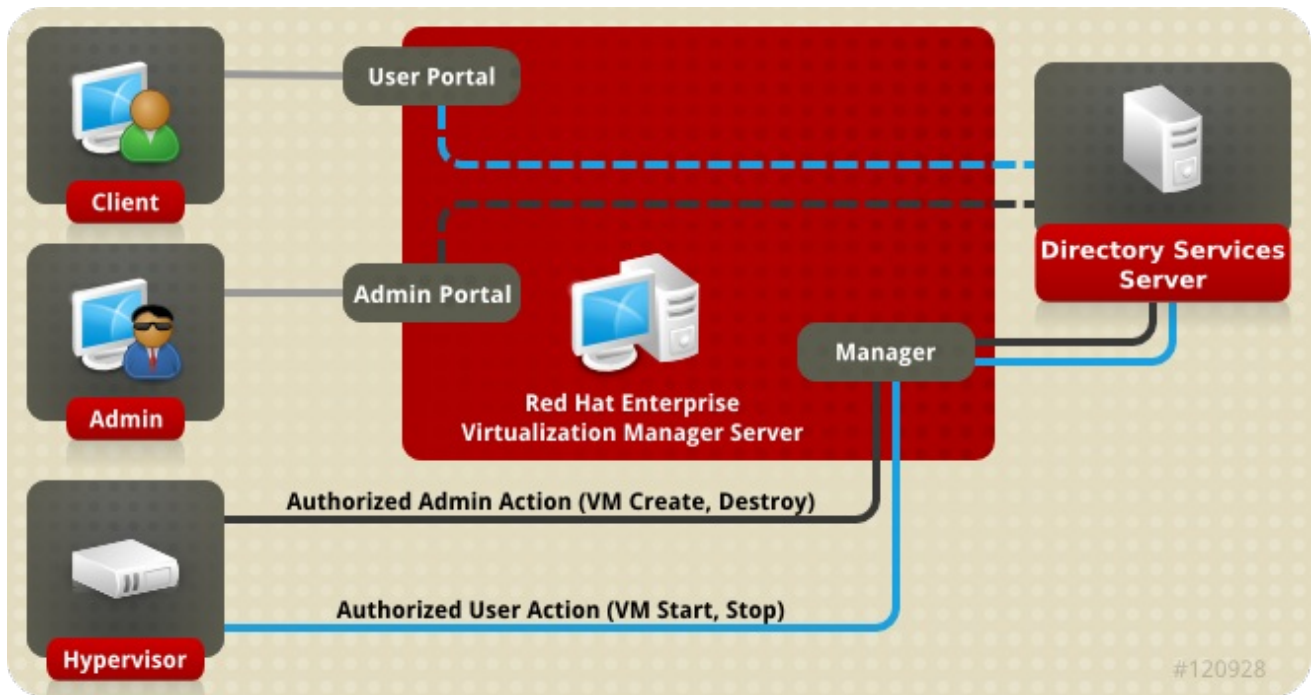
`admin@internal` ユーザーは、環境の初期設定用です。これには、ホストのインストールと受け入れ、外部 AD または IdM 認証ドメインの追加、外部ドメインからのユーザーへのアクセス許可の委任が含まれます。

### 6.3. GSSAPI を使用したリモート認証

Red Hat Virtualization のコンテキストでは、リモート認証は Red Hat Virtualization Manager からリモートで処理される認証を指します。リモート認証は、AD、IdM、または RHDS ドメイン内から Manager に到達するユーザーまたは API 接続に使用されます。Red Hat Virtualization Manager は、`engine-manage-domains` ツールを使用して管理者が RHDS、AD、または IdM ドメインの一部として設定する必要があります。これには、システムをドメインに参加させるのに十分な特権を持つ、ドメインの RHDS、AD、または IdM ディレクトリーサーバーからのアカウントの認証情報をマネージャーに提供する必要があります。ドメインが追加された後、ドメインユーザーは、パスワードを使用してディレクトリーサーバーに対して Red Hat Virtualization Manager によって認証されます。Manager は *Simple Authentication and Security Layer (SASL)* と呼ばれるフレームワークを使用し、次に *Generic Security Services Application Program Interface (GSSAPI)* を使用してユーザーの ID を安全に検証し、ユーザーが使用できる承認レベルを確認します。

#### 図6.1 GSSAPI 認証





## 第7章 テンプレートおよびプール

### 7.1. テンプレートおよびプール

Red Hat Virtualization 環境は、ユーザーへの仮想マシンのプロビジョニングを簡素化するツールを管理者に提供します。これらは テンプレート および プール です。テンプレートは、管理者がオペレーティングシステムのインストールと設定をバイパスして、既存の事前設定された仮想マシンに基づいて新しい仮想マシンをすばやく作成できるようにするショートカットです。これは、アプライアンスのように使用される仮想マシン、たとえば Web サーバー仮想マシンに特に役立ちます。組織が特定の Web サーバーの多くのインスタンスを使用する場合、管理者は、テンプレートとして使用される仮想マシンを作成し、オペレーティングシステム、Web サーバー、サポートパッケージをインストールし、固有の設定変更を適用できます。管理者は、必要に応じて新しい同一の仮想マシンを作成するために使用される、動作中の仮想マシンに基づいてテンプレートを作成できます。

仮想マシンプールは、ユーザーに迅速にプロビジョニングできる特定のテンプレートに基づく仮想マシンのグループです。プール内で仮想マシンを使用するためのアクセス許可は、プールレベルで付与されます。プールを使用する権限が付与されたユーザーには、プールから任意の仮想マシンが割り当てられます。仮想マシンプールに内在するのは、その中の仮想マシンの一時的な性質です。ユーザーには、過去にプール内のどの仮想マシンを使用したかに関わらず仮想マシンが割り当てられるため、プールはデータの永続性を必要とする目的には適していません。仮想マシンプールは、ユーザーデータが中央の場所に保存され、仮想マシンがそのデータにアクセスして使用する手段である場合、またはデータの永続性が重要でない場合に最適です。プールを作成すると、プールに配置される仮想マシンが停止状態で作成されます。これらは、ユーザーの要求に応じて開始されます。

### 7.2. テンプレート

テンプレートを作成するには、管理者が仮想マシンを作成してカスタマイズします。必要なパッケージがインストールされ、カスタマイズされた設定が適用され、デプロイ後に仮想マシンに加える必要のある変更を最小限に抑えるために、仮想マシンがその意図された目的のために準備されます。仮想マシンからテンプレートを作成する前の任意の推奨される手順は 一般化 です。一般化は、デプロイ時に変更されるシステムユーザー名、パスワード、タイムゾーン情報などの詳細を削除するために使用されます。一般化は、カスタマイズされた設定には影響しません。Red Hat Virtualization 環境の Windows および Linux ゲストの概要については、『Virtual Machine Management Guide』の [Templates](#) で詳しく説明されています。Red Hat Enterprise Linux ゲストは、`sys-unconfig` を使用して一般化されています。Windows ゲストは、`sys-prep` を使用して一般化されます。

テンプレートの基礎を提供する仮想マシンが十分に設定され、必要に応じて一般化され、停止されると、管理者は仮想マシンからテンプレートを作成できます。仮想マシンからテンプレートを作成すると、特別に設定された仮想ディスクイメージの読み取り専用コピーが作成されます。読み取り専用イメージは、そのテンプレートをベースとする後に作成されるすべての仮想マシンのバックアップイメージを形成します。つまり、テンプレートは基本的に、仮想ハードウェア設定が関連付けられたカスタマイズされた読み取り専用ディスクイメージです。ハードウェアは、テンプレートから作成された仮想マシンで変更できます。たとえば、1 ギガバイトの RAM を持つテンプレートから作成された仮想マシンに 2 ギガバイトの RAM をプロビジョニングする場合などです。ただし、テンプレートディスクイメージは変更できません。変更すると、テンプレートに基づいてすべての仮想マシンが変更されるためです。



テンプレートが作成されると、複数の仮想マシンのベースとして使用できます。仮想マシンは、シンプロビジョニング方式または クローン プロビジョニング方法を使用して、指定のテンプレートから作成されます。テンプレートからクローン作成される仮想マシンは、テンプレートベースイメージの完全な書き込み可能なコピーを取得し、テンプレートの存在に依存しないように、シン作成方法のスペース節約を犠牲にします。シン方法を使用してテンプレートから作成された仮想マシンは、テンプレートからの読み取り専用イメージをベースイメージとして使用するため、テンプレートと作成されたすべての仮想マシンを同じストレージドメインに保存する必要があります。データへの変更と新しく生成されたデータは、書き込みイメージのコピーに保存されます。テンプレートをベースとする各仮想マシンは、同じベース読み取り専用イメージと、仮想マシンに固有の書き込みイメージのコピーを使用します。これにより、同一のデータがストレージに保持される回数が制限されるため、ストレージを節約できます。さらに、読み取り専用のバックアップイメージを頻繁に使用すると、アクセスされるデータがキャッシュされ、ネットパフォーマンスが向上します。

### 7.3. POOLS

仮想マシンプールを使用すると、デスクトップとしてユーザーに多数の同一の仮想マシンを迅速にプロビジョニングできます。プールから仮想マシンにアクセスして使用する権限を付与されたユーザーは、要求のキュー内の位置に基づいて、使用可能な仮想マシンを受け取ります。プール内の仮想マシンは、データの永続性を許可しません。仮想マシンがプールから割り当てられるたびに、仮想マシンは基本状態で割り当てられます。これは、ユーザーデータが一元的に保存されている状況での使用に最適です。

仮想マシンプールはテンプレートから作成されます。プール内の各仮想マシンは、同じバックアップ読み取り専用イメージを使用し、書き込みイメージの一時的なコピーを使用して変更済みおよび新たに生成されたデータを保持します。プール内の仮想マシンは、ユーザーが生成して変更したデータを保持する書き込み層のコピーがシャットダウン時に失われるという点で、他の仮想マシンとは異なります。これは、仮想マシンプールが、それをサポートするテンプレートよりも多くのストレージと、使用中に生成または変更されたデータ用の領域を必要としないことを意味します。仮想マシンプールは、各ユーザーに専用の仮想デスクトップを提供するためのストレージコストをかけずに、一部のタスクでユーザーにコンピューティングパワーを提供する効率的な方法です。

#### 例7.1 プールの使用例

テクニカルサポート会社は 10 人のヘルプデスクスタッフを雇用しています。ただし、常に 5 つだけが機能しています。ヘルプデスクの従業員ごとに 1 つずつ、合計 10 台の仮想マシンを作成する代わりに、5 台の仮想マシンのプールを作成できます。ヘルプデスクの従業員は、シフトの開始時に仮想マシンを割り当て、終了時にそれをプールに戻します。

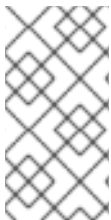
## 第8章 仮想マシンのスナップショット

### 8.1. スナップショット

スナップショットは、管理者が特定の時点で仮想マシンのオペレーティングシステム、アプリケーション、およびデータの復元ポイントを作成できるようにするストレージ機能です。スナップショットは、仮想マシンのハードディスクイメージに現在存在するデータを COW ボリュームとして保存し、スナップショットの作成時に存在していたデータへの復元を可能にします。スナップショットにより、現在のレイヤー上に新しい COW レイヤーが作成されます。スナップショットが作成された後に実行されるすべての書き込みアクションは、新しい COW レイヤーに書き込まれます。

仮想マシンのハードディスクイメージは 1 つ以上のボリュームのチェーンであることを理解することが重要です。仮想マシンの観点からは、これらのボリュームは単一のディスクイメージとして表示されます。仮想マシンは、そのディスクが複数のボリュームで設定されているという事実気づいていません。

COW ボリュームと COW レイヤーという用語は同じ意味で使用されますが、レイヤーはスナップショットの時間的性質をより明確に認識します。各スナップショットが作成され、管理者はスナップショットの作成後にデータに加えられた不適切な変更を破棄できます。スナップショットは、多くのワードプロセッサに存在する Undo 関数と同様の機能を提供します。



#### 注記

共有可能とマークされた仮想マシンのハードディスクのスナップショット、および直接 LUN 接続に基づくスナップショットは、ライブなどはサポートされていません。

3 つの主要なスナップショット操作は次のとおりです。

- 作成。仮想マシン用に作成された最初のスナップショットが含まれます。
- プレビュー。スナップショットをプレビューして、スナップショットが作成された時点でシステムデータを復元するかどうかを決定します。
- 削除。これには、不要になった復元ポイントの削除が含まれます。

スナップショット操作に関するタスクベースの情報は、『Red Hat Virtualization Virtual Machine Management Guide』の [Snapshots](#) を参照してください。

## 8.2. RED HAT VIRTUALIZATION のライブスナップショット

共有可能とマークされた仮想マシンのハードディスクのスナップショット、および直接 LUN 接続に基づくスナップショットは、ライブなどはサポートされていません。

複製または移行されていない他の仮想マシンは、実行中、一時停止中、または停止時にスナップショットを取得できます。

仮想マシンのライブスナップショットが開始すると、マネージャーは、SPM ホストが仮想マシンが使用する新しいボリュームを作成するように要求します。新しいボリュームの準備ができると、Manager は VDSM を使用して、仮想マシンを実行しているホスト上の libvirt および qemu と通信し、仮想マシンの書き込み操作に新しいボリュームの使用を開始する必要があります。仮想マシンが新しいボリュームに書き込める場合、スナップショット操作は成功したと見なされ、仮想マシンは前のボリュームへの書き込みを停止します。仮想マシンが新しいボリュームに書き込めない場合、スナップショット操作は失敗と見なされ、新しいボリュームが削除されます。

仮想マシンは、ライブスナップショットが開始してから新しいボリュームの準備が整うまで、現在のボリュームと新しいボリュームの両方にアクセスする必要があるため、両方のボリュームが読み取り/書き込みアクセスで開かれます。

静止をサポートするゲストエージェントがインストールされている仮想マシンは、スナップショット全体でファイルシステムの一貫性を確保できます。登録済みの Red Hat Enterprise Linux ゲストは `qemu-guest-agent` をインストールして、スナップショットの前に休止を有効にできます。

スナップショットの作成時に静止互換のゲストエージェントが仮想マシンに存在する場合、VDSM は libvirt を使用してエージェントと通信し、スナップショットの準備をします。未処理の書き込みアクションが完了すると、スナップショットが作成される前にファイルシステムがフリーズになります。スナップショットが完了し、libvirt が仮想マシンをディスク書き込みアクション用の新しいボリュームに切り替えると、ファイルシステムが解凍され、ディスクへの書き込みが再開されます。

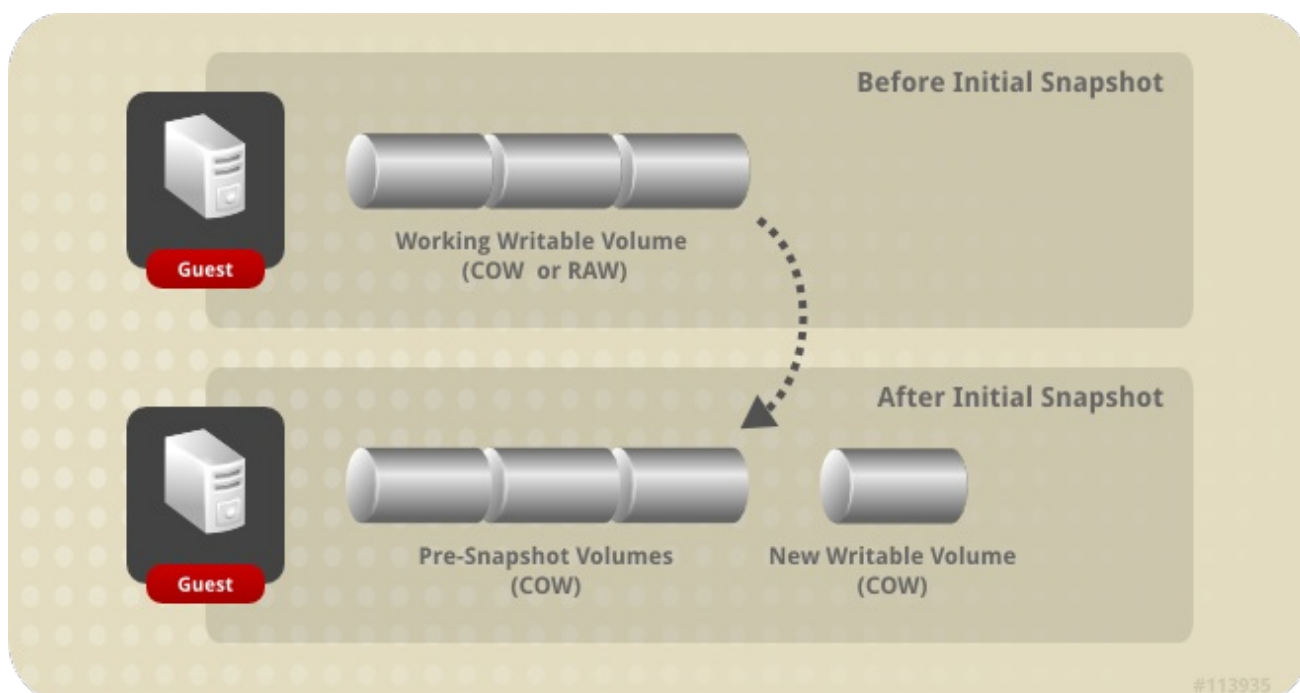
静止を有効にして試行されたすべてのライブスナップショット。互換性のあるゲストエージェントが存在しないために `snapshot` コマンドが失敗すると、ライブスナップショットは `use-quiescing` フラグなしで再度開始します。仮想マシンが静止ファイルシステムでスナップショット前の状態に戻ると、ファイルシステムチェックを必要とせずに正常に起動します。静止していないファイルシステムを使用して以前のスナップショットを元に戻すには、起動時にファイルシステムをチェックする必要があります。

## 8.3. スナップショットの作成

Red Hat Virtualization では、仮想マシンの初期スナップショットは、初期スナップショットが QCOW2 または RAW のいずれかの形式を保持するという点で後続のスナップショットとは異なります。仮想マシンの最初のスナップショットは、既存のボリュームをベースイメージとして指定します。追加のスナップショットは、前のスナップショット以降にイメージに保存されたデータに加えられた変更を追跡する追加の COW レイヤーです。

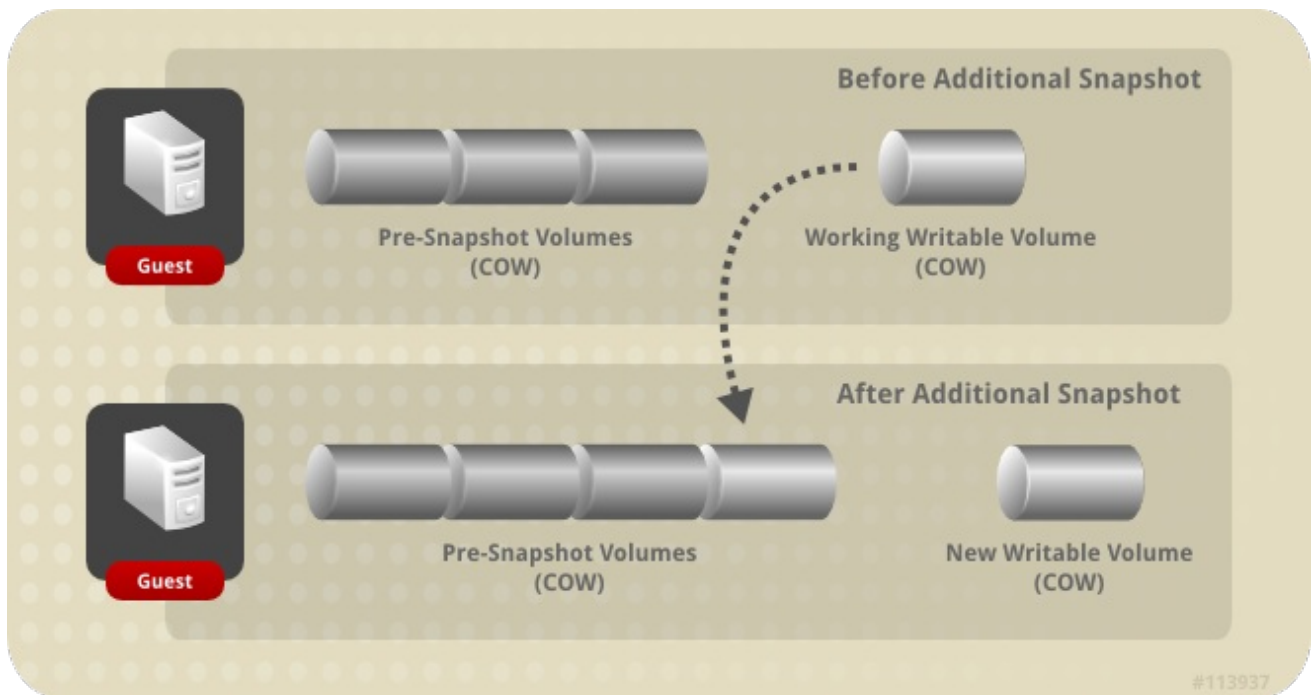
Red Hat Virtualization では、イメージがシンプロビジョニングされたイメージとして作成されるか、またはユーザーが QCOW2 に具体的に要求されない限り、ゲスト仮想マシンは通常 RAW ディスクイメージと対話します。図8.1「最初のスナップショットの作成」にあるように、スナップショットの作成により、仮想ディスクイメージを設定するボリュームが後続のすべてのスナップショットのベースイメージとして機能します。

図8.1 最初のスナップショットの作成



最初のスナップショットの後に作成されたスナップショットは、スナップショットの作成後に作成または変更されたデータが保存される新しい COW ボリュームの作成につながります。新しい COW レイヤーにはそれぞれ、COW メタデータのみが含まれます。スナップショットが新しい COW レイヤーに書き込まれた後に、仮想マシン経由で作成され、操作されるデータ。仮想マシンを使用して前の COW レイヤーに存在するデータを変更すると、データは前のレイヤーから読み取られ、最新のレイヤーに書き込まれます。仮想マシンは、仮想マシンに対して透過的に、最新から最古までの各 COW レイヤーをチェックすることによってデータを検索します。

図8.2 追加のスナップショットの作成



#### 8.4. スナップショットプレビュー

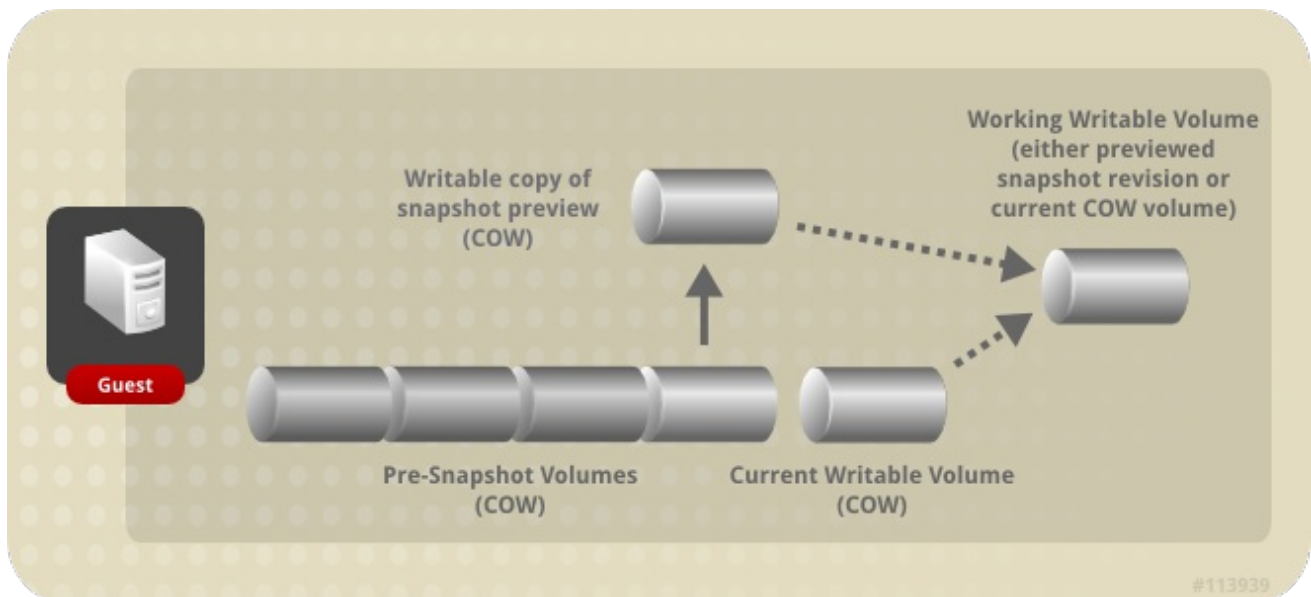
仮想ディスクイメージを元に戻すスナップショットを選択するには、管理者が以前に作成されたすべてのスナップショットをプレビューできます。

管理者は、ゲストごとに使用可能なスナップショットから、スナップショットボリュームを選択してその内容をプレビューできます。図8.3「スナップショットのプレビュー」にあるように、各スナップショットは COW ボリュームとして保存され、プレビューされると、プレビューされるスナップショットから新しいプレビューレイヤーがコピーされます。ゲストは、実際のスナップショットボリュームではなく、プレビューを操作します。

管理者が選択したスナップショットをプレビューした後、プレビューをコミットして、ゲストデータをスナップショットでキャプチャされた状態に復元できます。管理者がプレビューをコミットすると、ゲストはプレビューレイヤーに接続されます。

スナップショットをプレビューした後に、管理者は **Undo** を選択して、表示したスナップショットのプレビュー層を破棄できます。スナップショット自体を含むレイヤーは、プレビューレイヤーが破棄されても保持されます。

図8.3 スナップショットのプレビュー



## 8.5. スナップショットの削除

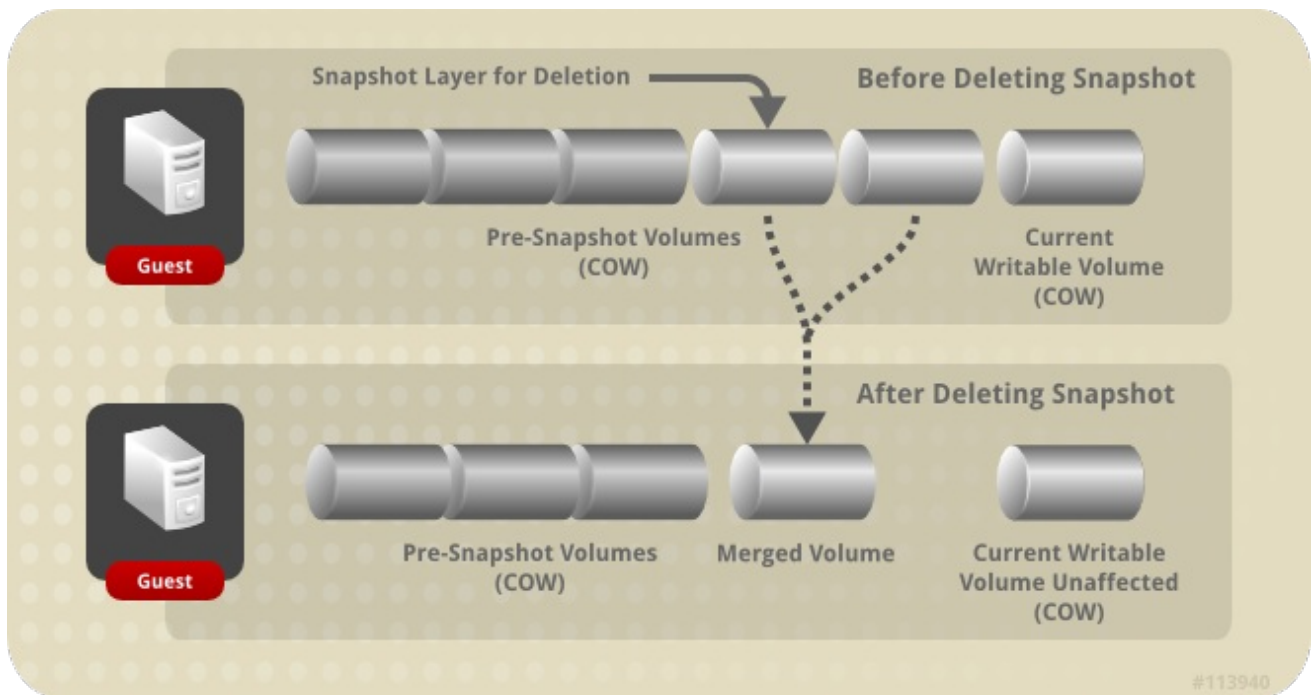
個々のスナップショット、または不要になった一連のスナップショットを削除できます。スナップショットを削除すると、仮想ディスクイメージをその特定の復元ポイントに復元する機能が削除されます。スナップショットによって消費されたディスクスペースを再利用する必要はなく、データを削除するわけでもありません。後続のスナップショットが削除されたスナップショットのデータを上書きした場合にのみ、ディスク領域が再利用されます。たとえば、5つのスナップショットのうち3番目のスナップショットを削除した場合、4番目と5番目のスナップショットを使用するには、3番目のスナップショットの変更されていないデータをディスクに保存する必要があります。ただし、4番目または5番目のスナップショットが3番目のデータを上書きした場合は、3番目のスナップショットが冗長化され、ディスク領域を再利用できます。ディスクスペースの再利用の可能性は別として、スナップショットを削除すると、仮想マシンのパフォーマンスも向上する可能性があります。

削除用にスナップショットを選択すると、QEMUは同じサイズの新しい論理ボリュームを作成し、削除されるスナップショットを後続のスナップショットとマージします。この新しい論理ボリュームのサイズが変更され、2つのスナップショット間のすべての違いに対応します。新しい論理ボリュームは、2つのスナップショットの合計の合計サイズになる可能性があります。2つのスナップショットをマージすると、後続のスナップショットの名前が変更され、削除のフラグが付けられ、新しい論理ボリュームに置き換えられ、その名前が使用されます。削除についてのフラグが付けられていたスナップショットと後続のスナップショットの両方が削除され、その場所は単一のマージされたスナップショットです。

たとえば、`snapshot Delete_snapshot` は 200 GB で、後続のスナップショット `Next_snapshot` は 100 GB です。`Delete_snapshot` が削除され、一時的に `Snapshot_merge` という名前の新しい論理ボリュームが作成され、サイズが 200 GB になります。`Snapshot_merge` は最終的に 300 GB にサイズ変更され、`Delete_snapshot` と `Next_snapshot` の両方のマージされたコンテンツの合計に対応します。`Next_snapshot` の名前が `Delete_me_too_snapshot` に変更され、`snapshot_merge` の名前が `Next_snapshot` に変更できるようにします。最後に、`Delete_snapshot` および `Delete_me_too_snapshot` が削除されます。



図8.4 スナップショットの削除



実行中の仮想マシンからスナップショットを削除するために使用されるロジックは、シャットダウンしている仮想マシンと若干異なります。ライブスナップショットの削除は非同期ブロックジョブとして処理され、VDSM が仮想マシンのリカバリーファイルで操作の記録を維持するため、操作中に VDSM が再起動したり、仮想マシンがシャットダウンされたりしても、ジョブを追跡できます。操作が開始すると、操作が失敗したり中断されたりした場合でも、削除されるスナップショットをプレビューしたり、復元ポイントとして使用したりすることはできません。アクティブレイヤーとその親レイヤーを統合する操作では、アクティブレイヤーから親レイヤーへのデータコピーと、アクティブレイヤーと親レイヤーへのディスクライトのミラーリングの 2 段階の処理に分割されます。最後に、削除されるスナップショットのデータが親スナップショットとマージされ、VDSM がイメージチェーン全体で変更を同期すると、ジョブは完了したと見なされます。

## 第9章 ハードウェアドライバーとデバイス

### 9.1. 仮想化されたハードウェア

Red Hat Virtualization は、仮想化されたゲストに 3 つの異なるタイプのシステムデバイスを提供します。これらのハードウェアデバイスはすべて、仮想化されたゲストに物理的に接続されたハードウェアデバイスとして表示されますが、デバイスドライバーはさまざまな方法で機能します。

#### エミュレートされたデバイス

エミュレートされたデバイス(仮想デバイスと呼ばれることもあります)は、完全にソフトウェアに存在します。エミュレートされたデバイスドライバーは、ホストで実行しているオペレーティングシステム(ソースデバイスを管理する)とゲストで実行しているオペレーティングシステムとの間の変換層です。エミュレートされたデバイスとの間でやり取りされるデバイスレベルの命令は、ハイパーバイザーによってインターセプトおよび変換されます。Linux カーネルによってエミュレートおよび認識されるのと同じタイプのデバイスは、エミュレートされたドライバーのバックギングソースデバイスとして使用できます。

#### 準仮想化デバイス

準仮想化デバイスでは、ゲストオペレーティングシステムにデバイスドライバーをインストールして、ホストマシンのハイパーバイザーと通信するためのインターフェイスを提供する必要があります。このインターフェイスは、仮想化環境の外部でディスク I/O などの従来の集中的なタスクを実行できるようにするために使用されます。この方法で仮想化に固有のオーバーヘッドを削減することは、ゲストオペレーティングシステムのパフォーマンスを、物理ハードウェアで直接実行する場合に期待されるパフォーマンスに近づけることを目的としています。

#### 物理的に共有されているデバイス

特定のハードウェアプラットフォームでは、仮想化されたゲストがさまざまなハードウェアデバイスやコンポーネントに直接アクセスできます。仮想化におけるこのプロセスは、パススルーまたはデバイス割り当てと呼ばれます。パススルーを使用すると、デバイスがゲストオペレーティングシステムに物理的に接続されているかのように表示および動作できます。

### 9.2. RED HAT VIRTUALIZATION における安定したデバイスアドレス

仮想ハードウェアの PCI アドレスの割り当ては、ovirt-engine データベースに永続化されます。

PCI アドレスは仮想マシンの作成時に QEMU によって割り当てられ、libvirt により VDSM に報告されます。VDSM は、ovirt-engine データベースに保存されている Manager に報告します。

仮想マシンが起動すると、Manager はデータベースからデバイスアドレスを VDSM に送信しま



す。VDSM は、仮想マシンの初回実行時に割り当てられた PCI デバイスアドレスを使用して仮想マシンを起動する libvirt に渡します。

デバイスが仮想マシンから削除されると、安定した PCI アドレスを含む、そのデバイスへのすべての参照も削除されます。削除したデバイスを交換するためにデバイスが追加されると、QEMU によって PCI アドレスが割り当てられます。これは、置き換えるデバイスと同じであるとは限りません。

### 9.3. 中央処理装置 (CPU)

クラスター内の各ホストには、多数の *仮想 CPU (vCPU)* があります。次に、仮想 CPU は、ホスト上で実行されているゲストに公開されます。クラスター内のホストによって公開されるすべての仮想 CPU は、クラスターが Red Hat Virtualization Manager を介して最初に作成されたときに選択されたタイプです。クラスター内で仮想 CPU タイプを混在させることはできません。

使用可能な各仮想 CPU タイプには、同じ名前の物理 CPU に基づく特性があります。仮想 CPU は、物理 CPU とゲストオペレーティングシステムを区別できません。



#### 注記

##### *x2APIC のサポート :*

Red Hat Enterprise Linux 7 ホストが提供するすべての仮想 CPU モデルには、*x2APIC* のサポートが含まれています。これにより、ハードウェア割り込みをより適切に処理するための *APIC (Advanced Programmable Interrupt Controller)* が提供されます。

### 9.4. システムデバイス

システムデバイスはゲストが実行するために重要であり、削除することはできません。ゲストに接続されている各システムデバイスも、使用可能な PCI スロットを使用します。デフォルトのシステムデバイスは次のとおりです。

- ホストブリッジ
- ISA ブリッジと USB ブリッジ (USB ブリッジと ISA ブリッジは同じデバイス)

- グラフィックカード(Cirrus ドライバーまたは qxl ドライバーのいずれかを使用)
- メモリーバルーンデバイス。

## 9.5. ネットワークデバイス

Red Hat Virtualization は、3つの異なるタイプのネットワークインターフェイスコントローラーをゲストに公開できます。ゲストに公開するネットワークインターフェイスコントローラーのタイプは、ゲストの作成時に選択されますが、Red Hat Virtualization Manager から変更できます。

- e1000 ネットワークインターフェイスコントローラーは、仮想化 Intel PRO/1000 (e1000) をゲストに公開します。
- virtio ネットワークインターフェイスコントローラーは、準仮想化ネットワークデバイスをゲストに公開します。
- rtl8139 ネットワークインターフェイスコントローラーは、仮想化された Realtek Semiconductor Corp RTL8139 をゲストに公開します。

ゲストごとに複数のネットワークインターフェイスコントローラーを使用できます。追加された各コントローラーは、ゲストで使用可能な PCI スロットを占有します。

## 9.6. グラフィックデバイス

2つのエミュレートされたグラフィックデバイスが提供されます。これらのデバイスは、SPICE プロトコルまたは VNC を使用して接続できます。

- ac97 は、Cirrus CLGD 5446 PCI VGA カードをエミュレートします。
- vga は、BochsVESA 拡張（すべての非標準モードを含むハードウェアレベル）を使用してダミー VGA カードをエミュレートします。

## 9.7. ストレージデバイス

ストレージデバイスとストレージプールは、ブロックデバイスドライバーを使用して、ストレージデバイスを仮想化ゲストに接続できます。ストレージドライバーはストレージデバイスではないことに注意してください。ドライバーは、バックিংストレージデバイス、ファイル、またはストレージプールボリュームを仮想化されたゲストに接続するのに使用されます。バックিংストレージデバイスは、サポートされている任意のタイプのストレージデバイス、ファイル、またはストレージプールボリュームにすることができます。

- IDE ドライバーは、エミュレートされたブロックデバイスをゲストに公開します。エミュレートされた IDE ドライバーは、最大 4 つの仮想化 IDE ハードディスクまたは仮想化 IDE CD-ROM ドライブの組み合わせを各仮想化ゲストに割り当てるために使用できます。エミュレートされた IDE ドライバーは、仮想化 DVD-ROM ドライブを提供するためにも使用されます。
- VirtIO ドライバーは、準仮想化ブロックデバイスをゲストに公開します。準仮想化ブロックドライバーは、仮想化ゲストに接続されたハイパーバイザーによってサポートされるすべてのストレージデバイス用のドライバーです (エミュレートする必要があるフロッピーディスクドライブを除く)。

## 9.8. サウンドデバイス

2 つのエミュレートされたサウンドデバイスが利用可能です。

- ac97 は、Intel 82801AA AC97 Audio と互換性のあるサウンドカードをエミュレートします。
- es1370 は、ENSONIQ AudioPCI ES1370 サウンドカードをエミュレートします。

## 9.9. シリアルドライバー

準仮想化シリアルドライバー(virtio-serial)は、バイト指向の文字ストリームドライバーです。準仮想化シリアルドライバーは、ネットワークが利用できない、または使用できない場合に、ホストのユーザー空間とゲストのユーザー空間との間の単純な通信インターフェイスを提供します。

## 9.10. バルーンドライバー

バルーンドライバーを使用すると、ゲストは必要なメモリー量をハイパーバイザーに表現できます。バルーンドライバーを使用すると、ホストはゲストに効率的にメモリーを割り当て、他のゲストやプロ

セスに空きメモリーを割り当てることができます。

バルーンドライバーを使用するゲストは、ゲストの RAM のセクションを未使用としてマークできます (バルーンの反転)。ハイパーバイザーはメモリーを解放し、そのメモリーを他のホストプロセスまたはそのホスト上の他のゲストに使用できます。ゲストが解放されたメモリーを再度必要とする場合、ハイパーバイザーは RAM をゲストに再割り当てできます (バルーンの収縮)。

## 第10章 最小要件および技術制限

### 10.1. 最小要件およびサポートされる制限

Red Hat Virtualization 環境に適用される物理および論理の制限が複数あります。この制限を超える設定がある環境では、現在サポートされていません。

### 10.2. リソースの制限

特定の制限は、ストレージドメインやホストなどのリソースに適用されます。

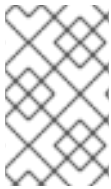
表10.1 リソースの制限

項目	制限事項
ストレージドメイン	<p>データセンターごとに少なくとも2つのストレージドメインが推奨されます。</p> <ul style="list-style-type: none"> <li>● データストレージドメインが必要です。</li> <li>● ISO ストレージドメインが推奨されます。</li> </ul>
ホスト	<p>Red Hat は、Red Hat Virtualization Manager ごとに最大 200 のホストをサポートします。</p>

### 10.3. クラスターの制限

クラスターは、一連の仮想マシンのリソースプールとして扱われる物理ホストのセットです。クラスター内のホストは、同じネットワークインフラストラクチャーと同じストレージを共有します。クラスターは移行ドメインで、仮想マシンをホストからホストに移行できます。安定性を確保するために、各クラスターにいくつかの制限が適用されます。

- すべての管理対象ハイパーバイザーがクラスター内にある必要があります。
- クラスター内のすべての管理対象ハイパーバイザーは、同じ CPU タイプを持つ必要があります。Intel と AMD CPU は同じクラスター内で共存できません。



## 注記

クラスターの詳細は、『Administration Guide』の [Clusters](#) を参照してください。

## 10.4. ストレージドメインの制限

ストレージドメインは、仮想ディスクイメージおよび ISO イメージのストレージの領域と、仮想マシンのインポートおよびエクスポートを行います。特定のデータセンター内に多数のストレージドメインが作成される場合がありますが、各ストレージドメインに適用される制限や推奨事項は複数あります。

表10.2 ストレージドメインの制限

項目	制限事項
ストレージタイプ	<p>サポート対象のストレージタイプは以下のとおりです。</p> <ul style="list-style-type: none"> <li>● ファイバーチャネルプロトコル (FCP)</li> <li>● Internet Small Computer System Interface (iSCSI)</li> <li>● Network File System (NFS)</li> <li>● POSIX 準拠ファイルシステム (POSIX)</li> <li>● Red Hat Gluster Storage (GlusterFS)</li> </ul> <p>Red Hat Virtualization 4.0 の新しい ISO およびエクスポートストレージドメインは、任意のファイルベースのストレージ (NFS、Posix、または GlusterFS) で提供できます。</p>
論理ユニット番号 (LUN)	<p>iSCSI または FCP が提供するストレージドメインごとに 300 を超える LUN は許可されません。</p>

項目	制限事項
論理ボリューム(LV)	<p>Red Hat Virtualization では、論理ボリュームは、仮想マシン、テンプレート、および仮想マシンスナップショットの仮想ディスクを表します。</p> <p>iSCSI または FCP が提供するストレージドメインごとに、350 を超える論理ボリュームは推奨されません。特定のストレージドメインの論理ボリュームの数がこの数を超える場合は、利用可能なストレージを 350 を超える論理ボリュームを持たない個別のストレージドメインに分割することが推奨されます。</p> <p>この制限の根本的な原因は、LVM メタデータのサイズです。論理ボリュームの数が増えると、その論理ボリュームに関連付けられた LVM メタデータも増加します。このメタデータのサイズが 1MB を超えると、新規ディスクの作成やスナップショットの作成などのプロビジョニング操作のパフォーマンスが減少し、QCOW ディスクの実行時に論理ボリュームのシンプロビジョニングに時間がかかることがあります。</p> <p>論理ボリュームの  <a href="https://access.redhat.com/solutions/441203">https://access.redhat.com/solutions/441203</a> 詳細は、を参照してください。</p>



### 注記

ストレージドメインの詳細は、『Administration Guide』の **Storage** を参照してください。

## 10.5. RED HAT VIRTUALIZATION MANAGER の制限

Red Hat Virtualization Manager サーバーは Red Hat Enterprise Linux 7 を実行する必要があります。さらに多くのハードウェア要件も満たしている必要があります。

表10.3 Red Hat Virtualization Manager の制限

項目	制限事項
RAM	<ul style="list-style-type: none"> <li>● 最小 4 GB の RAM が必要です。</li> </ul>
PCI デバイス	<ul style="list-style-type: none"> <li>● 最小帯域幅が 1Gbps のネットワークコントローラーが少なくとも 1Gbps であるネットワークコントローラーが推奨されます。</li> </ul>

項目	制限事項
ストレージ	<ul style="list-style-type: none"> <li>最低 25 GB の利用可能なローカルディスク領域が推奨されます。</li> </ul>



#### 注記

Red Hat Virtualization Manager の詳細は、[インストールガイド](#) を参照してください。

## 10.6. ハイパーバイザーの要件

Red Hat Virtualization Host (RHVH)には、多くのハードウェア要件とサポートされる制限があります。Red Hat Enterprise Linux ホストのストレージ要件は、既存の設定で使用されるディスク容量によって異なりますが、RHVH の要件よりも多くなるはずはです。

表10.4 Red Hat Virtualization ホストの要件とサポートされる制限

項目	サポートの制限
----	---------



項目	サポートの制限
CPU	<p>最低1つの物理 CPU が必要です。Red Hat Virtualization は、ホストでこれらの CPU モデルの使用をサポートします。</p> <ul style="list-style-type: none"> <li>● AMD Opteron G1</li> <li>● AMD Opteron G2</li> <li>● AMD Opteron G3</li> <li>● AMD Opteron G4</li> <li>● AMD Opteron G5</li> <li>● Intel Conroe</li> <li>● Intel Penryn</li> <li>● Intel Nehalem</li> <li>● Intel Westmere</li> <li>● Intel Haswell</li> <li>● Intel SandyBridge ファミリー</li> <li>● IBM POWER 8</li> </ul> <p>すべての CPU は、Intel® 64 または AMD64 CPU 拡張機能をサポートしており、AMD-V™ または Intel VT® ハードウェア仮想化拡張機能が有効になっている必要があります。<b>No eXecute</b> flag (NX) のサポートも必要です。</p>
RAM	<p>各仮想マシンに必要な RAM 容量は、以下によって異なります。</p> <ul style="list-style-type: none"> <li>● ゲストオペレーティングシステムの要件</li> <li>● ゲストアプリケーションの要件</li> <li>● 仮想マシンのメモリアクティビティと使用状況。</li> </ul> <p>また、KVM は仮想マシンの物理 RAM をオーバーコミットできます。これは、必要に応じて仮想マシンに RAM のみを割り当て、使用率の低い仮想マシンを swap に移動することで行います。</p> <p>サポートされている最大および最小 RAM については <a href="https://access.redhat.com/articles/rhel-limits">https://access.redhat.com/articles/rhel-limits</a>、を参照してください。</p>

項目	サポートの制限
ストレージ	<p>ホストでサポートされる最小内部ストレージは、以下の一覧の合計です。</p> <ul style="list-style-type: none"> <li>● ルート(/)パーティションには、少なくとも 6 GB のストレージが必要です。</li> <li>● /boot パーティションには、少なくとも 1 GB のストレージが必要です。</li> <li>● /var パーティションには少なくとも 15 GB のストレージが必要です。セルフホストエンジンのデプロイメントの場合、これは 60 GB 以上である必要があります。</li> <li>● スワップパーティションには、8 MB 以上のストレージが必要です。推奨されるスワップパーティションのサイズは、ホストがインストールされているシステムと、環境に必要なオーバーコミットのレベルによって異なります。詳細は、を参照 <a href="https://access.redhat.com/solutions/15244">https://access.redhat.com/solutions/15244</a> してください。</li> </ul> <p>これらは、ホストのインストールの <b>最小</b> ストレージ要件であることに注意してください。より多くのストレージ領域を使用するデフォルトの割り当てを使用することが推奨されます。</p>
PCI デバイス	1 Gbps の推奨最小帯域幅には、少なくとも 1 つのネットワークコントローラーが必要です。

## 重要

Red Hat Virtualization Host の起動時に、以下のようなメッセージが表示される場合があります。

```
Virtualization hardware is unavailable.
(No virtualization hardware was detected on this system)
```

この警告は、仮想化拡張機能が無効であるか、プロセッサに存在しないことを示しています。CPU が一覧表示された拡張機能をサポートし、システム BIOS で有効になっていることを確認します。

プロセッサに仮想化拡張機能があり、それらが有効になっていることを確認するには、以下を実行します。

- ホストブート画面では、任意のキーを押して、一覧から **Boot** または **Boot with serial console** エントリーを選択します。Tab を押して、選択したオブ

ションのカーネルパラメーターを編集します。最後のカーネルパラメーターの後に `Space` があり、`rescue` パラメーターを追加します。

- `Enter` を押して、レスキューモードで起動します。
- 表示されるプロンプトで、以下のコマンドを実行してプロセッサに仮想化拡張機能があり、それらが有効になっていることを確認します。

```
# grep -E 'svm|vmx' /proc/cpuinfo
```

何らかの出力が表示されれば、プロセッサはハードウェアの仮想化が可能です。出力が表示されない場合は、プロセッサがハードウェア仮想化をサポートしている可能性があります。状況によっては、製造元が BIOS で仮想化拡張機能を無効にする場合があります。この場合、製造元が提供するシステムの BIOS およびマザーボードマニュアルを参照してください。

- 追加のチェックとして、`kvm` モジュールがカーネルに読み込まれていることを確認します。

```
# lsmod | grep kvm
```

出力に `kvm_intel` または `kvm_amd` が含まれる場合は、`kvm` ハードウェア仮想化モジュールが読み込まれ、システムが要件を満たします。

## 10.7. ゲストの要件とサポートの制限

以下の要件とサポート制限は、Red Hat Virtualization Host (RHVH)で実行されるゲストに適用されます。

表10.5 仮想化されたハードウェア

項目	制限事項
CPU	Red Hat Enterprise Linux 7 では、ゲストごとに最大 240 の仮想化 CPU がサポートされます。

項目	制限事項
RAM	<p>ゲストごとに異なるメモリー要件があります。各ゲストに必要な RAM 容量は、ゲストオペレーティングシステムの要件と、ゲストが動作している負荷によって異なります。</p> <p>ゲストマシンの最大およびサポートされる最小 RAM については <a href="https://access.redhat.com/articles/rhel-kvm-limits">https://access.redhat.com/articles/rhel-kvm-limits</a>、を参照してください。</p>
PCI デバイス	<p>ゲストごとに最大 31 個の仮想化 PCI デバイスがサポートされます。この制限に対するシステムデバイスの数。一部は必須です。PCI デバイスの制限に対してカウントする必須デバイスには、PCI ホストブリッジ、ISA ブリッジ、USB ブリッジ、ボードブリッジ、グラフィックカード、IDE または VirtIO ブロックデバイスが含まれます。</p>
ストレージ	<p>ゲストごとに最大 28 個の仮想化ストレージデバイスがサポートされ、可能な 3 つの IDE および 25 Virtio で設定されます。</p>

## 10.8. SPICE の制限

SPICE が現在サポートしている最大解像度は 2560 x 1600 ピクセルです。

## 10.9. 追加の参考資料

これらの追加のドキュメントリソースは、Red Hat Virtualization ドキュメントスイートの一部ではありません。ただし、Red Hat Virtualization 環境を管理するシステム管理者に役立つ情報が含まれ、から入手 <https://access.redhat.com/documentation/en/red-hat-enterprise-linux/> できます。

### 『Red Hat Enterprise Linux - System Administrator's Guide』

Red Hat Enterprise Linux のデプロイメント、設定、および管理のガイド

### 『Red Hat Enterprise Linux - DM-Multipath Guide』

Red Hat Enterprise Linux での Device-Mapper Multipathing の使用に関するガイド

### 『Red Hat Enterprise Linux - Installation Guide』

Red Hat Enterprise Linux のインストールガイド

『Red Hat Enterprise Linux - Storage Administration Guide』

Red Hat Enterprise Linux でストレージデバイスおよびファイルシステムを管理するガイド

『Red Hat Enterprise Linux - Virtualization Deployment and Administration Guide』

Red Hat Enterprise Linux の仮想化テクノロジーのインストール、設定、管理、およびトラブルシューティングのガイド