



Red Hat

Red Hat Enterprise Linux 8

配置和管理虚拟化

设置您的主机、创建和管理虚拟机并了解 Red Hat Enterprise Linux 8 中的虚拟化功能

Red Hat Enterprise Linux 8 配置和管理虚拟化

设置您的主机、创建和管理虚拟机并了解 Red Hat Enterprise Linux 8 中的虚拟化功能

法律通告

Copyright © 2023 Red Hat, Inc.

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

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

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

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

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

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

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

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

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

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

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

All other trademarks are the property of their respective owners.

摘要

要将 Red Hat Enterprise Linux (RHEL) 系统用作虚拟化主机, 请按照本文档中的说明操作。提供的信息包括: 虚拟化的功能和用例 如何使用命令行工具以及 web 控制台管理您的主机和虚拟机 虚拟化的支持限制在各种系统构架中, 如 Intel 64、AMD64、IBM POWER 和 IBM Z

目录

使开源包含更多	5
对红帽文档提供反馈	6
第 1 章 介绍 RHEL 中的虚拟化	7
1.1. 什么是虚拟化？	7
1.2. 虚拟化的优点	7
1.3. 虚拟机组件及其交互	8
1.4. 用于虚拟化管理的工具和界面	9
1.5. 红帽虚拟化解决方案	10
第 2 章 虚拟化入门	11
2.1. 启用虚拟化	11
2.2. 创建虚拟机	12
2.3. 启动虚拟机	22
2.4. 连接至虚拟机	25
2.5. 关闭虚拟机	35
2.6. 删 除 虚 拟 机	37
2.7. 其它资源	38
第 3 章 在 IBM POWER 中使用虚拟化	39
3.1. 在 IBM POWER 中启用虚拟化	39
3.2. IBM POWER 中的虚拟化与 AMD64 和 INTEL 64 的不同	40
第 4 章 在 IBM Z 中使用虚拟化	43
4.1. 在 IBM Z 中启用虚拟化	43
4.2. 将 IBM Z 上的虚拟化从 RHEL 8.5 更新至 RHEL 8.6 或更高版本	44
4.3. IBM Z 上的虚拟化与 AMD64 和 INTEL 64 有什么不同	45
4.4. 后续步骤	48
4.5. 其它资源	48
第 5 章 在 WEB 控制台中管理虚拟机	49
5.1. 使用 WEB 控制台管理虚拟机概述	49
5.2. 设置 WEB 控制台以管理虚拟机	49
5.3. 使用 WEB 控制台重命名虚拟机	50
5.4. WEB 控制台中提供的虚拟机管理功能	51
5.5. 虚拟机管理器和 WEB 控制台中虚拟化功能的不同	52
第 6 章 查看有关虚拟机的信息	54
6.1. 使用命令行界面查看虚拟机信息	54
6.2. 使用 WEB 控制台查看虚拟机信息	56
6.3. 虚拟机 XML 配置示例	62
第 7 章 保存和恢复虚拟机	67
7.1. 保存和恢复虚拟机的工作方式	67
7.2. 使用命令行界面保存虚拟机	67
7.3. 使用命令行界面启动虚拟机	68
7.4. 使用 WEB 控制台启动虚拟机	69
第 8 章 克隆虚拟机	70
8.1. 克隆虚拟机的方式	70
8.2. 创建虚拟机模板	70
8.3. 使用命令行界面克隆虚拟机	74
8.4. 使用 WEB 控制台克隆虚拟机	75

第 9 章 迁移虚拟机	77
9.1. 迁移虚拟机的工作方式	77
9.2. 迁移虚拟机的好处	77
9.3. 迁移虚拟机的限制	78
9.4. 验证虚拟机迁移的主机 CPU 兼容性	79
9.5. 将虚拟机磁盘镜像与其他主机共享	82
9.6. 使用命令行界面迁移虚拟机	83
9.7. 使用 WEB 控制台实时迁移虚拟机	86
9.8. 虚拟机迁移支持的主机	88
9.9. 其它资源	89
第 10 章 管理虚拟设备	90
10.1. 虚拟设备的工作原理	90
10.2. 虚拟设备类型	91
10.3. 使用 CLI 管理附加到虚拟机的设备	92
10.4. 使用 WEB 控制台管理主机设备	96
10.5. 管理虚拟 USB 设备	100
10.6. 管理虚拟光驱	103
10.7. 管理 SR-IOV 设备	106
10.8. 将 DASD 设备附加到 IBM Z 中的虚拟机	110
第 11 章 为虚拟机管理存储	114
11.1. 了解虚拟机存储	114
11.2. 使用 CLI 管理虚拟机存储池	116
11.3. 使用 WEB 控制台管理虚拟机存储池	130
11.4. 创建存储池的参数	144
11.5. 使用 CLI 管理虚拟机存储卷	153
11.6. 使用 WEB 控制台管理虚拟机存储卷	156
11.7. 使用 WEB 控制台管理虚拟机存储磁盘	159
11.8. 使用 LIBVIRT SECRET 保护 iSCSI 存储池	164
11.9. 创建 VHBA	165
第 12 章 在虚拟机中管理 GPU 设备	168
12.1. 为虚拟机分配 GPU	168
12.2. 管理 NVIDIA VGPU 设备	170
第 13 章 配置虚拟机网络连接	177
13.1. 了解虚拟网络	177
13.2. 使用 WEB 控制台管理虚拟机网络接口	178
13.3. 推荐的虚拟机网络配置	181
13.4. 虚拟机网络连接的类型	185
13.5. 从 PXE 服务器启动虚拟机	188
13.6. 其它资源	192
第 14 章 在主机及其虚拟机间共享文件	193
14.1. 使用 NFS 在主机及其虚拟机间共享文件	193
第 15 章 保护虚拟机	196
15.1. 虚拟机中的安全性是如何工作的	196
15.2. 保护虚拟机的最佳实践	197
15.3. 创建 SECUREBOOT 虚拟机	198
15.4. 限制虚拟机用户可以使用哪些操作	199
15.5. 虚拟机安全性的自动功能	200
15.6. 用于虚拟化的 SELINUX 布尔值	201
15.7. 在 IBM Z 上设置 IBM SECURE EXECUTION	202

15.8. 将加密 COPROCESSORS 附加到 IBM Z 上的虚拟机	205
15.9. 在 WINDOWS 虚拟机中启用标准硬件安全性	208
15.10. 在 WINDOWS 虚拟机上启用增强的硬件安全	209
第 16 章 优化虚拟机性能	211
16.1. 影响虚拟机性能的因素	211
16.2. 使用 TUNED 优化虚拟机性能	211
16.3. 配置虚拟机内存	212
16.4. 优化虚拟机 I/O 性能	216
16.5. 优化虚拟机 CPU 性能	218
16.6. 优化虚拟机网络性能	228
16.7. 虚拟机性能监控工具	229
16.8. 其它资源	231
第 17 章 安装和管理 WINDOWS 虚拟机	232
17.1. 安装 WINDOWS 虚拟机	232
17.2. 优化 WINDOWS 虚拟机	233
17.3. 在 WINDOWS 虚拟机中启用标准硬件安全性	246
17.4. 在 WINDOWS 虚拟机上启用增强的硬件安全	247
17.5. 后续步骤	248
第 18 章 创建嵌套虚拟机	249
18.1. 在 INTEL 上创建嵌套虚拟机	249
18.2. 在 AMD 上创建嵌套虚拟机	250
18.3. 在 IBM Z 上创建嵌套的虚拟机	252
18.4. 在 IBM POWER9 上创建嵌套虚拟机	252
18.5. 嵌套虚拟化的限制	254
第 19 章 诊断虚拟机问题	256
19.1. 生成 LIBVIRT 调试日志	256
19.2. 转储虚拟机内核	258
19.3. 回溯虚拟机进程	260
第 20 章 RHEL 8 虚拟化的功能支持和限制	262
20.1. RHEL 虚拟化支持如何工作	262
20.2. RHEL 8 虚拟化中的推荐功能	262
20.3. RHEL 8 虚拟化不支持的功能	263
20.4. RHEL 8 虚拟化中的资源分配限制	267
20.5. RHEL 8 支持虚拟化功能概述	267

使开源包含更多

红帽承诺替换我们的代码、文档和网页属性中存在问题的语言。我们从这四个术语开始：master、slave、blacklist 和 whitelist。这些更改将在即将发行的几个发行本中逐渐实施。如需了解更多详细信息，请参阅 [CTO Chris Wright 信息](#)。

对红帽文档提供反馈

我们感谢您对我们文档的反馈。让我们了解如何改进它。

根据具体内容提交评论

1. 查看 **Multi-page HTML** 格式的文档，并确保在页面完全加载后看到右上角的 **反馈** 按钮。
2. 使用光标突出显示您要评论的文本部分。
3. 点击在高亮文本旁的 **添加反馈** 按钮。
4. 添加您的反馈并点 **Submit**。

通过 Bugzilla 提交反馈（需要帐户）

1. 登录到 [Bugzilla](#) 网站。
2. 从 **Version** 菜单中选择正确的版本。
3. 在 **Summary** 字段中输入描述性标题。
4. 在 **Description** 字段中输入您的建议以改进。包括文档相关部分的链接。
5. 点 **Submit Bug**。

第1章 介绍 RHEL 中的虚拟化

如果您对虚拟化概念或其在 Linux 中的实施不熟悉，以下小节提供了 RHEL 8 虚拟化的一般概述：红帽提供的基本知识、优势、组件和其他可能的虚拟化解决方案。

1.1. 什么是虚拟化？

RHEL 8 提供虚拟化功能，它让运行 RHEL 8 的机器能够 *托管* (*host*) 多个虚拟机 (VM)，也称为 *客户机* (*guest*)。VM 使用主机的物理硬件和计算资源，在主机操作系统中作为用户空间进程运行一个独立的虚拟操作系统（*客户机操作系统*）。

换句话说，虚拟化功能使在一个操作系统中执行其他操作系统成为可能。

VM 可让您安全地测试软件配置和功能，运行旧的软件或优化硬件的工作负载效率。有关优点的更多信息，请参阅 [虚拟化的优势](#)。

有关虚拟化的更多信息，请参阅[红帽客户门户网站](#)。

后续步骤

- 要在 Red Hat Enterprise Linux 8 中尝试虚拟化，[请参阅开始使用虚拟化](#)。
- 除了 Red Hat Enterprise Linux 8 虚拟化外，红帽还提供了多种专用的虚拟化解决方案，每种解决方案都有不同的用户重点和特性。如需更多信息，请参阅 [Red Hat 虚拟化解决方案](#)。

1.2. 虚拟化的优点

与使用物理机器相比，使用虚拟机 (VM) 有以下优点：

• 灵活精细的资源分配

一个虚拟机在主机机器（通常是物理机器）上运行，主机的硬件也可以分配给客户机操作系统使用。但是，物理资源分配是在软件级别上完成的，因此非常灵活。虚拟机使用的主机内存、CPU 或存储空间的比例是可以配置的，可以指定非常精细的资源请求。

例如：客户机操作系统的磁盘可以在主机的文件系统中以一个文件代表，且该磁盘的大小限制比物理磁盘的限制要小。

• 软件控制的配置

虚拟机的整个配置都作为数据保存在主机上，并由软件控制。因此，虚拟机可轻松创建、删除、克隆、迁移、远程操作或连接到远程存储。

• 与主机分离

在虚拟内核中运行的客户机操作系统与主机操作系统分开。这意味着可在虚拟机中安装任何操作系统，即使虚拟机操作系统不稳定或受损，主机也不会受到任何影响。

• 空间及成本效率

单个物理机器可以托管大量虚拟机。因此，无需多个物理机器执行同样的任务，因此降低了与物理硬件关联的空间、电源和维护的要求。

• 软件兼容性

因为虚拟机可以使用不同于其主机的操作系统，所以通过虚拟化，可以运行最初没有为主机操作系统发布的应用程序。例如，使用 RHEL 7 客户机操作系统，您可以在 RHEL 8 主机系统上运行为 RHEL 7 发布的应用程序。



注意

不是所有操作系统都可以作为 RHEL 8 主机中的客户机操作系统被支持。详情请查看 [RHEL 8 虚拟化中的推荐功能](#)。

1.3. 虚拟机组件及其交互

RHEL 8 中的虚拟化由以下主要软件组件组成：

虚拟机监控程序（Hypervisor）

在 RHEL 8 中创建虚拟机（VM）的基础是（hypervisor（虚拟机监控程序）），它是一个软件层用来控制硬件并在主机中运行多个操作系统。

虚拟机监控程序包括 **基于内核的虚拟机(KVM)** 模块和虚拟化内核驱动程序。这些组件可确保主机中的 Linux 内核为用户空间软件提供虚拟化资源。

在用户空间级别，**QEMU** 模拟器会模拟一个客户机操作系统可以在上面运行的完整虚拟硬件平台，并管理如何在主机中分配资源并提供给客户机。

此外，**libvirt** 软件套件充当管理和通信层，使与 QEMU 更容易交互、实施安全规则，并提供用于配置和运行 VM 的许多其他工具。

XML 配置

基于主机的 XML 配置文件（也称域 XML 文件）决定了特定虚拟机中的所有设置和设备。配置包括：

- 元数据，如虚拟机名称、时区和其他有关虚拟机的信息。
- 对虚拟机中的设备的描述，包括虚拟 CPU（vCPUs）、存储设备、输入/输出设备、网络接口卡及其他真实和虚拟硬件。
- 虚拟机设置，如它可以使用的最大内存量、重启设置和其他有关虚拟机行为的设置。

有关 XML 配置内容的更多信息，请参阅 [示例虚拟机 XML 配置](#)。

组件交互

当虚拟机启动时，虚拟机监控程序使用 XML 配置在主机上以用户空间进程的形式创建虚拟机实例。hypervisor 还使虚拟机进程能被基于主机的接口访问，如 **virsh**、**virt-install** 和 **guestfish** 工具，或者 Web 控制台 GUI。

当使用这些虚拟化工具时，libvirt 会将它们的输入转换成 QEMU 的指令。QEMU 将指令信息发送到 KVM，这样可确保内核正确分配执行该指令所需的资源。因此，QEMU 可以执行相应的用户空间更改，如创建或修改虚拟机或在虚拟机的客户机操作系统中执行操作。

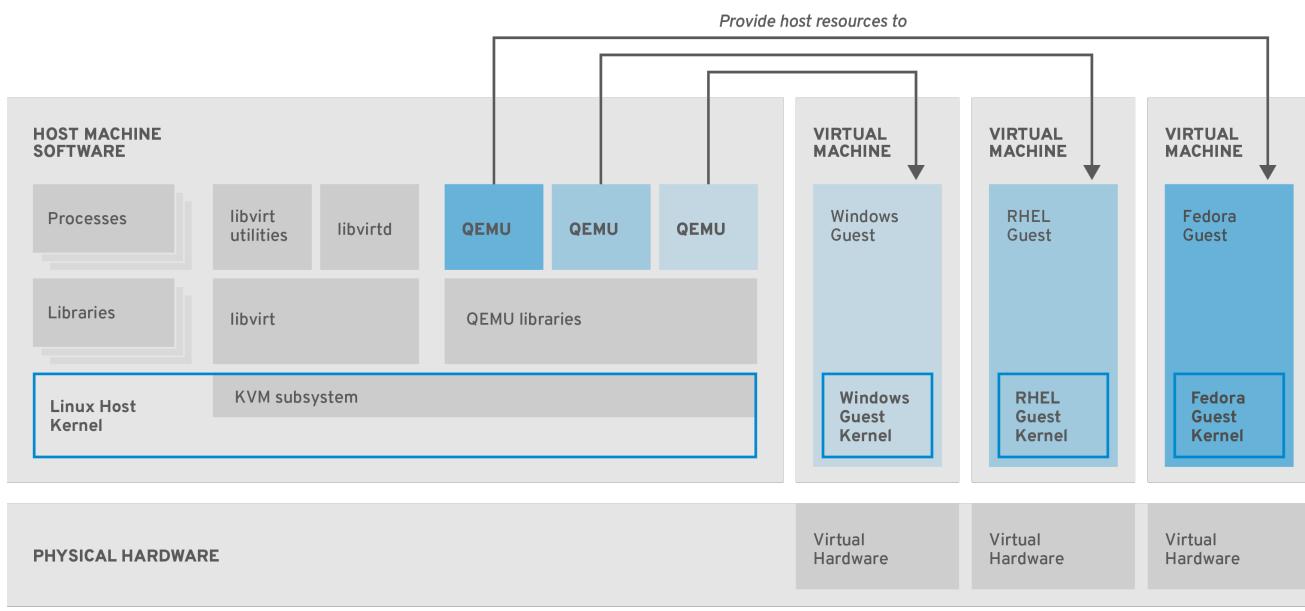


注意

尽管 QEMU 是构架中的基本组件，但出于安全考虑，它并没有被设计为直接在 RHEL 8 系统中使用。因此，红帽不支持使用 **qemu-*** 命令，强烈建议您使用 libvirt 与 QEMU 进行交互。

有关基于主机的接口的更多信息，请参阅 [虚拟化管理的工具和接口](#)。

图 1.1. RHEL 8 虚拟化架构



RHEL_7_0319

1.4. 用于虚拟化管理的工具和界面

您可以使用命令行界面（CLI）或几个图形用户界面（GUI）管理 RHEL 8 中的虚拟化。

命令行界面

CLI 是在 RHEL 8 中管理虚拟化的最强大的方法。虚拟机（VM）管理的 CLI 命令包括：

- **virsh** - 一个用途广泛的虚拟化命令行工具，根据提供的参数，可以实现不同功能。例如：
 - 启动和关闭虚拟机 - **virsh start** 和 **virsh shutdown**
 - 列出可用的虚拟机 - **virsh list**
 - 从配置文件创建虚拟机 - **virsh create**
 - 进入虚拟化 shell - **virsh**

如需更多信息，请参阅 **virsh(1)** 手册页。

- **virt-install** - 用于创建新虚拟机的 CLI 工具。如需更多信息，请参阅 **virt-install(1)** 手册页。
- **virt-xml** - 用于编辑虚拟机配置的工具。
- **guestfish** - 用于检查和修改虚拟机磁盘镜像的工具。如需更多信息，请参阅 **guestfish(1)** 手册页。

图形界面

您可以使用以下 GUI 在 RHEL 8 中管理虚拟化：

- **RHEL 8 web 控制台**（也称为 Cockpit）提供了一个远程的、易于使用的图形用户界面管理虚拟机和虚拟化主机。
有关使用 web 控制台进行基本虚拟化管理的步骤，请参阅 [web 控制台中管理虚拟机](#)。

- 虚拟机管理器 (**virt-manager**) 应用程序提供管理虚拟机和虚拟化主机的专用 GUI。



重要

虽然 RHEL 8 中仍在支持，但 **virt-manager** 已被弃用。在以后的版本中计划使用 Web 控制台替换它。因此，建议您熟悉通过 web 控制台使用 GUI 管理虚拟化。

但是，在 RHEL 8 中，某些功能只能从 **virt-manager** 或命令行访问。详情请参阅[虚拟机管理器和 web 控制台中虚拟化功能之间的差别](#)。

- **Gnome Boxes** 应用程序是一个轻量级图形界面，可用来查看和访问虚拟机和远程系统。GNOME Boxes 主要用于桌面系统。



重要

GNOME Boxes 作为 GNOME 桌面环境的一部分提供，在 RHEL 8 中被支持，但红帽建议您通过 web 控制台使用 GUI 管理虚拟化。

其它资源

- [虚拟化入门](#)

1.5. 红帽虚拟化解决方案

以下红帽产品基于 RHEL 8 虚拟化功能构建，并扩展了 RHEL 8 中的 KVM 虚拟化功能。另外，许多[RHEL 8 虚拟化的限制](#) 不适用于这些产品：

OpenShift Virtualization

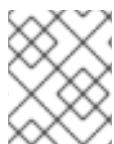
OpenShift Virtualization 基于 KubeVirt 技术，作为 Red Hat OpenShift Container Platform 的一部分，并可在容器中运行虚拟机。

有关 OpenShift Virtualization 的更多信息，请参阅[红帽混合云](#)页面。

Red Hat OpenStack Platform (RHOSP)

Red Hat OpenStack Platform 为创建、部署并扩展一个安全可靠的公共或私有 [OpenStack](#) 云环境提供了一个集成的基础。

如需有关 Red Hat OpenStack Platform 的更多信息，请参阅[红帽客户门户网站](#)或 [Red Hat OpenStack Platform 文档套件](#)。



注意

有关 RHEL 不支持的虚拟化功能，但在其他 Red Hat 虚拟化解决方案中被支持，请参阅：[RHEL 8 虚拟化中不支持的功能](#)

第 2 章 虚拟化入门

要开始使用 RHEL 8 中的虚拟化，请按照以下步骤操作。这样做的默认方法是使用命令行界面(CLI)，但为了方便用户，一些步骤可以在 Web 控制台 GUI 中完成。

1. 启用虚拟化模块并安装虚拟化软件包 - [请参阅启用虚拟化](#)。
2. 创建虚拟机：
 - 对于 CLI，请参阅 [使用命令行界面 创建虚拟机](#)。
 - 对于 GUI，请参阅 [创建虚拟机并使用 web 控制台安装客户机操作系统](#)。
3. 启动虚拟机：
 - 对于 CLI，请参阅 [使用命令行界面 启动虚拟机](#)。
 - 对于 GUI，请参阅 [使用 web 控制台启动虚拟机](#)。
4. 连接到虚拟机：
 - 对于 CLI，请参阅 [使用 SSH 连接到虚拟机](#)，或 [使用 Virt Viewer 打开虚拟机图形控制台](#)。
 - 对于 GUI，请参阅 [使用 web 控制台与虚拟机交互](#)。



注意

Web 控制台目前仅提供一部分虚拟机管理功能，因此对于 RHEL 8 中虚拟化的高级使用，建议使用命令行。

2.1. 启用虚拟化

要在 RHEL 8 中使用虚拟化，您必须启用虚拟化模块，安装虚拟化软件包，并确保您的系统被配置为可托管虚拟机(VM)。

先决条件

- RHEL 8 已安装并在主机中注册。
- 您的系统满足以下硬件要求以作为虚拟主机工作：
 - 有以下最小系统资源可用：
 - 主机有 6 GB 可用磁盘空间，以及每个预期的虚拟机需要额外 6 GB 空间。
 - 主机需要 2 GB RAM，以及每个预期的虚拟机需要额外 2 GB。
 - 主机上有 4 个 CPU。虚拟机通常可以使用单个分配的 vCPU 运行，但红帽建议为每个虚拟机分配 2 个或更多 vCPU，以避免虚拟机在高负载期间变得无响应。
 - 主机的构架 支持 KVM 虚拟化。
 - 值得注意的是，RHEL 8 不支持 64 位 ARM 架构(ARM 64)上的虚拟化。
 - 以下流程适用于 AMD64 和 Intel 64 架构(x86_64)。要在具有不同的受支持构架的主机上启用虚拟化，请参阅以下部分之一：

- 在 IBM POWER 中启用虚拟化
- 在 IBM Z 中启用虚拟化

流程

1. 在 RHEL 8 虚拟化模块中安装软件包：

```
# yum module install virt
```

2. 安装 **virt-install** 和 **virt-viewer** 软件包：

```
# yum install virt-install virt-viewer
```

3. 启动 **libvirtd** 服务。

```
# systemctl start libvirtd
```

验证

1. 确认您的系统已准备好成为虚拟化主机：

```
# virt-host-validate
[...]
QEMU: Checking for device assignment IOMMU support      : PASS
QEMU: Checking if IOMMU is enabled by kernel       : WARN (IOMMU appears to be
disabled in kernel. Add intel_iommu=on to kernel cmdline arguments)
LXC: Checking for Linux >= 2.6.26                  : PASS
[...]
LXC: Checking for cgroup 'blkio' controller mount-point : PASS
LXC: Checking if device /sys/fs/fuse/connections exists : FAIL (Load the 'fuse' module to
enable /proc/ overrides)
```

2. 如果所有 **virt-host-validate** 检查返回 **PASS** 值，则您的系统已准备好 [创建虚拟机](#)。
如果有任何检查返回 **FAIL** 值，请按照显示的说明来修复问题。

如果有任何检查返回 **WARN** 值，请考虑按照显示的说明改进虚拟化功能。

故障排除

- 如果您的主机 CPU 不支持 KVM 虚拟化，**virt-host-validate** 会产生以下输出：

```
QEMU: Checking for hardware virtualization: FAIL (Only emulated CPUs are available,
performance will be significantly limited)
```

但是，在这样的主机系统上的虚拟机将无法引导，而不存在性能问题。

要临时解决这个问题，您可以将虚拟机的 XML 配置中的 **<domain type>** 的值改为 **qemu**。但请注意，红帽不支持使用 **qemu** 域类型的虚拟机，在生产环境中不建议这样做。

2.2. 创建虚拟机

要在 RHEL 8 中创建虚拟机(VM)，请使用 [命令行界面](#) 或 [RHEL 8 web 控制台](#)。

先决条件

- 已在您的系统中[安装并启用](#)虚拟化。
- 您有足够的系统资源来分配给您的虚拟机，如磁盘空间、RAM 或 CPU。根据虚拟机的预期任务和工作负载，推荐的值可能会有很大不同。



警告

RHEL 8 无法从主机 CD-ROM 或者 DVD-ROM 设备安装。如果您在使用 RHEL 8 中的任何 VM 安装方法时选择了 CD-ROM 或 DVD-ROM 作为安装源，则安装会失败。如需更多信息，请参阅[红帽知识库](#)。

2.2.1. 使用命令行界面创建虚拟机

要使用 **virt-install** 工具在 RHEL 8 主机上创建虚拟机(VM)，请按照以下说明操作。

先决条件

- 虚拟化已在您的主机系统中[启用](#)。
- 您有足够的系统资源可以分配给虚拟机，如磁盘空间、RAM 或 CPU。根据虚拟机的预期任务和工作负载，推荐的值可能会有很大不同。
- 操作系统 (OS) 安装源可存在于本地或者网络中。可以是以下之一：
 - 安装介质的 ISO 镜像
 - 现有虚拟机安装的磁盘镜像



警告

在 RHEL 8 中无法从主机 CD-ROM 或者 DVD-ROM 设备安装。当使用 RHEL 8 中的任何虚拟机安装方法时，如果选择了 CD-ROM 或者 DVD-ROM 作为安装源，则安装将失败。如需更多信息，请参阅[红帽知识库](#)。

- 可选：对于快速、简单的配置安装，可以使用 Kickstart 文件。

流程

要创建虚拟机并启动其操作系统安装，请使用 **virt-install** 命令以及以下强制参数：

- 新机器的名称(**--name**)
- 分配的内存量(**--memory**)
- 分配的虚拟 CPU 数量(**--vcpus**)

- 所分配的存储的类型和大小(**--disk**)
- OS 安装源的类型和位置 (**--cdrom** 或 **--location**)

根据所选安装方法，所需选项和值可能会有所不同。请参阅以下的示例：

- 下面创建一个名为 **demo-guest1** 的虚拟机，它从本地存储在 `/home/username/Downloads/Win10install.iso` 文件中的 ISO 镜像安装 Windows 10 OS。此虚拟机还可分配 2048 MiB RAM 和 2 个 vCPU，为虚拟机自动配置 80 GiB qcow2 虚拟磁盘。

```
# virt-install \
--name demo-guest1 --memory 2048 \
--vcpus 2 --disk size=80 --os-variant win10 \
--cdrom /home/username/Downloads/Win10install.iso
```

- 下面创建了一个名为 **demo-guest2** 的虚拟机，该虚拟机使用 `/home/username/Downloads/rhel8.iso` 镜像从实时 CD 运行 RHEL 8 操作系统。没有为这个虚拟机分配磁盘空间，因此在此会话中所做的更改不会被保留。另外，虚拟机被分配 4096 MiB RAM 和 4 个 vCPU。

```
# virt-install \
--name demo-guest2 --memory 4096 --vcpus 4 \
--disk none --livecd --os-variant rhel8.0 \
--cdrom /home/username/Downloads/rhel8.iso
```

- 下面创建一个名为 **demo-guest3** 的 RHEL 8 虚拟机，它连接到现有磁盘镜像 `/home/username/backup/disk.qcow2`。这和在不同的机器间物理地移动硬盘驱动器类似，因此 **demo-guest3** 可用的操作系统和数据由之前处理镜像的方式决定。另外，这个虚拟机还会分配 2048 MiB RAM 和 2 个 vCPU。

```
# virt-install \
--name demo-guest3 --memory 2048 --vcpus 2 \
--os-variant rhel8.0 --import \
--disk /home/username/backup/disk.qcow2
```

请注意，在导入磁盘镜像时，强烈建议使用 **--os-variant** 选项。如果没有提供，创建虚拟机的性能将会受到负面影响。

- 下面创建了一个名为 **demo-guest4** 的虚拟机，该虚拟机可从 `http://example.com/OS-install` URL 安装。要使安装成功启动，URL 必须包含可正常工作的操作系统安装树。另外，操作系统是使用 `/home/username/ks.cfg` kickstart 文件自动配置的。此虚拟机还可分配 2048 MiB RAM、2 个 vCPU 和 160 GiB qcow2 虚拟磁盘。

```
# virt-install \
--name demo-guest4 --memory 2048 --vcpus 2 --disk size=160 \
--os-variant rhel8.0 --location http://example.com/OS-install \
--initrd-inject /home/username/ks.cfg --extra-args="inst.ks=file:/ks.cfg console=tty0 \
console=ttyS0,115200n8"
```

- 下面创建了一个名为 **demo-guest5** 的虚拟机，该虚拟机以纯文本模式从 **RHEL8.iso** 镜像文件安装，而不是图形化安装。它将客户端控制台连接到串行控制台。虚拟机有 16384 MiB 内存、16 个 vCPU 和 280 GiB 磁盘。当通过慢速网络连接连接到主机时这种安装很有用。

```
# virt-install \
--name demo-guest5 --memory 16384 --vcpus 16 --disk size=280 \
```

```
--os-variant rhel8.0 --location RHEL8.iso \
--graphics none --extra-args='console=ttyS0'
```

- 下面创建一个名为 **demo-guest6** 的虚拟机，其配置与 **demo-guest5** 相同，但会位于 10.0.0.1 远程主机上。

```
# virt-install \
--connect qemu+ssh://root@10.0.0.1/system --name demo-guest6 --memory 16384 \
--vcpus 16 --disk size=280 --os-variant rhel8.0 --location RHEL8.iso \
--graphics none --extra-args='console=ttyS0'
```

验证

- 如果成功创建虚拟机，则使用虚拟机的图形控制台打开 **virt-viewer** 窗口并启动客户端操作系统安装。

故障排除

- 如果 **virt-install** 失败，且出现 **cannot find default network** 错误：

- 确定 **libvirt-daemon-config-network** 软件包已安装：

```
# yum info libvirt-daemon-config-network
Installed Packages
Name        : libvirt-daemon-config-network
[...]
```

- 验证 **libvirt** 默认网络是否处于活动状态，并且已配置为自动启动：

```
# virsh net-list --all
Name      State  Autostart  Persistent
-----
default   active  yes       yes
```

- 如果没有，激活默认网络并将其设置为 auto-start：

```
# virsh net-autostart default
Network default marked as autostarted

# virsh net-start default
Network default started
```

- 如果激活默认网络失败并显示以下错误，则代表还没有正确安装 **libvirt-daemon-config-network** 软件包。

```
error: failed to get network 'default'
error: Network not found: no network with matching name 'default'
```

要解决这个问题，请重新安装 **libvirt-daemon-config-network**。

```
# yum reinstall libvirt-daemon-config-network
```

- ii. 如果激活默认网络失败并显示类似如下的错误，则默认网络子网和主机上现有接口之间出现了冲突。

```
error: Failed to start network default  
error: internal error: Network is already in use by interface ens2
```

要解决这个问题，请使用 **virsh net-edit default** 命令，并将配置中的 **192.168.122.*** 值更改为主机上尚未使用的子网。

其它资源

- [man virt-install 命令](#)
- [克隆虚拟机](#)

2.2.2. 使用 web 控制台创建虚拟机并安装客户端操作系统

要在 RHEL 8 主机上的 GUI 中管理虚拟机(VM)，请使用 web 控制台。以下章节提供了如何使用 RHEL 8 web 控制台创建虚拟机并在其上安装客户机操作系统的信息。

2.2.2.1. 使用 web 控制台创建虚拟机

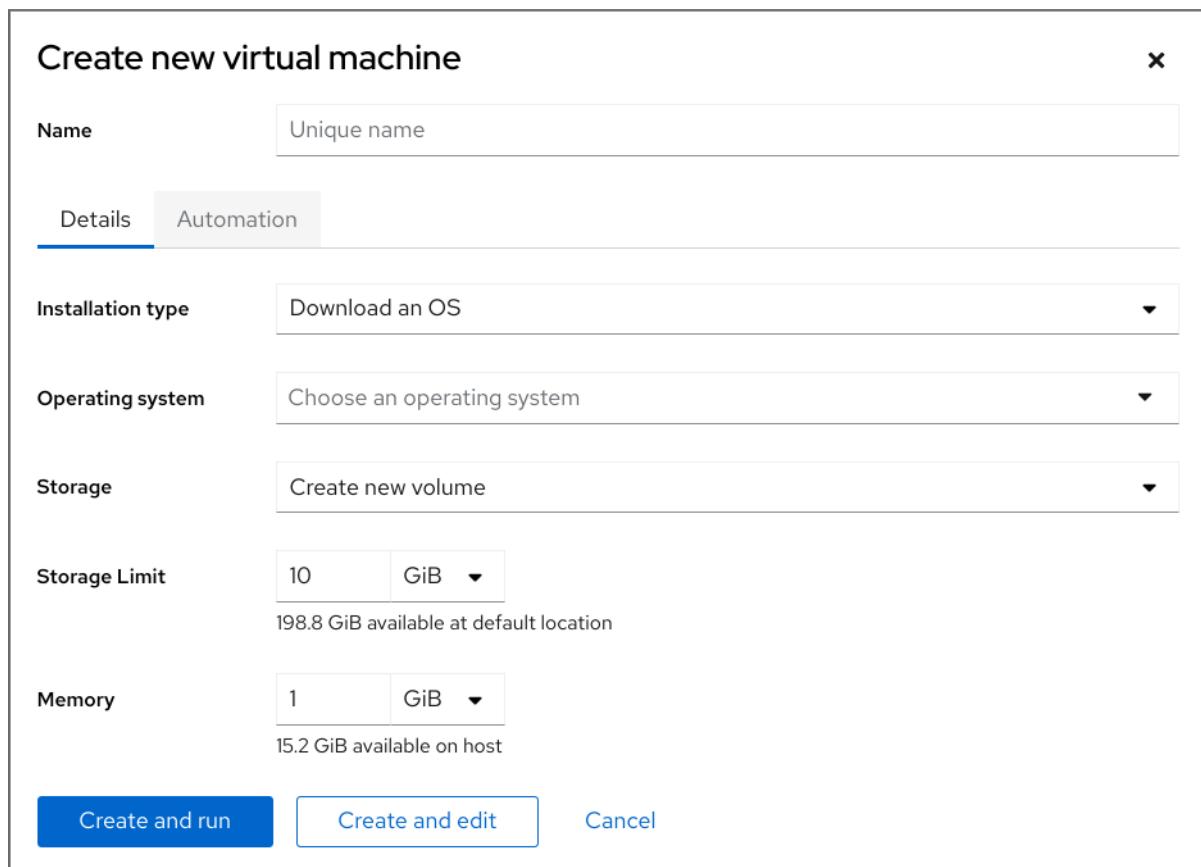
要在连接 web 控制台的主机机器中创建虚拟机（VM），请遵循以下步骤。

先决条件

- 虚拟化已在您的主机系统中启用。
- Web 控制台 VM 插件 [已安装在您的系统上](#)。
- 您有足够的系统资源可以分配给虚拟机，如磁盘空间、RAM 或 CPU。根据虚拟机的预期任务和工作负载，推荐的值可能会有很大不同。

流程

1. 在 web 控制台的 **Virtual Machines** 界面中，点 **Create VM**。
此时会出现 Create new virtual machine 对话框。



2. 输入您要创建的虚拟机的基本配置。

- **Name** - 虚拟机的名称。
- **安装类型** - 安装可以使用本地安装介质、URL、PXE 网络引导、云基础镜像，或者从有限的操作系统集中下载操作系统。
- **Operating system** - 虚拟机的操作系统.请注意，红帽只对 [一组有限的客户机操作系统提供支持](#)。
- **Storage** - 用于配置虚拟机的存储的类型。
- **Storage Limit** - 配置虚拟机的存储空间量。
- **Memory** - 用于配置虚拟机的内存量。

3. 创建虚拟机：

- 如果您希望虚拟机自动安装操作系统，点 **Create 并运行**。
- 如果要在安装操作系统前编辑虚拟机，请点击 **Create 并编辑**。

其它资源

- [在虚拟机上安装操作系统](#)

2.2.2.2. 使用 web 控制台导入磁盘镜像来创建虚拟机

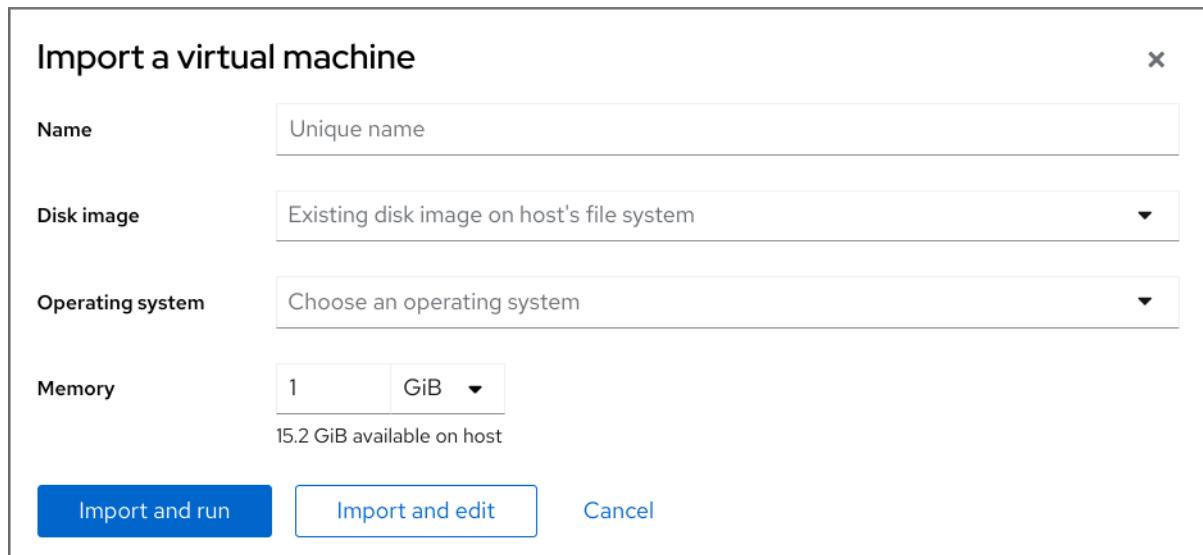
要通过导入现有虚拟机安装的磁盘镜像来创建虚拟机（VM），请遵循以下步骤。

先决条件

- Web 控制台 VM 插件 [已安装在您的系统上](#)。
- 您有足够的系统资源可以分配给虚拟机，如磁盘空间、RAM 或 CPU。根据虚拟机的预期任务和工作负载，推荐的值可能会有很大不同。
- 确保已有虚拟机安装的磁盘镜像

流程

1. 在 web 控制台的 **Virtual Machines** 界面中，单击 **Import VM**。
这时将出现 Import a virtual machine 对话框。



2. 输入您要创建的虚拟机的基本配置。
 - **Name** - 虚拟机的名称。
 - **Disk image** - 主机系统上虚拟机现有磁盘映像的路径。
 - **Operating system** - 虚拟机的操作系统.请注意，红帽只对 [一组有限的客户机操作系统提供支持](#)。
 - **Memory** - 用于配置虚拟机的内存量。
3. 导入虚拟机：
 - 如果您希望虚拟机自动安装操作系统，点 **Import and run**。
 - 如果要在安装操作系统前编辑虚拟机，点 **Import and edit**。

2.2.2.3. 使用 Web 控制台安装客户端操作系统

第一次载入虚拟机（VM）时，您必须在虚拟机上安装操作系统。



注意

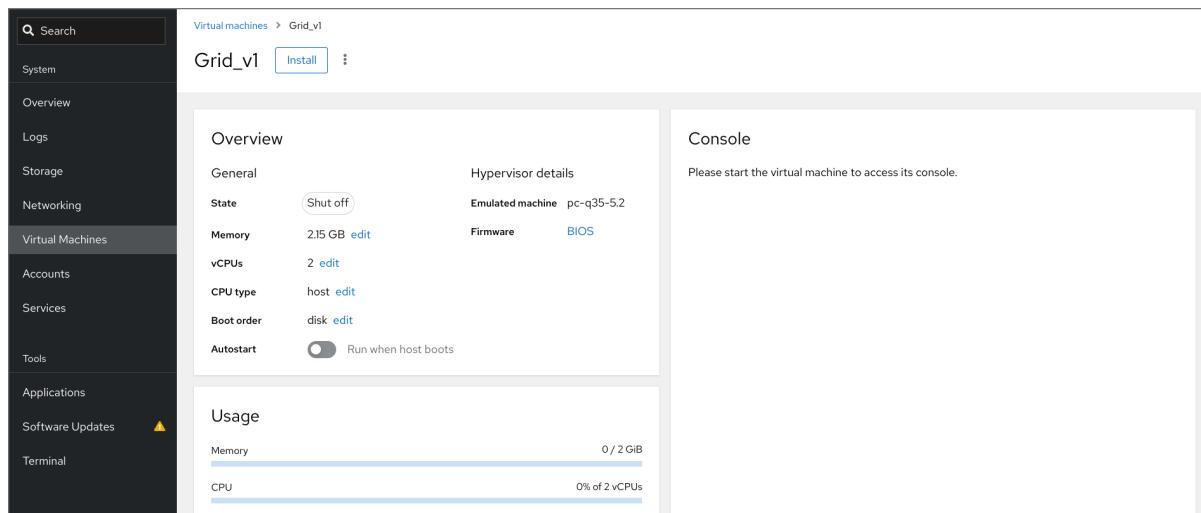
如果您在创建新虚拟机时点 **Create and run** or **Import and run**，则操作系统的安装过程会在创建虚拟机时自动启动。

先决条件

- Web 控制台 VM 插件 已安装在您的系统上。
- 在其中安装操作系统的虚拟机必须可用。

流程

1. 在 **Virtual Machines** 界面中，单击要在其上安装客户机操作系统的虚拟机。
此时将打开一个新页面，其中包含有关所选虚拟机的基本信息，以及管理虚拟机各方面的控制。



2. 可选：修改固件。



注意

只有在选择 **Create and edit** 或 **Import** 并在创建新虚拟机时 编辑了固件，且虚拟机上还没有安装操作系统时，才能更改固件。

- a. 点击固件。
- b. 在 Change Firmware 窗口中，选择所需固件。



- c. 点 Save。

3. 点 **Install**。
在 VM 控制台中运行的操作系统的安装过程。

故障排除

- 如果安装过程失败，则必须删除虚拟机并重新创建。

2.2.3. 通过 web 控制台使用云镜像身份验证创建虚拟机

默认情况下，发行版云镜像没有登录帐户。但是，使用 RHEL web 控制台，您现在可以创建虚拟机(VM)，并指定 root 和用户帐户登录凭证，然后它们被传给 cloud-init。

先决条件

- Web 控制台 VM 插件 [已安装在您的系统上](#)。
- 虚拟化已在您的主机系统上 [启用](#)。
- 您有足够的系统资源来分配给虚拟机，如磁盘空间、RAM 或 CPU。根据虚拟机的预期任务和工作负载，推荐的值可能会有很大不同。

流程

1. 在 web 控制台的 **Virtual Machines** 界面中，点 **Create VM**。
此时会出现 Create new virtual machine 对话框。

The screenshot shows the 'Create new virtual machine' dialog box. It has a 'Details' tab selected. The 'Name' field contains 'Unique name'. Under 'Installation type', 'Download an OS' is selected. Under 'Operating system', 'Choose an operating system' is selected. Under 'Storage', 'Create new volume' is selected. The 'Storage Limit' is set to 10 GiB. Below it, a note says '198.8 GiB available at default location'. The 'Memory' setting is 1 GiB, with a note below stating '15.2 GiB available on host'. At the bottom are three buttons: 'Create and run' (blue), 'Create and edit' (white), and 'Cancel'.

2. 在 **Name** 字段中输入虚拟机的名称。
3. 在 **Details** 标签页中，在 **Installation type** 字段中选择 **Cloud base image**。

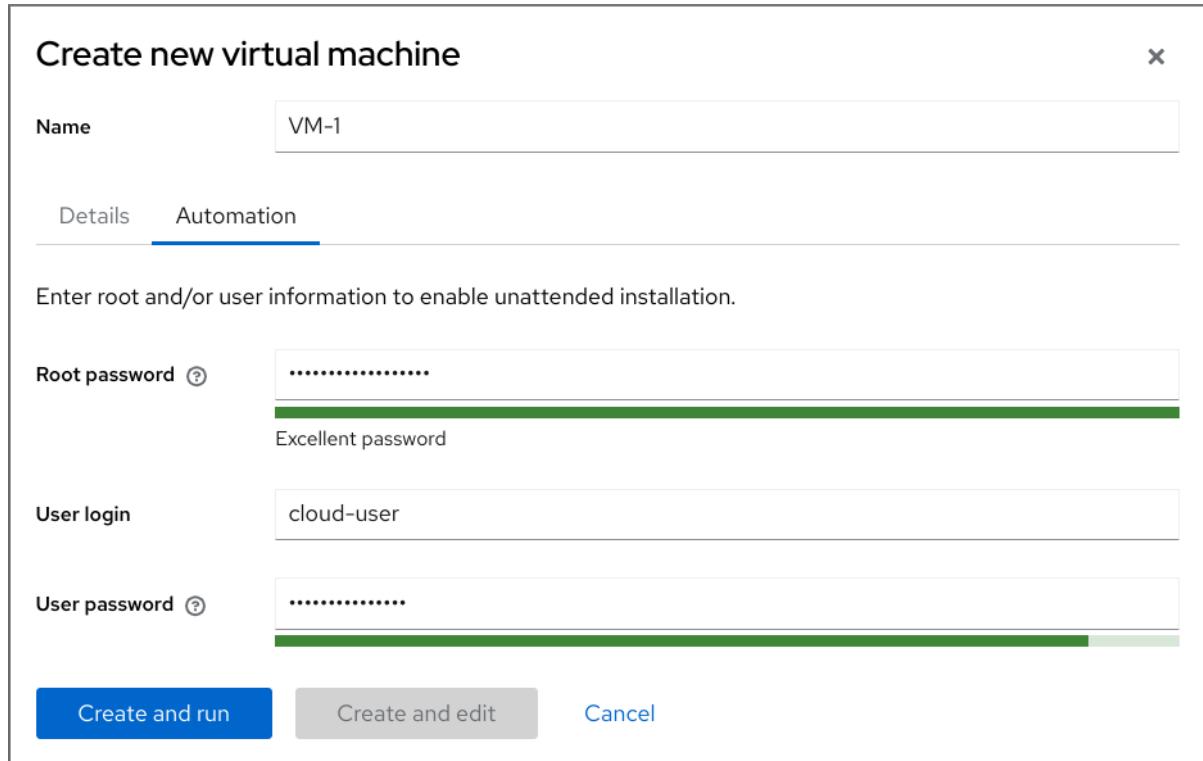
Create new virtual machine

Name	VM-1
Details Automation	
Installation type	Cloud base image
Installation source	/home/test/ X
Operating system	Choose an operating system
Storage	Create new volume
Storage Limit	10 GiB ▼ 198.8 GiB available at default location
Memory	1 GiB ▼ 15.2 GiB available on host
Create and run Create and edit Cancel	

4. 在 **Installation source** 字段中，设置主机系统上镜像文件的路径。
5. 输入您要创建的虚拟机的配置。
 - **Operating system** - 虚拟机的操作系统.请注意，红帽只对它们提供支持

一组有限的客户机操作系统。

- **Storage** - 用于配置虚拟机的存储的类型。
- **Storage Limit** - 配置虚拟机的存储空间量。
- **Memory** - 用于配置虚拟机的内存量。
 1. 点 **Automation** 标签页。
设置云身份验证凭证。
- **Root password** - 输入虚拟机的 root 密码。如果您不想设置 root 密码，请将该字段留空。
- **User login** - 输入 cloud-init 用户登录名。如果您不想创建用户帐户，请将此字段留空。
- **User password** - 输入密码.如果您不想创建用户帐户，请将此字段留空。



1. 点 **Create 并运行。**
虚拟机已创建。

其它资源

- [在虚拟机上安装操作系统](#)

2.3. 启动虚拟机

要在 RHEL 8 中启动虚拟机(VM)，您可以使用 [命令行界面](#) 或 [web 控制台 GUI](#)。

先决条件

- 在启动虚拟机前，它必须被创建，理想情况下，还要使用操作系统进行安装。有关操作的说明，请参阅 [创建虚拟机](#)。

2.3.1. 使用命令行界面启动虚拟机

您可以使用命令行界面(CLI)启动关闭的虚拟机(VM)或恢复保存的虚拟机。使用 CLI，您可以启动本地和远程虚拟机。

先决条件

- 已定义的一个不活跃地虚拟机。
- 虚拟机的名称。
- 对于远程虚拟机：
 - 虚拟机所在主机的 IP 地址。
 - 对主机的 root 访问权限。

流程

- 对于本地虚拟机，请使用 **virsh start** 工具。
例如，以下命令启动 *demo-guest1* 虚拟机。

```
# virsh start demo-guest1
Domain 'demo-guest1' started
```

- 对于位于远程主机上的虚拟机，请使用 **virsh start** 工具以及与主机的 QEMU+SSH 连接。
例如，以下命令在 192.168.123.123 主机上启动 *demo-guest1* 虚拟机。

```
# virsh -c qemu+ssh://root@192.168.123.123/system start demo-guest1
root@192.168.123.123's password:
Domain 'demo-guest1' started
```

其它资源

- virsh start --help** 命令
- [设置对远程虚拟化主机的简单访问](#)
- [当主机启动时自动启动虚拟机](#)

2.3.2. 使用 web 控制台启动虚拟机

如果虚拟机(VM) 处于关闭状态，您可以使用 RHEL 8 web 控制台启动它。您还可以将虚拟机配置为在主机启动时自动启动。

先决条件

- Web 控制台 VM 插件 [已安装在您的系统上](#)。
- 已定义的一个不活跃地虚拟机。
- 虚拟机的名称。

流程

- 在 **Virtual Machines** 界面中，点击您要启动的虚拟机。
此时将打开一个新页面，其中包含有关所选虚拟机的详细信息，以及关闭和删除虚拟机的控制。
- 点 **Run**。
虚拟机启动，您可以[连接到其控制台或图形输出](#)。
- 可选：要将虚拟机配置为在主机启动时自动启动，请切换 **Overview** 部分中的 **Autostart** 复选框。
如果使用不由 libvirt 管理的网络接口，您还必须对 systemd 配置进行额外的更改。否则，受影响的虚拟机可能无法启动，请参阅 [当主机启动时自动启动虚拟机](#)。

其它资源

- [在 web 控制台中关闭虚拟机](#)

- 使用 web 控制台重启虚拟机

2.3.3. 当主机启动时自动启动虚拟机

当一个带有运行的虚拟机(VM)的主机重启时，虚拟机将关闭， 默认必须手动启动。要确保虚拟机在其主机运行时处于活跃状态，您可以将虚拟机配置为自动启动。

先决条件

- 已创建的虚拟机

流程

1. 使用 **virsh autostart** 工具将虚拟机配置为在主机启动时自动启动。

例如，以下命令将 *demo-guest1* 虚拟机配置为自动启动：

```
# virsh autostart demo-guest1
Domain 'demo-guest1' marked as autostarted
```

2. 如果您使用不是由 **libvirt** 管理的网络接口，则也必须对 **systemd** 配置进行额外的更改。否则，受影响的虚拟机可能无法启动。



注意

例如，这些接口包括：

- **NetworkManager** 创建的网桥设备
- 网络配置为使用 `<forward mode='bridge'>`

- a. 在 **systemd** 配置目录树中，如果 **libvirdt.service.d** 目录尚不存在，则创建该目录。

```
# mkdir -p /etc/systemd/system/libvirdt.service.d
```

- b. 在之前创建的目录中创建一个 **10-network-online.conf** **systemd** 单元覆盖文件。此文件的内容覆盖 **libvirdt** 服务的默认 **systemd** 配置。

```
# touch /etc/systemd/system/libvirdt.service.d/10-network-online.conf
```

- c. 将以下行添加到 **10-network-online.conf** 文件中：这个配置更改可确保 **systemd** 仅在主机上的网络就绪后启动 **libvirdt** 服务。

```
[Unit]
After=network-online.target
```

验证

1. 查看虚拟机配置，并检查是否启用了 *autostart* 选项。

例如，以下命令显示有关 *demo-guest1* 虚拟机的基本信息，包括 *autostart* 选项：

```
# virsh dominfo demo-guest1
Id: 2
```

```
Name: demo-guest1
UUID: e46bc81c-74e2-406e-bd7a-67042bae80d1
OS Type: hvm
State: running
CPU(s): 2
CPU time: 385.9s
Max memory: 4194304 KiB
Used memory: 4194304 KiB
Persistent: yes
Autostart: enable
Managed save: no
Security model: selinux
Security DOI: 0
Security label: system_u:system_r:svirt_t:s0:c873,c919 (enforcing)
```

2. 如果您使用不由 libvirt 管理的网络接口, 请检查 **10-network-online.conf** 文件的内容是否与以下输出匹配 :

```
$ cat /etc/systemd/system/libvirtd.service.d/10-network-online.conf
[Unit]
After=network-online.target
```

其它资源

- **virsh autostart --help** 命令
- 使用 web 控制台启动虚拟机。

2.4. 连接至虚拟机

要在 RHEL 8 中与虚拟机 (VM) 交互, 您需要通过以下方法之一连接它 :

- 使用 Web 控制台界面时, 请在 web 控制台界面中使用 Virtual Machines 窗格。如需更多信息, [请参阅使用 web 控制台与虚拟机交互](#)。
- 如果您需要在不使用 Web 控制台的情况下与虚拟机图形显示交互, 请使用 Virt Viewer 应用程序。详情请查看 [使用 Virt Viewer 打开虚拟机图形控制台](#)。
- 如果不需要图形显示, 请使用 [SSH 终端连接](#)。
- 当使用网络无法从您的系统访问虚拟机时, 请使用 [virsh 控制台](#)。

如果您要连接的虚拟机位于远程主机而不是本地主机, 您可以选择配置您的系统[以更方便地访问远程主机](#)。

先决条件

- 已安装并启动您要与之交互的虚拟机。

2.4.1. 使用 web 控制台与虚拟机交互

要在 RHEL 8 web 控制台中与虚拟机 (VM) 交互, 您需要连接到虚拟机的控制台。这包括图形和串行控制台。

- 要在 web 控制台中与虚拟机的图形界面交互, 请使用[图形控制台](#)。

- 要在远程 viewer 中与虚拟机的图形界面交互，请使用[remote viewers 中的图形控制台](#)。
- 要在 web 控制台中与虚拟机的 CLI 交互，请使用[串行控制台](#)。

2.4.1.1. 在 web 控制台中查看虚拟机图形控制台

使用虚拟机（VM）控制台界面，您可以在 RHEL 8 web 控制台中查看所选虚拟机的图形输出。

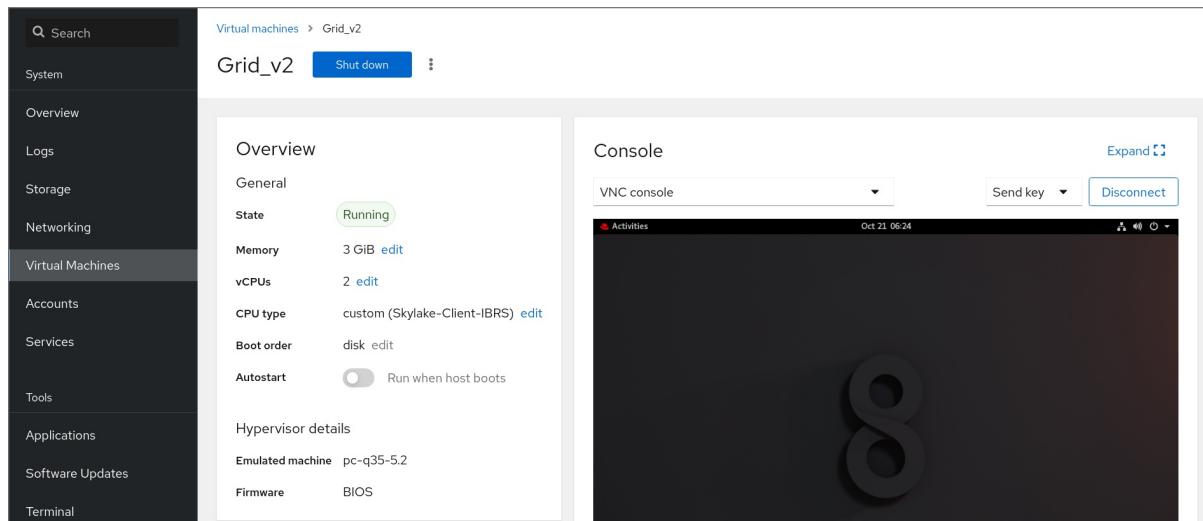
先决条件

- Web 控制台 VM 插件[已安装在您的系统上](#)。
- 确保主机和虚拟机支持图形界面。

流程

- 在 **Virtual Machines** 界面中，单击您要查看其图形控制台的虚拟机。此时将打开一个新页面，其中包含虚拟机的 **Overview** 和 **Console** 部分。
- 在控制台下拉菜单中选择 **VNC 控制台**。VNC 控制台在 Web 界面中的菜单下方显示。

图形控制台会出现在 web 界面中。



3. 单击 **Expand**

现在，您可以使用鼠标和键盘与虚拟机控制台进行交互，其方式与您与真实机器进行交互的方式相同。VM 控制台中的显示反映了虚拟机上正在执行的操作。



注意

运行 web 控制台的主机可能会截获特定的组合键，如 **Ctrl+Alt+Del**，阻止它们发送到虚拟机。

要发送此类组合键，请单击 **Send key** 菜单并选择要发送的键序列。

例如，要将 **Ctrl+Alt+Del** 组合发送给虚拟机，请单击 **Send key** 并选择 **Ctrl+Alt+Del** 菜单条目。

故障排除

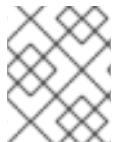
- 如果在图形控制台中点击没有任何效果，请将控制台扩展为全屏。这是一个已知的鼠标光标偏移问题。

其它资源

- [使用 Web 控制台在远程 viewer 中查看图形控制台](#)
- [在 web 控制台中查看虚拟机串口控制台](#)

2.4.1.2. 使用 Web 控制台在远程 viewer 中查看图形控制台

使用 web 控制台界面，您可以在远程查看器（如 Virt Viewer）中显示所选虚拟机(VM)的图形控制台。

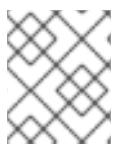


注意

您可以在 web 控制台中启动 Virt Viewer。可以手动启动其他 VNC 和 SPICE remote viewer。

先决条件

- Web 控制台 VM 插件 [已安装在您的系统上](#)。
- 确保主机和虚拟机支持图形界面。
- 在 Virt Viewer 中查看图形控制台前，您必须在 web 控制台连接的机器上安装 Virt Viewer。
 - 单击 **Launch remote viewer**。
virt viewer .vv、文件下载。
 - 打开文件以启动 Virt Viewer。

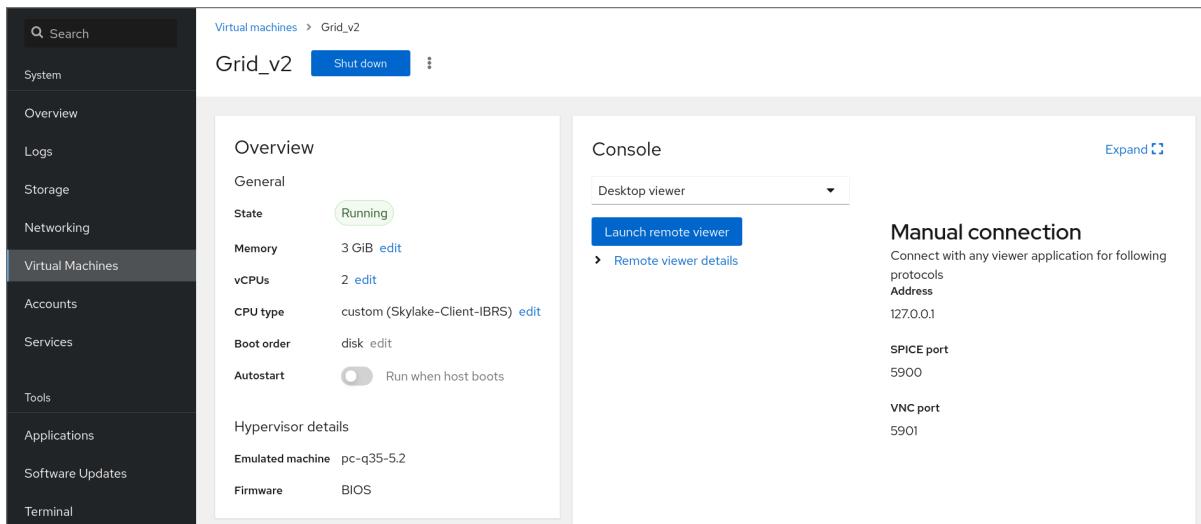


注意

远程查看器在大多数操作系统上提供。但是，一些浏览器扩展和插件不允许 Web 控制台打开 Virt Viewer。

流程

- 在 **Virtual Machines** 界面中，单击您要查看其图形控制台的虚拟机。
此时将打开一个新页面，其中包含虚拟机的 **Overview** 和 **Console** 部分。
- 在控制台下拉菜单中选择 **Desktop Viewer**。



3. 点击 **Launch Remote Viewer**。

图形控制台在 Virt Viewer 中打开。

您可以使用鼠标和键盘与虚拟机控制台以与实际机器交互的方式与 VM 控制台进行交互。VM 控制台中的显示反映了虚拟机上正在执行的操作。



注意

运行 web 控制台的服务器可以截获特定的组合键，如 **Ctrl+Alt+Del**，阻止它们发送到虚拟机。

要发送这样的组合键，点 **Send key** 菜单并选择要使用地键序列。

例如：要将 **Ctrl+Alt+Del** 组合发送到 VM，点 **Send key** 菜单并选择 **Ctrl+Alt+Del** 菜单。

故障排除

- 如果在图形控制台中点击没有任何效果，请将控制台扩展为全屏。这是一个已知的鼠标光标偏移问题。
- 如果在 web 控制台中启动 Remote Viewer 不起作用或者不佳，您可以使用以下协议手动与任何 viewer 应用程序连接：
 - Address** - 默认地址为 **127.0.0.1**。您可以修改 `/etc/libvirt/qemu.conf` 中的 **vnc_listen** 或 **spice_listen** 参数，以将其更改为主机的 IP 地址。
 - SPICE port** - 5900
 - VNC port** - 5901

其它资源

- [在 web 控制台中查看虚拟机图形控制台](#)
- [在 web 控制台中查看虚拟机串口控制台](#)

2.4.1.3. 在 web 控制台中查看虚拟机串口控制台

您可以在 RHEL 8 web 控制台中查看所选虚拟机(VM)的串行控制台。这在主机机器或者虚拟机没有使用图形界面配置时很有用。

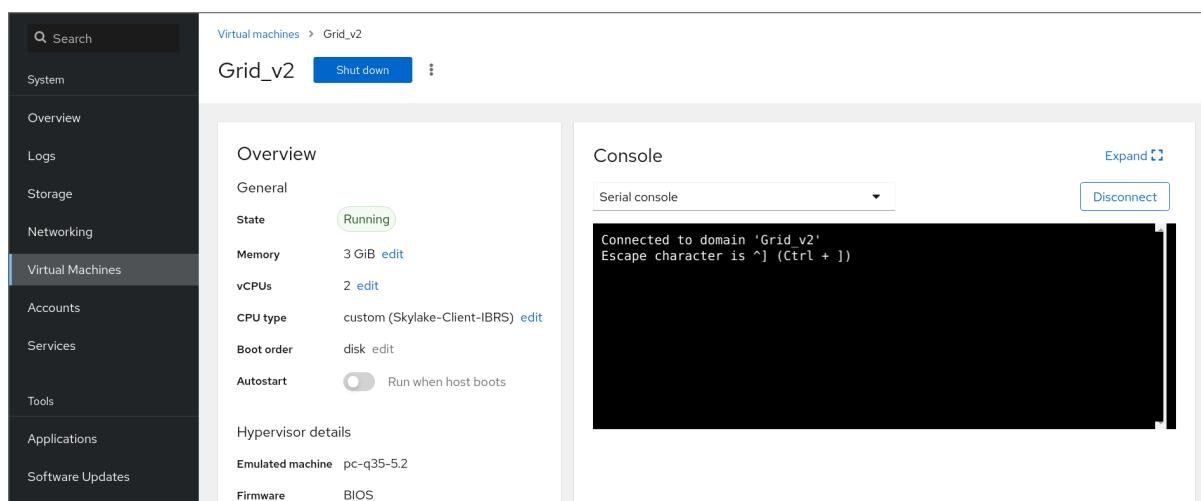
有关串行控制台的更多信息，请参阅 [打开虚拟机串口控制台](#)。

先决条件

- Web 控制台 VM 插件 [已安装在您的系统上](#)。

流程

1. 在 **Virtual Machines** 窗格中，单击您要查看其串行控制台的虚拟机。
此时将打开一个新页面，其中包含虚拟机的 **Overview** 和 **Console** 部分。
2. 在控制台下拉菜单中选择 **Serial console**。
图形控制台会出现在 web 界面中。



您可以断开串行控制台与虚拟机的连接和重新连接。

- 要断开串行控制台与虚拟机的连接，请点 **Disconnect**。
- 要将串行控制台重新连接到虚拟机，请点 **Reconnect**。

其它资源

- 在 web 控制台中 [查看虚拟机图形控制台](#)
- 使用 Web 控制台在 [远程 viewer 中查看图形控制台](#)

2.4.2. 使用 Virt Viewer 打开虚拟机图形控制台

要连接到 KVM 虚拟机(VM)的图形控制台并在 **Virt Viewer** 桌面应用程序中打开它，请按照以下流程操作。

先决条件

- 您的系统以及您要连接的虚拟机必须支持图形显示。
- 如果目标虚拟机位于远程主机上，则需要对主机有连接和 root 访问权限。
- 可选：如果目标虚拟机位于远程主机上，请设置 libvirt 和 SSH [以更方便地访问远程主机](#)。

流程

- 要连接到本地虚拟机，请使用以下命令，并将 *guest-name* 替换为您要连接的虚拟机的名称：

```
# virt-viewer guest-name
```

- 要连接到远程虚拟机，请使用 **virt-viewer** 命令及 SSH 协议。例如，以下命令以 root 身份连接到位于远程系统 10.0.0.1 的名为 *guest-name* 的虚拟机。此连接还需要为 10.0.0.1 进行 root 身份验证。

```
# virt-viewer --direct --connect qemu+ssh://root@10.0.0.1/system guest-name
root@10.0.0.1's password:
```

验证

如果连接正常工作，则虚拟机将显示在 **Virt Viewer** 窗口中。

您可以使用鼠标和键盘与虚拟机控制台以与实际机器交互的方式与 VM 控制台进行交互。VM 控制台中的显示反映了虚拟机上正在执行的操作。

故障排除

- 如果在图形控制台中点击没有任何效果，请将控制台扩展为全屏。这是一个已知的鼠标光标偏移问题。

其它资源

- virt-viewer** 手册页
- [设置对远程虚拟化主机的简单访问](#)
- [使用 web 控制台与虚拟机交互](#)

2.4.3. 使用 SSH 连接到虚拟机

要使用 SSH 连接协议与虚拟机(VM)的终端交互，请遵循以下步骤。

先决条件

- 有对目标虚拟机的网络连接和 root 访问权限。
- 如果目标虚拟机位于远程主机上，您也可以拥有对该主机的连接和 root 访问权限。
- 您的虚拟机网络由 **libvirt** 生成的 **dnsmasq** 分配 IP 地址。这是 **libvirt NAT 网络** 中的示例。值得注意的是，如果虚拟机使用以下网络配置之一，则无法使用 SSH 连接到虚拟机：
 - **hostdev** 接口
 - 直接接口
 - 网桥间
- 在虚拟机主机上已安装并启用了 **libvirt-nss** 组件。如果没有，请执行以下操作：
 - 安装 **libvirt-nss** 软件包：

```
# yum install libvirt-nss
```

- b. 编辑 `/etc/nsswitch.conf` 文件，并将 `libvirt_guest` 添加到 `hosts` 行中：

```
[...]
passwd: compat
shadow: compat
group: compat
hosts: files libvirt_guest dns
[...]
```

流程

- 当连接到远程虚拟机时，请首先 SSH 到其物理主机。以下示例演示了使用其 root 凭证连接到主机 10.0.0.1：

```
# ssh root@10.0.0.1
root@10.0.0.1's password:
Last login: Mon Sep 24 12:05:36 2021
root~#
```

- 使用虚拟机的名称和用户访问凭证来连接它。例如，以下命令使用其 root 凭证连接到 `testguest1` 虚拟机：

```
# ssh root@testguest1
root@testguest1's password:
Last login: Wed Sep 12 12:05:36 2018
root~]#
```

故障排除

- 如果您不知道虚拟机的名称，可以使用 `virsh list --all` 命令列出主机上所有可用的虚拟机：

```
# virsh list --all
 Id  Name          State
 -----
 2   testguest1    running
 -   testguest2    shut off
```

其它资源

- [上游 libvirt 文档](#)

2.4.4. 打开虚拟机串口控制台

使用 `virsh console` 命令，可以连接到虚拟机的串行控制台(VM)。

但虚拟机有以下情况时很有用：

- 没有提供 VNC 或 SPICE 协议，因此没有为 GUI 工具提供视频显示。
- 没有网络连接，因此无法使用 SSH 进行交互。

先决条件

- 虚拟机必须配有一个串口控制台设备，如 **console type='pty'**。要进行验证，请执行以下操作：

```
# virsh dumpxml vm-name | grep console
<console type='pty' tty='/dev/pts/2'>
</console>
```

- 虚拟机必须在内核命令行中配置串口控制台。要进行验证，VM 上的 **cat /proc/cmdline** 命令输出应包含 **console=ttyS0**。例如：

```
# cat /proc/cmdline
BOOT_IMAGE=/vmlinuz-3.10.0-948.el7.x86_64 root=/dev/mapper/rhel-root ro console=tty0
console=ttyS0,9600n8 rd.lvm.lv=rhel/root rd.lvm.lv=rhel/swap rhgb
```

如果没有在虚拟机中正确设置串口控制台，请使用 **virsh** 控制台连接到虚拟机，请将您连接到无响应的客户端控制台。但是，您仍然可以使用 **Ctrl+]** 快捷键退出无响应控制台。

- 要在虚拟机上设置串行控制台，请执行以下操作：

- 在虚拟机上启用 **console=ttyS0** 内核选项：

```
# grubby --update-kernel=ALL --args="console=ttyS0"
```

- 清除可能会阻止您更改生效的内核选项。

```
# grub2-editenv - unset kernelopts
```

- 重启虚拟机。

流程

1. 在您的主机系统上，使用 **virsh console** 命令。如果 libvirt 驱动程序支持安全控制台处理，以下示例连接到 *guest1* 虚拟机：

```
# virsh console guest1 --safe
Connected to domain 'guest1'
Escape character is ^]

Subscription-name
Kernel 3.10.0-948.el7.x86_64 on an x86_64

localhost login:
```

2. 您还可以使用与标准命令行界面相同的方法与 virsh 控制台互动。

其它资源

- **virsh** 手册页

2.4.5. 设置对远程虚拟化主机的简单访问

当使用 libvirt 工具在远程主机系统上管理虚拟机时，建议使用 **-c qemu+ssh://root@hostname/system** 语法。例如，要在 10.0.0.1 主机上以 root 用户身份使用 **virsh list** 命令：

```
# virsh -c qemu+ssh://root@10.0.0.1/system list
```

root@10.0.0.1's password:

Id	Name	State
1	remote-guest	running

然而，为方便起见，您可以通过修改 SSH 和 libvirt 配置来完全删除需要指定连接详情。例如，您可以：

```
# virsh -c remote-host list
```

root@10.0.0.1's password:

Id	Name	State
1	remote-guest	running

要进行改进，请按照以下步骤操作。

流程

1. 编辑或创建 `~/.ssh/config` 文件，并添加以下内容，其中 `host-alias` 是与特定的远程主机相关联的短名称，`hosturl` 是主机的 URL 地址。

```
Host host-alias
  User      root
  Hostname  hosturl
```

例如，下面的命令为 `root@10.0.0.1` 设置 `tyrannosaurus` 别名：

```
Host tyrannosaurus
  User      root
  Hostname  10.0.0.1
```

2. 编辑或创建 `/etc/libvirt/libvirt.conf` 文件，并添加以下内容，其中 `qemu-host-alias` 是 QEMU 和 libvirt 工具将与预期主机关联的主机别名：

```
uri_aliases = [
  "qemu-host-alias=qemu+ssh://host-alias/system",
]
```

例如，以下命令使用上一步中配置的 `tyrannosaurus` 别名来设置 `t-rex` 别名，它代表 `qemu+ssh://10.0.0.1/system`：

```
uri_aliases = [
  "t-rex=qemu+ssh://tyrannosaurus/system",
]
```

验证

1. 确认通过在本地系统上使用基于 libvirt 的工具，并添加 `-c qemu-host-alias` 参数，您可以来管理远程虚拟机。这会在远程主机中使用 SSH 自动执行命令。

例如，验证以下列出的在 10.0.0.1 远程主机上的虚拟机，在上一步中将连接设为 `t-rex`：

```
$ virsh -c t-rex list
root@10.0.0.1's password:

Id  Name      State
-----
1   velociraptor    running
```



注意

除了 `virsh` 之外，`-c`（或 `--connect`）选项以及上述远程主机访问配置也可以由以下工具使用：

- [virt-install](#)
- [virt-viewer](#)

后续步骤

- 如果要在单一远程主机中只使用 libvirt 工具，您也可以为基于 libvirt 的实用程序设置特定的连接作为默认目标。为此，请编辑 `/etc/libvirt/libvirt.conf` 文件，并将 `uri_default` 参数的值设为 `qemu-host-alias`。例如：以下命令使用在前面的步骤中设置的 `t-rex` 主机别名作为默认的 libvirt 目标。

```
# These can be used in cases when no URI is supplied by the application
# (@uri_default also prevents probing of the hypervisor driver).
#
uri_default = "t-rex"
```

因此，所有基于 libvirt 的命令都会在指定的远程主机中自动执行。

```
$ virsh list
root@10.0.0.1's password:

Id  Name      State
-----
1   velociraptor    running
```

但是，如果您也要管理本地主机或不同远程主机上的虚拟机，则不建议这样做。

- 当连接到远程主机时，您可以避免向远程系统提供 root 密码。要做到这一点，请使用以下一个或多个方法：
 - [设置对远程主机的基于密钥的 SSH 访问](#)
 - 使用 SSH 连接多路复用连接到远程系统
 - [身份管理中的 Kerberos 身份验证](#)
- `-c`（或 `--connect`）选项可用于在远程主机上运行 `virt-install`、`virt-viewer`、`virsh` 和 `virt-manager` 命令。

2.5. 关闭虚拟机

要关闭托管于 RHEL 8 上的正在运行的虚拟机，请使用 [命令行界面](#) 或 [Web 控制台 GUI](#)。

2.5.1. 使用命令行界面关闭虚拟机

要关闭响应的虚拟机（VM），请执行以下操作之一：

- 在[连接到客户端](#)时，使用适合客户端操作系统的 shutdown 命令。
- 在主机上使用 **virsh shutdown** 命令：
 - 如果虚拟机位于本地主机上：

```
# virsh shutdown demo-guest1
Domain 'demo-guest1' is being shutdown
```

- 如果虚拟机位于远程主机上，在本例中为 10.0.0.1:

```
# virsh -c qemu+ssh://root@10.0.0.1/system shutdown demo-guest1
root@10.0.0.1's password:
Domain 'demo-guest1' is being shutdown
```

要强制虚拟机关机（例如，如果其机已变得无响应），请在主机上使用 **virsh destroy** 命令：

```
# virsh destroy demo-guest1
Domain 'demo-guest1' destroyed
```



注意

virsh destroy 命令实际上不会删除虚拟机配置或磁盘镜像。它只会终止虚拟机的正在运行的虚拟机实例，类似于从物理机拔掉电源。因此，在个别情况下，**virsh destroy** 可能会导致虚拟机文件系统崩溃，因此仅在所有其他关闭方法都失败时才建议使用这个命令。

2.5.2. 使用 web 控制台关闭和重启虚拟机

使用 RHEL 8 web 控制台，您可以[关闭](#)或[重启](#)正在运行的虚拟机。您还可以向无响应的虚拟机发送不可屏蔽中断。

2.5.2.1. 在 web 控制台中关闭虚拟机

如果虚拟机(VM)处于 **running** 状态，您可以使用 RHEL 8 web 控制台关闭它。

先决条件

- Web 控制台 VM 插件[已安装在您的系统上](#)。

流程

1. 在 **Virtual Machines** 接口中，找到您要关闭的虚拟机行。
2. 在行的右侧，点击 **Shut Down**。

虚拟机关机。

故障排除

- 如果虚拟机没有关闭，请点 **Shut Down** 按钮旁边的 Menu : 按钮，然后选择 **Force Shut Down**。
- 要关闭无响应虚拟机，您还可以[发送一个不可屏蔽的中断](#)。

其它资源

- [使用 web 控制台启动虚拟机](#)
- [使用 web 控制台重启虚拟机](#)

2.5.2.2. 使用 web 控制台重启虚拟机

如果虚拟机(VM)处于 **running** 状态，您可以使用 RHEL 8 web 控制台重启它。

先决条件

- Web 控制台 VM 插件 [已安装在您的系统上](#)。

流程

- 在 **Virtual Machines** 接口中，找到您要重启的虚拟机行。
- 在行的右侧，点击菜单按钮 :。此时会显示一个下拉操作菜单。
- 在下拉菜单中，单击 **Reboot**。
虚拟机将关机并重启。

故障排除

- 如果虚拟机没有重启，点 **Reboot** 按钮旁边的 Menu : 按钮，然后选择 **Force Reboot**。
- 要关闭无响应虚拟机，您还可以[发送一个不可屏蔽的中断](#)。

其它资源

- [使用 web 控制台启动虚拟机](#)
- [在 web 控制台中关闭虚拟机](#)

2.5.2.3. 使用 web 控制台向虚拟机发送不可屏蔽中断

发送不可屏蔽中断（NMI）可能会导致无响应运行的虚拟机（VM）响应或关闭。例如，您可以将 **Ctrl+Alt+Del** NMI 发送到不响应标准输入的虚拟机。

先决条件

- Web 控制台 VM 插件 [已安装在您的系统上](#)。

流程

1. 在 **Virtual Machines** 接口中，找到您要给其发送 NMI 的虚拟机行。
2. 在行的右侧，点击菜单按钮 。此时会显示一个下拉操作菜单。
3. 在下拉菜单中，点击 **Send non-maskable interrupt**。一个 NMI 发送到虚拟机。

其它资源

- [使用 web 控制台启动虚拟机](#)
- [使用 web 控制台重启虚拟机](#)
- [在 web 控制台中关闭虚拟机](#)

2.6. 删除虚拟机

要在 RHEL 8 中删除虚拟机，请使用 [命令行界面](#) 或 [web 控制台 GUI](#)。

2.6.1. 使用命令行界面删除虚拟机

要删除虚拟机（VM），您可以使用命令行从主机中删除它的 XML 配置和相关存储文件。按照以下步骤操作：

先决条件

- 备份虚拟机中的重要数据。
- 关闭虚拟机。
- 确保没有其他虚拟机使用相同的关联的存储。

流程

- 使用 **virsh undefine** 工具。
例如：以下命令删除 *guest1* 虚拟机、与其关联的存储卷以及非电压 RAM（若有）。

```
# virsh undefine guest1 --remove-all-storage --nvram
Domain 'guest1' has been undefined
Volume 'vda'(/home/images/guest1.qcow2) removed.
```

其它资源

- [virsh undefine --help 命令](#)
- [virsh 手册页](#)

2.6.2. 使用 web 控制台删除虚拟机

要从 RHEL 8 web 控制台连接的主机中删除虚拟机（VM）及其关联的存储文件，请按照以下步骤操作：

先决条件

- Web 控制台 VM 插件 [已安装在您的系统上。](#)
- 备份虚拟机中的重要数据。
- 确保没有其他虚拟机使用相同的关联存储。
- 可选：关闭虚拟机。

流程

1. 在**虚拟机**界面中，点击您要删除的虚拟机的 Menu 按钮 。此时会出现一个下拉菜单，控制各种虚拟机操作。
2. 点击 **Delete**。此时会出现确认对话框。



3. 可选：要删除与虚拟机关联的所有或部分存储文件，请选择您要删除的存储文件旁的复选框。
4. 点击 **Delete**。虚拟机和任何选择的存储文件都将被删除。

2.7. 其它资源

- [在 IBM POWER 中使用虚拟化](#)
- [在 IBM Z 中使用虚拟化](#)

第3章 在 IBM POWER 中使用虚拟化

当在 IBM POWER8 或 POWER9 硬件上使用 RHEL 8 时，您可以使用 KVM 虚拟化。但是，与 AMD64 和 Intel64 构架中的虚拟化相比，在您的系统中[启用 KVM 虚拟机管理程序](#)需要额外的步骤。某些 RHEL 8 虚拟化功能在 IBM POWER 上也有[不同或受限的功能](#)。

除以下部分中的信息外，在 IBM POWER 中使用虚拟化与 AMD64 和 Intel64 相同。因此，在 IBM POWER 上使用虚拟化时，您可以查看其他 RHEL 8 虚拟化文档。

3.1. 在 IBM POWER 中启用虚拟化

要在运行 RHEL 8 的 IBM POWER8 或 IBM POWER9 系统上设置 KVM hypervisor 并创建虚拟机(VM)，请按照以下说明操作。

先决条件

- 在您的主机上已安装并注册了 RHEL 8。
- 有以下最小系统资源可用：
 - 主机有 6 GB 可用磁盘空间，以及每个预期的虚拟机需要额外 6 GB 空间。
 - 主机需要 2 GB RAM，以及每个预期的虚拟机需要额外 2 GB。
 - 主机上有 4 个 CPU。虚拟机通常可以使用单个分配的 vCPU 运行，但红帽建议为每个虚拟机分配 2 个或更多 vCPU，以避免虚拟机在高负载期间变得无响应。
- 您的 CPU 机器类型必须支持 IBM POWER 虚拟化。
要进行验证，请在 `/proc/cpuinfo` 文件中查询平台信息。

```
# grep ^platform /proc/cpuinfo/
platform      : PowerNV
```

如果这个命令的输出包含 **PowerNV** 条目，则您正在运行 PowerNV 机器类型，并可在 IBM POWER 上使用虚拟化。

流程

1. 载入 KVM-HV 内核模块

```
# modprobe kvm_hv
```

2. 验证是否载入了 KVM 内核模块

```
# lsmod | grep kvm
```

如果 KVM 成功加载，这个命令的输出会包括 **kvm_hv**。

3. 在虚拟化模块中安装软件包：

```
# yum module install virt
```

4. 安装 **virt-install** 软件包：

```
# yum install virt-install
```

- 启动 **libvirtd** 服务。

```
# systemctl start libvirtd
```

验证

- 确认您的系统已准备好成为虚拟化主机：

```
# virt-host-validate
[...]
QEMU: Checking if device /dev/vhost-net exists      : PASS
QEMU: Checking if device /dev/net/tun exists        : PASS
QEMU: Checking for cgroup 'memory' controller support : PASS
QEMU: Checking for cgroup 'memory' controller mount-point : PASS
[...]
QEMU: Checking for cgroup 'blkio' controller support   : PASS
QEMU: Checking for cgroup 'blkio' controller mount-point : PASS
QEMU: Checking if IOMMU is enabled by kernel       : PASS
```

- 如果所有 **virt-host-validate** 检查返回 **PASS** 值，则您的系统已准备好 [创建虚拟机](#)。如果有任何检查返回 **FAIL** 值，请按照显示的说明来修复问题。

如果有任何检查返回 **WARN** 值，请考虑按照显示的说明改进虚拟化功能。

故障排除

- 如果您的主机 CPU 不支持 KVM 虚拟化，**virt-host-validate** 会产生以下输出：

```
QEMU: Checking for hardware virtualization: FAIL (Only emulated CPUs are available,
performance will be significantly limited)
```

但是，在这样的主机系统上的虚拟机将无法引导，而不存在性能问题。

要临时解决这个问题，您可以将虚拟机的 XML 配置中的 **<domain type>** 的值改为 **qemu**。但请注意，红帽不支持使用 **qemu** 域类型的虚拟机，在生产环境中不建议这样做。

3.2. IBM POWER 中的虚拟化与 AMD64 和 INTEL 64 的不同

IBM POWER 系统上 RHEL 8 中的 KVM 虚拟化在许多方面与 AMD64 和 Intel 64 系统中的 KVM 不同，特别是：

内存要求

IBM POWER 上的虚拟机消耗更多内存。因此，推荐在 IBM POWER 主机上为虚拟机（VM）分配最小内存为 2GB RAM。

显示协议

IBM POWER 系统不支持 SPICE 协议。要显示虚拟机的图形化输出，请使用 **VNC** 协议。另外，只支持以下虚拟图形卡设备：

- vga** - 仅在 **-vga std** 模式中支持，在 **-vga cirrus** 模式中不支持。
- virtio-vga**

- **virtio-gpu**

SMBIOOS

SMBIOS 配置不可用。

内存分配错误

POWER8 虚拟机（包括兼容模式虚拟机）可能会因为类似如下的错误而失败：

```
qemu-kvm: Failed to allocate KVM HPT of order 33 (try smaller maxmem?): Cannot allocate
memory
```

这在使用 RHEL 7.3 及之前作为虚拟机操作系统的虚拟机上发生的可能性更大。

要解决这个问题，请通过将 **kvm_cma_resv_ratio= memory** 添加到主机的内核命令行来增加客户机哈希页表(HPT)的可用 CMA 内存池，其中 *memory* 是应为 CMA 池保留的主机内存的百分比（默认为 5）。

巨页

透明巨页 (THP) 在 IBM POWER8 虚拟机上不提供任何显著的性能优势。但是 IBM POWER9 虚拟机可以如预期从 THP 中受益。

另外，IBM POWER8 系统中静态巨页的大小为 16 MiB 和 16 GiB，而 AMD64、Intel 64 和 IBM POWER9 上的 2 MiB 和 1GiB 不同。因此，要将配置有静态巨页的虚拟机从 IBM POWER8 主机迁移 到 IBM POWER9 主机，您必须首先在虚拟机上设置 1GiB 巨页。

kvm-clock

不需要为 IBM POWER9 上的虚拟机的时间管理配置 **kvm-clock** 服务。

pvpnac

IBM POWER9 系统不支持 **pvpnac** 设备。但是在这个构架中默认使用并激活等同的功能。要在虚拟机中启用它，请使用带有 **preserve** 值的 **<on_crash>** XML 配置元素。

另外，请确保从 **<devices>** 部分中删除 **<panic>** 元素，因为它的存在可能会导致虚拟机无法在 IBM POWER 系统上引导。

单线程主机

在 IBM POWER8 系统中，主机机器必须以 **单线程模式** 运行，才能支持虚拟机。如果安装了 *qemu-kvm* 软件包，则会自动进行配置。但是，在单线程主机上运行的虚拟机仍然可以使用多个线程。

外设设备

IBM POWER 系统中不支持在 AMD64 和 Intel 64 系统上支持很多外设设备，或者支持不同的设备替换。

- 不支持用于 PCI-E 层次结构的设备，包括 **ioh3420** 和 **xio3130-downstream**。此功能由 **spapr-pci-host-bridge** 设备提供的多个独立的 PCI root 网桥所取代。
- 不支持 UHCI 和 EHCI PCI 控制器。使用 OHCI 和 XHCI 控制器。
- 不支持 IDE 设备，包括虚拟 IDE CD-ROM(**ide-cd**)和虚拟 IDE 磁盘(**ide-hd**)。改为使用 **virtio-scsi** 和 **virtio-blk** 设备。
- 不支持仿真 PCI NIC (**rtl8139**)。改为使用 **virtio-net** 设备。
- 不支持声音设备，包括 **intel-hda**、**hda-output** 和 **AC97**。
- 不支持 USB 重定向设备，包括 **usb-redir** 和 **usb-tablet**。

v2v 和 p2v

virt-v2v 和 **virt-p2v** 工具仅在 AMD64 和 Intel 64 构架上被支持，且在 IBM POWER 上不提供。

其他资源

- 有关在红帽支持的系统构架中选择支持的和不支持的虚拟化功能的比较，请参阅 [RHEL 8 支持的虚拟化功能概述](#)。

第 4 章 在 IBM Z 中使用虚拟化

在 IBM Z 硬件上使用 RHEL 8 时，您可以使用 KVM 虚拟化。但是，与 AMD64 和 Intel 64 构架中的虚拟化相比，在您的系统中[启用 KVM 虚拟机管理程序](#)需要额外的步骤。某些 RHEL 8 虚拟化功能在 IBM Z 上也具有[不同的或受限的功能](#)。

除以下部分中的信息外，在 IBM Z 中使用虚拟化与 AMD64 和 Intel64 相同。因此，在使用 IBM Z 上的虚拟化时，您可以查看其他 RHEL 8 虚拟化文档以了解更多信息。



注意

不支持在 z/VM 操作系统上运行 KVM。

4.1. 在 IBM Z 中启用虚拟化

要在运行 RHEL 8 的 IBM Z 系统上设置 KVM hypervisor 并创建虚拟机(VM)，请按照以下说明操作。

先决条件

- RHEL 8.6 或更高版本已安装并在主机上已注册。



重要

如果您已在使用 RHEL 8.5 或更早版本的 IBM Z 机器上启用了虚拟化，您应该重新配置虚拟化模块并更新您的系统。具体步骤请查看[IBM Z 中的虚拟化与 AMD64 和 Intel 64 不同](#)。

- 有以下最小系统资源可用：
 - 主机有 6 GB 可用磁盘空间，以及每个预期的虚拟机需要额外 6 GB 空间。
 - 主机需要 2 GB RAM，以及每个预期的虚拟机需要额外 2 GB。
 - 主机上有 4 个 CPU。虚拟机通常可以使用单个分配的 vCPU 运行，但红帽建议为每个虚拟机分配 2 个或更多 vCPU，以避免虚拟机在高负载期间变得无响应。
- 您的 IBM Z 主机系统使用 z13 CPU 或更高版本。
- RHEL 8 安装在逻辑分区(LPAR)上。另外，LPAR 支持[启动阶段执行 \(SIE\)](#) 虚拟化功能。要进行验证，请在 `/proc/cpuinfo` 文件中搜索 `sie`。

```
# grep sie /proc/cpuinfo/
features      : esan3 zarch stle msa ldisp eimm dfp edat etf3eh highgprs te sie
```

流程

1. 载入 KVM 内核模块：

```
# modprobe kvm
```

2. 验证是否载入了 KVM 内核模块：

```
# lsmod | grep kvm
```

如果 KVM 成功加载，这个命令的输出包括会 **kvm**。

3. 在 **virt:rhel/common** 模块中安装软件包：

```
# yum module install virt:rhel/common
```

4. 启动虚拟化服务：

```
# for drv in qemu network nodedev nwfilter secret storage interface; do systemctl start virt${drv}d{,-ro,-admin}.socket; done
```

验证

1. 验证您的系统已准备好成为虚拟化主机。

```
# virt-host-validate
[...]
QEMU: Checking if device /dev/kvm is accessible : PASS
QEMU: Checking if device /dev/vhost-net exists : PASS
QEMU: Checking if device /dev/net/tun exists : PASS
QEMU: Checking for cgroup 'memory' controller support : PASS
QEMU: Checking for cgroup 'memory' controller mount-point : PASS
[...]
```

2. 如果所有 **virt-host-validate** 检查返回 **PASS** 值，则您的系统已准备好 [创建虚拟机](#)。如果有任何检查返回 **FAIL** 值，请按照显示的说明来修复问题。

如果有任何检查返回 **WARN** 值，请考虑按照显示的说明改进虚拟化功能。

故障排除

- 如果您的主机 CPU 不支持 KVM 虚拟化，**virt-host-validate** 会产生以下输出：

```
QEMU: Checking for hardware virtualization: FAIL (Only emulated CPUs are available,
performance will be significantly limited)
```

但是，在这样的主机系统上的虚拟机将无法引导，而不存在性能问题。

要临时解决这个问题，您可以将虚拟机的 XML 配置中的 **<domain type>** 的值改为 **qemu**。但请注意，红帽不支持使用 **qemu** 域类型的虚拟机，在生产环境中不建议这样做。

4.2. 将 IBM Z 上的虚拟化从 RHEL 8.5 更新至 RHEL 8.6 或更高版本

如果您在 RHEL 8.6 之前的 IBM Z 硬件上安装 RHEL 8，则需要从 AV 流获取虚拟化 RPM，与 RHEL 8 基本 RPM 流分开。从 RHEL 8.6 开始，以前从 AV 流提供的虚拟化 RPM 只在基本 RHEL 流中可用。另外，在以后的 RHEL 8 次要发行本中将停用 AV 流。因此，不再支持使用 AV 流。

按照下面的说明，您将取消激活您的 AV 流，并允许访问 RHEL 8.6 及更高版本中提供的虚拟化 RPM。

先决条件

- 您会在 IBM Z 上使用 RHEL 8.5，安装 **virt:av** 模块。确认情况确实如此：

```
# hostnamectl | grep "Operating System"
Operating System: Red Hat Enterprise Linux 8.5 (Ootpa)
# yum module list --installed
[...]
Advanced Virtualization for RHEL 8 IBM Z Systems (RPMs)
Name           Stream          Profiles       Summary
virt           av [e]        common [i]    Virtualization module
```

流程

- 禁用 **virt:av** 模块。

```
# yum disable virt:av
```

- 删除您系统上已经包含的预先存在的虚拟化软件包和模块。

```
# yum module reset virt -y
```

- 将软件包升级到最新的 RHEL 版本。

```
# yum update
```

这也会在系统上自动启用 **virt:rhel** 模块。

验证

- 确定您的系统中的 **virt** 模块由 **rhel** 流提供。

```
# yum module info virt
```

Name	:	virt
Stream	:	rhel [d][e][a]
Version	:	8050020211203195115
[...]		

其它资源

- [IBM Z 上的虚拟化与 AMD64 和 Intel 64 的不同](#)

4.3. IBM Z 上的虚拟化与 AMD64 和 INTEL 64 有什么不同

IBM Z 系统上 RHEL 8 中的 KVM 虚拟化与 AMD64 和 Intel 64 系统上的 KVM 在以下方面有所不同：

PCI 和 USB 设备

IBM Z 不支持虚拟 PCI 和 USB 设备。这意味着不支持 **virtio-*pci** 设备，应该改为使用 **virtio-*ccw** 设备。例如，使用 **virtio-net-ccw** 而不是 **virtio-net-pci**。
请注意，支持直接附加 PCI 设备（也称 PCI 透传）。

支持的客户端操作系统

如果红帽只使用 RHEL 7、8 或 9 作为客户机操作系统，红帽只支持在 IBM Z 上托管的虚拟机。

设备引导顺序

IBM Z 不支持 `<boot dev='device'>` XML 配置元素。要定义设备引导顺序, 请使用 XML 的 `<devices>` 部分中的 `<boot order='number'>` 元素。

另外, 您可以使用 `<boot>` 元素中特定于架构的 **loadparm** 属性选择所需的 引导条目。例如, 以下决定在引导序列中最先使用磁盘, 以及该磁盘上是否有 Linux 发行版可用, 它将选择第二个引导条目 :

```
<disk type='file' device='disk'>
<driver name='qemu' type='qcow2' />
<source file='/path/to/qcow2' />
<target dev='vda' bus='virtio' />
<address type='ccw' cssid='0xfe' ssid='0x0' devno='0x0000' />
<boot order='1' loadparm='2' />
</disk>
```



注意

在 AMD64 和 Intel 64 主机上也首选使用 `<boot order='number'>` 进行引导顺序管理。

内存热插拔

在 IBM Z 上无法将内存添加到正在运行的虚拟机。请注意, 在 IBM Z 上, 以及 AMD64 和 Intel64 上, 从正在运行的虚拟机 (**内存热插拔**) 中不可能删除内存。

NUMA 拓扑

在 IBM Z 上, **libvirt** 不支持 CPU 的非统一内存访问(NUMA)拓扑。因此, 在这些系统上无法使用 NUMA 来调节 vCPU 性能。

vfio-ap

IBM Z 主机上的虚拟机可以使用 **vfio-ap** 加密设备透传, 其它构架都不支持它。

vfio-ccw

IBM Z 主机上的虚拟机可以使用 **vfio-ccw** 磁盘设备透传, 其它构架都不支持它。

SMBIOOS

IBM Z 不提供 SMBIOS 配置。

watchdog 设备

如果在 IBM Z 主机上的虚拟机中使用看门狗设备, 请使用 **diag288** 模型。例如 :

```
<devices>
<watchdog model='diag288' action='poweroff' />
</devices>
```

kvm-clock

kvm-clock 服务特定于 AMD64 和 Intel 64 系统, 不需要为 IBM Z 上的虚拟机时间管理配置。

v2v 和 **p2v**

virt-v2v 和 **virt-p2v** 工具仅在 AMD64 和 Intel 64 构架上支持, 在 IBM Z 上不提供。

嵌套虚拟化

创建嵌套虚拟机需要在 IBM Z 上进行与 AMD64 和 Intel64 的不同设置。详情请参阅 [创建嵌套虚拟机](#)。

早期版本中没有图形输出

当在主机上使用 RHEL 8.3 或更早的次版本时, 当使用 VNC 协议连接到虚拟机时, 无法显示虚拟机图形输出。这是因为 IBM Z 上的早期 RHEL 版本不支持 **gnome-desktop** 工具。此外, SPICE 显示协议不适用于 IBM Z。

迁移

要成功迁移到以后的主机型号（例如从 IBM z14 到 z15），或更新 hypervisor，请使用 **host-model** CPU 模式。不建议使用 **host-passthrough** 和 **maximum** CPU 模式，因为它们通常不适合迁移。如果要在 **custom** CPU 模式中指定明确的 CPU 模式，请按照以下步骤操作：

- 不要使用以 **-base** 结尾的 CPU 模式。
- 不要使用 **qemu**、**max** 或 **host** CPU 模式。

要成功迁移到较旧的主机型号（如从 z15 到 z14），或者迁移到早期版本的 QEMU、KVM 或 RHEL 内核，请使用结尾没有 **-base** 的最老的可用的主机型号的 CPU 类型。

- 如果源主机和目标主机正在运行，您可以在目标主机上使用 **virsh hypervisor-cpu-baseline** 命令来获取合适的 CPU 模型。详情请参阅 [验证虚拟机迁移的主机 CPU 兼容性](#)。
- 有关 RHEL 8 中支持的机器类型的更多信息，请参阅 [RHEL 8 虚拟化中推荐的功能](#)。

PXE 安装和引导

当使用 PXE 在 IBM Z 上运行虚拟机时，**pxelinux.cfg/default** 文件需要一个特定的配置。例如：

```
# pxelinux
default linux
label linux
kernel kernel.img
initrd initrd.img
append ip=dhcp inst.repo=example.com/redhat/BaseOS/s390x/os/
```

安全执行

您可以通过在虚拟机 XML 配置中定义 `<launchSecurity type="s390-pv" />` 来引导带有准备好的安全客户机镜像的虚拟机。这会加密虚拟机的内存来保护它不受虚拟机监控程序不需要的访问。

请注意，在以安全执行模式运行虚拟机时不支持以下功能：

- 使用 **vfio** 的设备透传
- 使用 **virsh domstats** 和 **virsh memstat** 获取内存信息
- **memballoon** 和 **virtio-rng** 虚拟设备
- 使用巨页的内存支持
- 实时和非实时迁移
- 保存和恢复虚拟机
- 虚拟机快照，包括内存快照（使用 **--memspec** 选项）
- 完整内存转储。反之，为 **virsh dump** 命令指定 **--memory-only** 选项。
- 248 或更多 vCPU。安全客户机的 vCPU 限值为 247。
- 嵌套虚拟化

其它资源

- 跨构架支持的虚拟化功能概述

4.4. 后续步骤

- 在 IBM Z 系统上设置虚拟机时，建议保护客户机操作系统不受 "Spectre" 漏洞的影响。要做到这一点，请使用 **virsh edit** 命令修改虚拟机的 XML 配置，并使用以下方法之一配置其 CPU：

- 使用主机 CPU 模式：

```
<cpu mode='host-model' check='partial'>
  <model fallback='allow'/>
</cpu>
```

这使得 **ppa15** 和 **bpb** 功能可供虚拟客户机使用（如果主机支持这两个功能）。

- 如果使用特定的主机模式，请添加 **ppa15** 和 **bpb** 功能。以下示例使用 zEC12 CPU 模型：

```
<cpu mode='custom' match='exact' check='partial'>
  <model fallback='allow'>zEC12</model>
  <feature policy='force' name='ppa15' />
  <feature policy='force' name='bpb' />
</cpu>
```

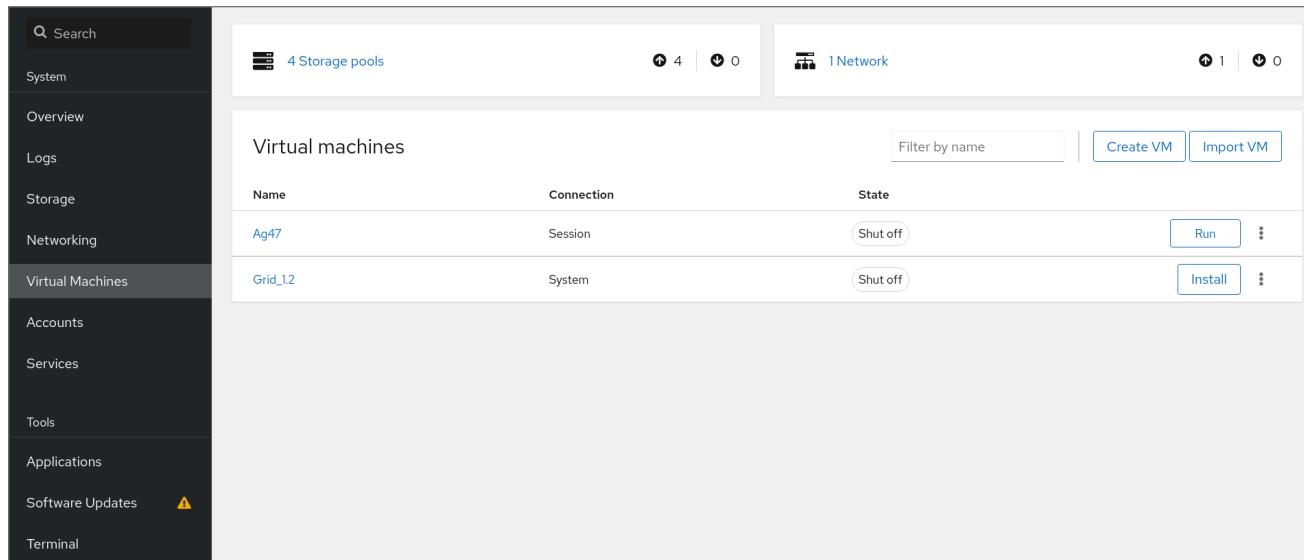
请注意，当在使用 z12 CPU 的主机机器上将 **ppa15** 功能与 **z114** 和 **z196** CPU 模式一起使用时，您还需要使用最新的微码级别（bundle 95 或更高版本）。

4.5. 其它资源

- 将 DASD 设备附加到 IBM Z 中的虚拟机
- 将加密 coprocessors 附加到 IBM Z 上的虚拟机
- 在 IBM Z 上设置 IBM Secure Execution
- 在 IBM Z 上创建嵌套的虚拟机
- 在 IBM Z 中配置透传 PCI 设备

第5章 在WEB控制台中管理虚拟机

要在 RHEL 8 主机上的图形界面中管理虚拟机，您可以使用 RHEL 8 web 控制台中的 **Virtual Machines** 窗格。



5.1. 使用 WEB 控制台管理虚拟机概述

RHEL 8 web 控制台是一个用于系统管理的基于 web 的界面。作为其功能之一，Web 控制台提供主机系统中虚拟机（VM）的图形视图，并可创建、访问和配置这些虚拟机。

请注意，要使用 Web 控制台管理 RHEL 8 上的虚拟机，您必须首先安装用于虚拟化的 [web 控制台插件](#)。

后续步骤

- 有关在 web 控制台中启用虚拟机管理的说明，请参阅 [设置 Web 控制台来管理虚拟机](#)。
- 有关 web 控制台提供的虚拟机管理操作的完整列表，请参阅 [web 控制台中提供的虚拟机管理功能](#)。
- 有关 web 控制台中目前不可用但在 `virt-manager` 应用程序中可以使用的功能列表，请参阅 [虚拟机管理器和 Web 控制台中虚拟化功能之间的区别](#)。

5.2. 设置 WEB 控制台以管理虚拟机

在使用 RHEL 8 web 控制台管理虚拟机(VM)之前，您必须在主机上安装 web 控制台虚拟机插件。

先决条件

- 确保机器上安装并启用了 Web 控制台。

```
# systemctl status cockpit.socket
cockpit.socket - Cockpit Web Service Socket
Loaded: loaded (/usr/lib/systemd/system/cockpit.socket)
[...]
```

如果此命令返回 **Unit cockpit.socket could not be found**，请按照 [安装 Web 控制台](#) 文档来启用 Web 控制台。

流程

- 安装 **cockpit-machines** 插件。

```
# yum install cockpit-machines
```

验证

- 访问 Web 控制台，例如在浏览器中输入 <https://localhost:9090> 地址。
- 登录。
- 如果安装成功，则 **Virtual Machines** 会显示在 web 控制台侧菜单中。

The screenshot shows the Red Hat Enterprise Linux 8 web control panel. On the left, there is a dark sidebar menu with the following items: Search, System, Overview, Logs, Storage, Networking, Virtual Machines (which is highlighted), Accounts, Services, Tools, Applications, Software Updates (with a yellow warning icon), and Terminal. The main content area has three tabs at the top: 'Storage pools' (4 pools, 0 free), 'Network' (1 network, 0 free), and 'Virtual machines'. Below these tabs, there is a table titled 'Virtual machines' with two rows. The first row contains 'Ag47' under 'Name', 'Session' under 'Connection', and 'Shut off' under 'State'. The second row contains 'Grid_12' under 'Name', 'System' under 'Connection', and 'Shut off' under 'State'. To the right of the table are buttons for 'Create VM' and 'Import VM'. There is also a 'Filter by name' input field.

其它资源

- [使用 RHEL 8 web 控制台管理系统](#)

5.3. 使用 WEB 控制台重命名虚拟机

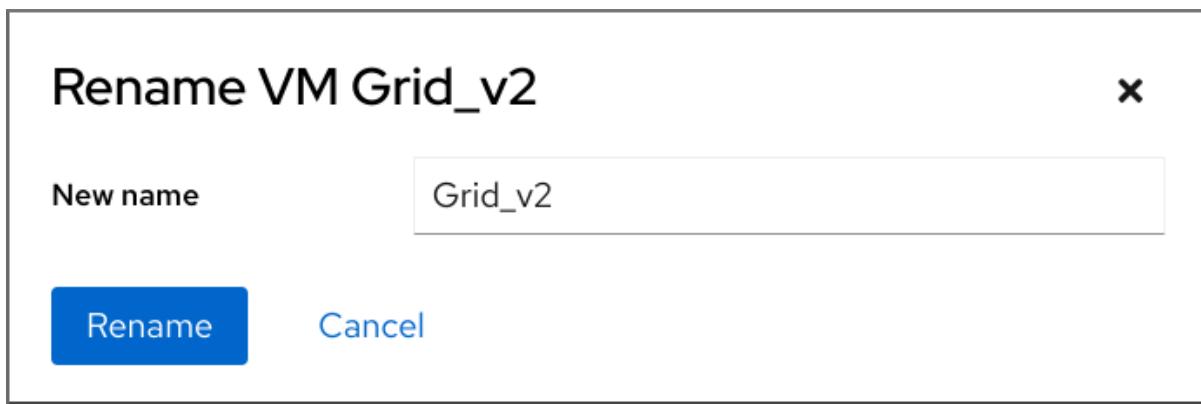
创建虚拟机(VM)后，您可能想要重命名虚拟机以避免冲突，或者根据您的用例分配一个新的唯一的名称。您可以使用 RHEL web 控制台来重命名虚拟机。

先决条件

- Web 控制台 VM 插件 [已安装在您的系统上](#)。
- 确保虚拟机已关闭。

流程

- 在 **Virtual Machines** 界面中，点击您要重命名的虚拟机的菜单按钮 **:**。此时会出现一个下拉菜单，控制各种虚拟机操作。
- 点击 **Rename**。此时会出现重命名虚拟机对话框。



3. 在 **New name** 字段中输入虚拟机的名称。

4. 点击 **Rename**。

验证

- 新虚拟机名称应该出现在 **Virtual Machines** 界面中。

5.4. WEB 控制台中提供的虚拟机管理功能

使用 RHEL 8 web 控制台，您可以执行以下操作来管理系统上的虚拟机(VM)。

表 5.1. 您可以在 RHEL 8 web 控制台中执行的虚拟机管理任务

任务	详情请查看
创建虚拟机并将其安装到客户端操作系统	使用 web 控制台创建虚拟机并安装客户端操作系统
删除虚拟机	使用 web 控制台删除虚拟机
启动、关闭和重启虚拟机	使用 web 控制台启动虚拟机 以及 使用 web 控制台关闭和重启虚拟机
使用各种控制台连接到虚拟机并与虚拟机交互	使用 web 控制台与虚拟机交互
查看有关虚拟机的各种信息	使用 web 控制台查看虚拟机信息
调整分配给虚拟机的主机内存	使用 web 控制台添加和删除虚拟机内存
管理虚拟机的网络连接	使用 web 控制台管理虚拟机网络接口
管理主机上可用的虚拟机存储，并将虚拟磁盘附加到虚拟机	使用 web 控制台管理虚拟机的存储
配置虚拟机的虚拟 CPU 设置	使用 web 控制台管理 virtual CPU
实时迁移虚拟机	使用 web 控制台实时迁移虚拟机
管理主机设备	使用 Web 控制台管理主机设备

5.5. 虚拟机管理器和 WEB 控制台中虚拟化功能的不同

RHEL 8 支持虚拟机管理器(**virt-manager**)应用程序，但已弃用。Web 控制台旨在在以后的主发行版本中替换。因此，建议您熟悉通过 web 控制台使用 GUI 管理虚拟化。

但是，在 RHEL 8 中，某些虚拟机管理任务只能在 **virt-manager** 或命令行中执行。下表主要介绍 **virt-manager** 中可用但没有在 RHEL 8.0 web 控制台中提供的功能。

如果在 RHEL 8 的后续次要版本中提供了此功能，则在 *Support in web console introduced* 列中会显示最低的 RHEL 8 版本。

表 5.2. 在 RHEL 8.0 中，无法通过 Web 控制台执行的虚拟机管理任务

任务	引入的 Web 控制台支持	使用 CLI 的替代方法
将虚拟机设置为在主机引导时启动	RHEL 8.1	virsh autostart
挂起虚拟机	RHEL 8.1	virsh suspend
恢复挂起的虚拟机	RHEL 8.1	virsh restore
创建文件系统目录存储池	RHEL 8.1	virsh pool-define-as
创建 NFS 存储池	RHEL 8.1	virsh pool-define-as
创建物理磁盘设备存储池	RHEL 8.1	virsh pool-define-as
创建 LVM 卷组存储池	RHEL 8.1	virsh pool-define-as
创建基于分区的存储池	当前不可用	virsh pool-define-as
创建基于 GlusterFS 的存储池	当前不可用	virsh pool-define-as
使用 SCSI 设备创建基于 vHBA 的存储池	当前不可用	virsh pool-define-as
创建基于多路径的存储池	当前不可用	virsh pool-define-as
创建基于 RBD 的存储池	当前不可用	virsh pool-define-as
创建新存储卷	RHEL 8.1	virsh vol-create
添加一个新的虚拟网络	RHEL 8.1	virsh net-create 或 virsh net-define
删除虚拟网络	RHEL 8.1	virsh net-undefine
从主机机器接口创建到虚拟机的桥接	当前不可用	virsh iface-bridge

任务	引入的 Web 控制台支持	使用 CLI 的替代方法
创建快照	当前不可用	virsh snapshot-create-as
恢复到快照	当前不可用	virsh snapshot-revert
删除快照	当前不可用	virsh snapshot-delete
克隆虚拟机	RHEL 8.4	virt-clone
将虚拟机迁移到另一台主机机器中	RHEL 8.5	virsh migrate
将主机设备附加到虚拟机	RHEL 8.5	virt-xml --add-device
从虚拟机中删除主机设备	RHEL 8.5	virt-xml --remove-device

其它资源

- 在 RHEL 7 中使用虚拟机管理器(在 RHEL 8 及更高版本中已弃用)

第 6 章 查看有关虚拟机的信息

当需要在 RHEL 8 上调整或排除虚拟化部署的任何方面时，您需要执行的第一个步骤通常是查看有关虚拟机当前状态和配置的信息。要做到这一点，您可以使用 [命令行界面](#) 或 [Web 控制台](#)。您还可以查看虚拟机 [XML 配置](#) 中的信息。

6.1. 使用命令行界面查看虚拟机信息

要检索主机上虚拟机（VM）的信息，请使用以下一个或多个命令。

流程

- 获取主机上的虚拟机列表：

```
# virsh list --all
Id  Name      State
-----
1   testguest1    running
-   testguest2    shut off
-   testguest3    shut off
-   testguest4    shut off
```

- 要获取有关特定虚拟机的基本信息：

```
# virsh dominfo testguest1
Id:          1
Name:        testguest1
UUID:        a973666f-2f6e-415a-8949-75a7a98569e1
OS Type:     hvm
State:       running
CPU(s):      2
CPU time:   188.3s
Max memory: 4194304 KiB
Used memory: 4194304 KiB
Persistent:  yes
Autostart:   disable
Managed save: no
Security model: selinux
Security DOI: 0
Security label: system_u:system_r:svirt_t:s0:c486,c538 (enforcing)
```

- 要获得特定虚拟机的完整 XML 配置：

```
# virsh dumpxml testguest2
<domain type='kvm' id='1'>
  <name>testguest2</name>
  <uuid>a973434f-2f6e-4ěša-8949-76a7a98569e1</uuid>
  <metadata>
  [...]
```

- 有关虚拟机磁盘和其它块设备的详情：

```
# virsh domblklist testguest3
```

```
Target  Source
-----
vda    /var/lib/libvirt/images/testguest3.qcow2
sda    -
sdb    /home/username/Downloads/virt-p2v-1.36.10-1.el7.iso
```

有关管理虚拟机存储的说明，请参阅 [管理虚拟机的存储](#)。

- 要获取有关虚拟机文件系统及其挂载点的信息：

```
# virsh domfsinfo testguest3
Mountpoint Name Type Target
-----
/      dm-0  xfs
/boot   vda1 xfs
```

- 要获取有关特定虚拟机 vCPU 的详细信息：

```
# virsh vcpuinfo testguest4
VCPU:      0
CPU:       3
State:     running
CPU time:  103.1s
CPU Affinity: yyyy

VCPU:      1
CPU:       0
State:     running
CPU time:  88.6s
CPU Affinity: yyyy
```

要在虚拟机中配置和优化 vCPU，请参阅 [优化虚拟机 CPU 性能](#)。

- 列出主机上的所有虚拟网络接口：

```
# virsh net-list --all
Name      State  Autostart Persistent
-----
default   active yes      yes
labnet    active yes      yes
```

有关特定接口的详情：

```
# virsh net-info default
Name:      default
UUID:     c699f9f6-9202-4ca8-91d0-6b8cb9024116
Active:    yes
Persistent: yes
Autostart: yes
Bridge:   virbr0
```

有关网络接口、虚拟机网络和配置它们的说明，请参阅 [配置虚拟机网络连接](#)。

- 有关查看主机上存储池和存储卷信息的说明，请参阅 [使用 CLI 查看虚拟机存储信息](#)。

6.2. 使用 WEB 控制台查看虚拟机信息

使用 RHEL 8 web 控制台，您可以查看 web 控制台会话可以访问的所有[虚拟机](#)和[存储池](#)的信息。

您可以查看 web 控制台会话连接到的所选虚拟机的信息。这包括其[磁盘](#)、[虚拟网络接口](#)和[资源使用情况](#)的信息。

6.2.1. 在 web 控制台中查看虚拟化概述

使用 Web 控制台，您可以访问虚拟化概述，其中包含有关可用虚拟机(VM)、存储池和网络的汇总信息。

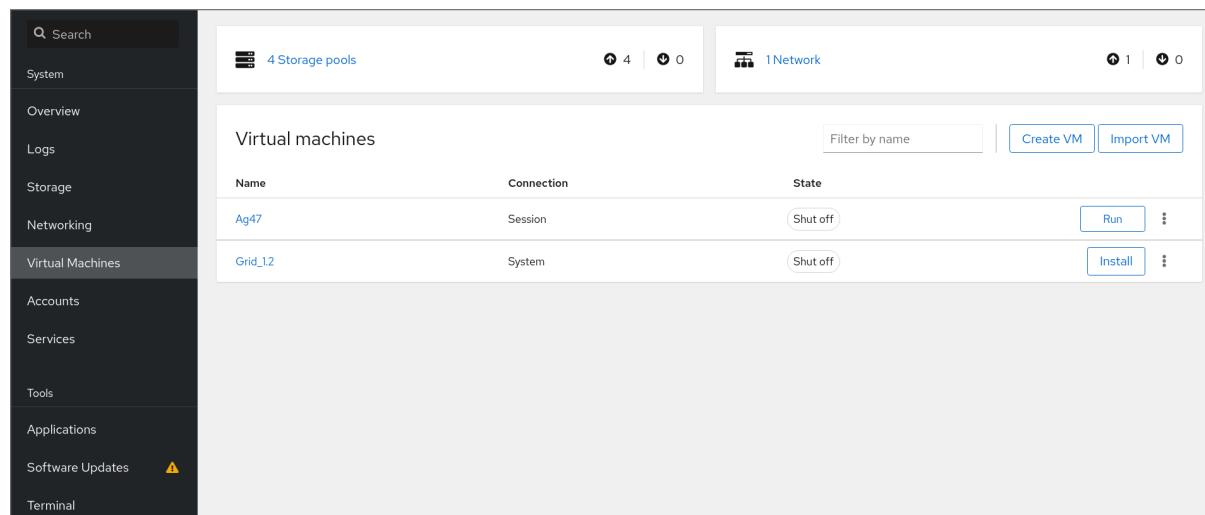
先决条件

- Web 控制台 VM 插件[已安装在您的系统上](#)。

流程

- 在 web 控制台侧菜单中点击 **Virtual Machines**。

此时将显示一个对话框，其中包含有关 Web 控制台所连接的可用存储池、可用网络和虚拟机的信息。



该信息包括：

- **存储池** - 可以通过 Web 控制台及其状态访问的活动的或非活动的存储池数量。
- **网络** - 可以通过 Web 控制台及其状态访问的活动的或非活动的网络数量。
- **Name** - 虚拟机的名称。
- **Connection** - libvirt 连接、系统或者会话的类型。
- **State** - 虚拟机的状态。

其它资源

- [使用 web 控制台查看虚拟机信息](#)

6.2.2. 使用 Web 控制台查看存储池信息

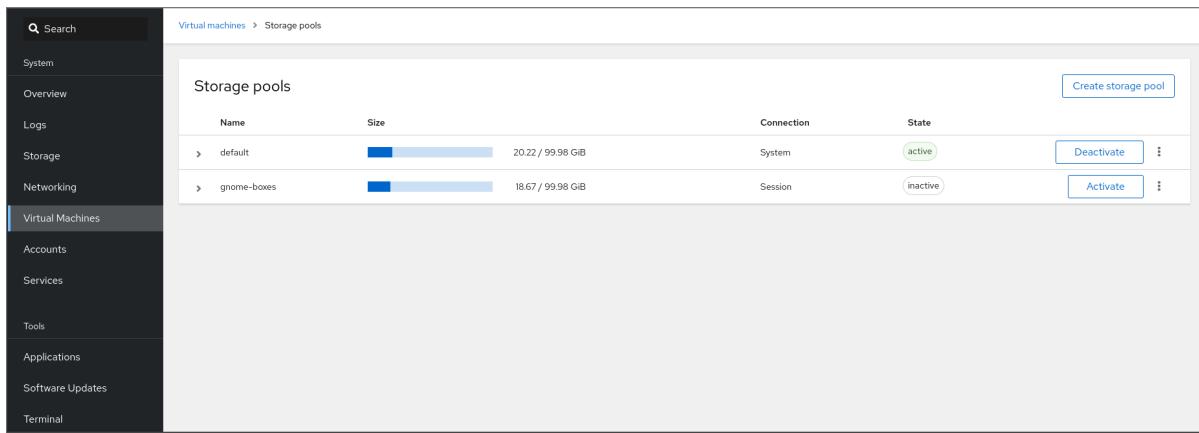
使用 Web 控制台，您可以查看关于系统上可用存储池的详细信息。存储池可用于为您的虚拟机创建磁盘映像。

先决条件

- Web 控制台 VM 插件 已安装在您的系统上。

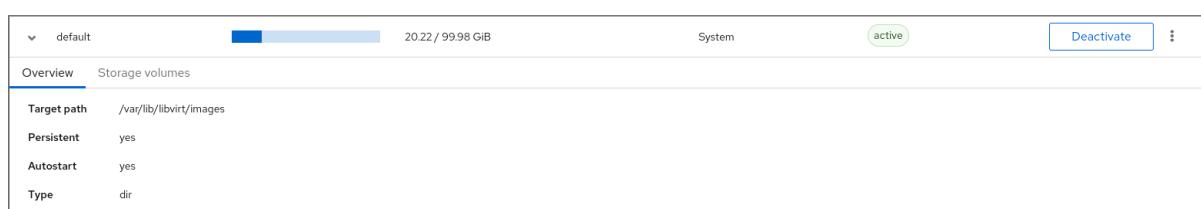
流程

1. 点 Virtual Machines 接口顶部的 **Storage Pools**。
此时会出现 存储池 窗口，显示配置的存储池列表。



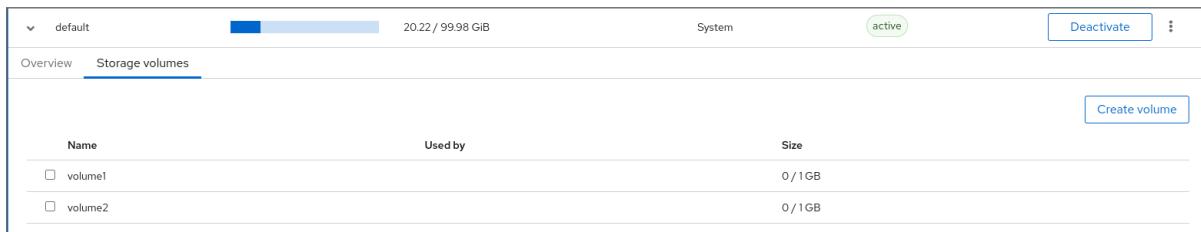
该信息包括：

- **Name** - 存储池的名称。
 - **Size** - 当前分配和存储池的总容量。
 - **connection** - 用于访问存储池的连接。
 - **State** - 存储池的状态。
2. 点击您要查看其信息的存储池旁的箭头。
行会展开，以显示概述窗格，其中含有所选存储池的详细信息。



该信息包括：

- **Target path** - 目录支持的存储池类型的源，如 **dir** 或 **netfs**。
 - **Persistent** - 表示存储池是否具有持久性配置。
 - **Autostart** - 表示在系统启动时存储池是否自动启动。
 - **类型** - 存储池的类型。
3. 要查看与存储池相关联的存储卷的列表，请单击 **Storage Volumes**。
这时将出现存储卷窗格，显示已配置的存储卷的列表。



该信息包括：

- **Name** - 存储卷的名称。
- **Used by** - 当前使用存储卷的虚拟机。
- **Size** - 卷的大小。

其它资源

- 使用 web 控制台查看虚拟机信息

6.2.3. 在 web 控制台中查看基本虚拟机信息

使用 web 控制台，您可以查看关于所选虚拟机(VM)的基本信息，如分配的资源或 hypervisor 详情。

先决条件

- Web 控制台 VM 插件 已安装在您的系统上。

流程

1. 在 web 控制台侧菜单中点击 **Virtual Machines**。

2. 点击您要查看其信息的虚拟机。

此时将打开一个新页面，其中包含关于所选虚拟机基本信息的概述部分，以及用于访问虚拟机的图形界面的控制台部分。

概述部分包括以下常规虚拟机详情：

- **State** - VM 的状态，运行或关闭。

- **Memory** - 分配给虚拟机的内存量。
- **vCPUs** - 为虚拟机配置的虚拟 CPU 数量。
- **CPU Type** - 为虚拟机配置的虚拟 CPU 的构架。
- **Boot Order** - 为虚拟机配置的引导顺序。
- **Autostart** - 是否为虚拟机启用自动启动。

该信息还包括以下管理程序详情：

- **Emulated Machine** - 虚拟机模拟的机器类型。
- **Firmware** - 虚拟机的固件。

其它资源

- [使用 web 控制台查看虚拟机信息](#)
- [使用 Web 控制台管理虚拟 CPU](#)

6.2.4. 在 web 控制台中查看虚拟机资源使用情况

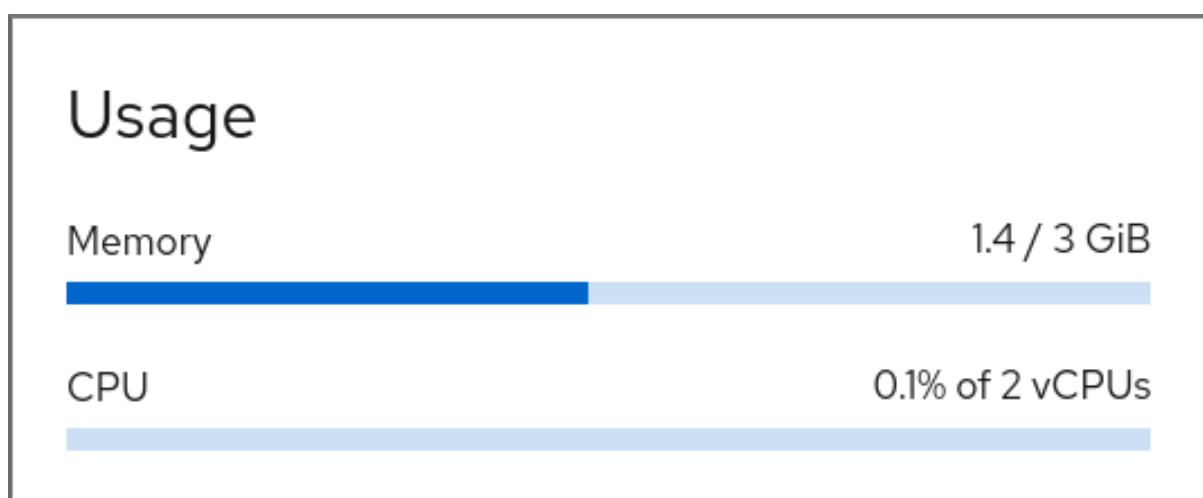
使用 web 控制台，您可以查看所选虚拟机(VM)的内存和虚拟 CPU 使用情况。

先决条件

- Web 控制台 VM 插件 [已安装在您的系统上](#)。

流程

1. 在 **Virtual Machines** 界面中，点击您要查看其信息的虚拟机。
此时将打开一个新页面，其中包含关于所选虚拟机基本信息的概述部分，以及用于访问虚拟机的图形界面的控制台部分。
2. 滚动到 **使用情况**。
使用情况部分显示有关虚拟机内存和虚拟 CPU 使用情况的信息。



其它资源

- 使用 web 控制台查看虚拟机信息

6.2.5. 在 web 控制台中查看虚拟机磁盘信息

使用 web 控制台，您可以查看分配给所选虚拟机(VM)的磁盘的详细信息。

先决条件

- Web 控制台 VM 插件 [已安装在您的系统上](#)。

流程

1. 点击您要查看其信息的虚拟机。

此时将打开一个新页面，其中包含关于所选虚拟机基本信息的概述部分，以及用于访问虚拟机的图形界面的控制台部分。

2. 滚动到 **磁盘**。

磁盘部分显示分配给虚拟机的磁盘以及 **添加、删除 或 编辑** 磁盘的选项。

Disks						
Device	Used	Capacity	Bus	Access	Source	
disk	8.9 GiB	10 GiB	virtio	Writeable	File	/var/lib/libvirt/images/Grid_v2.qcow2
disk	0 GiB	15 GiB	virtio	Writeable	Pool	default
				Volume	v2	

该信息包括：

- **Device** - 磁盘的设备类型。
- **Used** - 当前分配的磁盘数量。
- **Capacity** - 存储卷的最大大小。
- **Bus** - 模拟的磁盘设备的类型。
- **Access** - 磁盘是否是 **可写** 或 **只读**。对于 **raw** 磁盘，您还可以将访问权限设为 **可写和共享**。
- **Source** - 磁盘设备或者文件。

其它资源

- 使用 web 控制台查看虚拟机信息

6.2.6. 在 web 控制台中查看和编辑虚拟网络接口信息

使用 RHEL 8 web 控制台，您可以在所选虚拟机(VM)上查看和修改虚拟网络接口：

先决条件

- Web 控制台 VM 插件 [已安装在您的系统上](#)。

流程

- 在 **Virtual Machines** 界面中，点击您要查看其信息的虚拟机。此时将打开一个新页面，其中包含关于所选虚拟机基本信息的概述部分，以及用于访问虚拟机的图形界面的控制台部分。
- 滚动 **到网络接口**。网络接口部分显示为虚拟机配置的虚拟网络接口的信息，以及 **添加、删除、编辑或拔掉** 网络接口的选项。

Network interfaces						Add network interface
Type	Model type	MAC address	IP address	Source	State	
network	virtio	52:54:00	inet 192.168.122.9/24	default	up	Delete Unplug Edit

该信息包括：

- 类型** - 虚拟机的网络接口类型。类型包括虚拟网络、到 LAN 的桥接和直接连接。



注意

RHEL 8 及更新版本不支持通用以太网连接。

- 型号类型** - 虚拟网络接口的型号。
- MAC 地址** - 虚拟网络接口的 MAC 地址。
- IP 地址** - 虚拟网络接口的 IP 地址。
- Source** - 网络接口源。这取决于网络类型。
- State** - 虚拟网络接口的状态。

- 要编辑虚拟网络接口设置，请点 **Edit**。此时会打开「虚拟网络接口设置」对话框。

52:54:00:b4:2a:62 virtual network interface settings

Interface type	Virtual network
Source	default
Model	(Linux, perf)
MAC address	52:64:00:b4:2a:63

[Save](#) [Cancel](#)

- 更改接口类型、源、型号或 MAC 地址。
- 点 **Save**。已修改网络接口。

**注意**

对虚拟网络接口设置的更改仅在重启虚拟机后生效。

此外，MAC 地址只能在虚拟机关闭时修改。

其它资源

- [使用 web 控制台查看虚拟机信息](#)

6.3. 虚拟机 XML 配置示例

虚拟机的 XML 配置（也称为 *域 XML*）决定虚拟机的设置和组件。下表显示了虚拟机（VM）的 XML 配置示例并解释了其内容。

要获取虚拟机的 XML 配置，您可以使用 **virsh dumpxml** 命令，后跟虚拟机的名称。

```
# virsh dumpxml testguest1
```

表 6.1. XML 配置示例

域 XML 部分	描述
<pre><domain type='kvm'> <name>Testguest1</name> <uuid>ec6fbcaa1-3eb4-49da-bf61-bb02fbec4967</uuid> <memory unit='KiB'>1048576</memory> <currentMemory unit='KiB'>1048576</currentMemory></pre>	<p>这是一个名为 <i>Testguest1</i> 的 KVM 虚拟机，内存为 1024 MiB。</p>
<pre><vcpu placement='static'>1</vcpu></pre>	<p>虚拟机被分配为单个虚拟 CPU（vCPU）。</p> <p>有关配置 vCPU 的详情，请参考 优化虚拟机 CPU 性能。</p>
<pre><os> <type arch='x86_64' machine='pc-q35-4.1'>hvm</type> <boot dev='hd' /> </os></pre>	<p>机器构架被设置为 AMD64 和 Intel 64 架构，并使用 Intel Q35 机器类型来决定功能兼容性。将操作系统设置为从硬盘引导。</p> <p>有关创建带有安装的操作系统的虚拟机的详情，请参考 使用 web 控制台 创建虚拟机并安装客户机操作系统。</p>
<pre><features> <acpi/> <apic/> </features></pre>	<p>acpi 和 apic hypervisor 功能被禁用。</p>

域 XML 部分	描述
<cpu mode='host-model' check='partial'/>	功能 XML 中的主机 CPU 定义（可使用 virsh capabilities 获得）会自动复制到虚拟机的 XML 配置中。因此，当虚拟机引导时， libvirt 会选择与主机 CPU 类似的 CPU 型号，然后添加额外的功能来尽可能接近主机型号。
<clock offset='utc'> <timer name='rtc' tickpolicy='catchup' /> <timer name='pit' tickpolicy='delay' /> <timer name='hpet' present='no' /> </clock>	VM 的虚拟硬件时钟使用 UTC 时区。另外，设置了三个不同的计时器以便与 QEMU 管理程序同步。
<on_poweroff>destroy</on_poweroff> <on_reboot>restart</on_reboot> <on_crash>destroy</on_crash>	当虚拟机关闭或其操作系统意外终止时， libvirt 会终止虚拟机并释放其所有分配的资源。虚拟机重启后， libvirt 会使用同样的配置重新启动它。
<pm> <suspend-to-mem enabled='no' /> <suspend-to-disk enabled='no' /> </pm>	这个虚拟机禁用 S3 和 S4 ACPI 睡眠状态。
<devices> <emulator>/usr/bin/qemu-kvm</emulator> <disk type='file' device='disk'> <driver name='qemu' type='qcow2' /> <source file='/var/lib/libvirt/images/Testguest.qcow2' /> <target dev='hda' bus='ide' /> </disk> <disk type='file' device='cdrom'> <driver name='qemu' type='raw' /> <target dev='hdb' bus='ide' /> <readonly /> </disk>	<p>虚拟机使用 /usr/bin/qemu-kvm 二进制文件进行模拟，它连接了两个磁盘设备。</p> <p>第一个磁盘是基于主机上存储的 /var/lib/libvirt/images/Testguest.qcow2 的虚拟硬盘，其逻辑设备名称设为 hda。</p> <p>第二个磁盘是虚拟 CD-ROM，其逻辑设备名称设为 hdb。</p>

域 XML 部分	描述
<pre><controller type='usb' index='0' model='qemu-xhci' ports='15' /> <controller type='sata' index='0' /> <controller type='pci' index='0' model='pcie-root' /> <controller type='pci' index='1' model='pcie-root-port' /> <model name='pcie-root-port' /> <target chassis='1' port='0x10' /> </controller> <controller type='pci' index='2' model='pcie-root-port' /> <model name='pcie-root-port' /> <target chassis='2' port='0x11' /> </controller> <controller type='pci' index='3' model='pcie-root-port' /> <model name='pcie-root-port' /> <target chassis='3' port='0x12' /> </controller> <controller type='pci' index='4' model='pcie-root-port' /> <model name='pcie-root-port' /> <target chassis='4' port='0x13' /> </controller> <controller type='pci' index='5' model='pcie-root-port' /> <model name='pcie-root-port' /> <target chassis='5' port='0x14' /> </controller> <controller type='pci' index='6' model='pcie-root-port' /> <model name='pcie-root-port' /> <target chassis='6' port='0x15' /> </controller> <controller type='pci' index='7' model='pcie-root-port' /> <model name='pcie-root-port' /> <target chassis='7' port='0x16' /> </controller> <controller type='virtio-serial' index='0' /></pre>	<p>VM 使用单个控制器来附加 USB 设备，而用于 PCI-Express (PCIe) 设备的根控制器。此外，提供了 virtio-serial 控制器，它使虚拟机能够以各种方式与主机进行交互，如串行控制台。</p> <p>有关虚拟设备的更多信息，请参阅虚拟设备的类型。</p>
<pre><interface type='network' /> <mac address='52:54:00:65:29:21' /> <source network='default' /> <model type='rtl8139' /> </interface></pre>	<p>网络接口是在虚拟机中设置的，它使用 default 虚拟网络和 rtl8139 网络设备型号。</p> <p>有关配置网络接口的详情，请参考优化虚拟机网络性能。</p>

域 XML 部分	描述
<pre data-bbox="219 265 964 826"><serial type='pty'> <target type='isa-serial' port='0'> <model name='isa-serial'/> </target> </serial> <console type='pty'> <target type='serial' port='0'> </console> <channel type='unix'> <target type='virtio' name='org.qemu.guest_agent.0' /> <address type='virtio-serial' controller='0' bus='0' port='1' /> </channel> <channel type='spicevmc'> <target type='virtio' name='com.redhat.spice.0' /> <address type='virtio-serial' controller='0' bus='0' port='2' /> </channel></pre>	<p data-bbox="1006 220 1414 428">pty 串行控制台是在虚拟机上设置的，可实现基本虚拟机与主机的通信。控制台使用端口 1 上的UNIX通道，以及端口 2 上的半虚拟化SPICE。这个设置是自动设置的，我们不推荐修改这些设置。</p> <p data-bbox="1006 462 1425 530">有关与虚拟机交互的更多信息，请参阅使用 web 控制台与虚拟机交互。</p>
<pre data-bbox="219 938 699 1102"><input type='tablet' bus='usb'> <address type='usb' bus='0' port='1' /> </input> <input type='mouse' bus='ps2' /> <input type='keyboard' bus='ps2' /></pre>	<p data-bbox="1006 893 1425 1073">虚拟机使用虚拟 usb 端口，该端口设定为接收表格输入，并设置了一个虚拟 ps2 端口以接收鼠标和键盘输入。这个设置是自动设置的，我们不推荐修改这些设置。</p>
<pre data-bbox="203 1215 917 1500"><graphics type='spice' autoport='yes' listen='127.0.0.1'> <listen type='address' address='127.0.0.1' /> <image compression='off' /> </graphics> <graphics type='vnc' port='1' autoport='yes' listen='127.0.0.1'> <listen type='address' address='127.0.0.1' /> </graphics></pre>	<p data-bbox="1006 1170 1409 1275">虚拟机使用 VNC 和 SPICE 协议来渲染其图形输出，图形压缩是关闭的。</p>
<pre data-bbox="203 1612 826 1971"><sound model='ich6'> <address type='pci' domain='0x0000' bus='0x00' slot='0x04' function='0x0' /> </sound> <video> <model type='qxl' ram='65536' vram='65536' vgamem='16384' heads='1' primary='yes' /> <address type='pci' domain='0x0000' bus='0x00' slot='0x02' function='0x0' /> </video></pre>	<p data-bbox="1006 1574 1425 1747">为虚拟机设置了 ICH6 HDA 声音设备，并将 QEMU QXL 半虚拟化帧缓冲器设备设为视频加速器。这个设置是自动设置的，我们不推荐修改这些设置。</p>

域 XML 部分	描述
<pre><redirdev bus='usb' type='spicevmc'> <address type='usb' bus='0' port='1' /> </redirdev> <redirdev bus='usb' type='spicevmc'> <address type='usb' bus='0' port='2' /> </redirdev> <memballoon model='virtio'> <address type='pci' domain='0x0000' bus='0x00' slot='0x07' function='0x0' /> </memballoon> </devices> </domain></pre>	虚拟机有两个重新目录来远程附加 USB 设备，打开内存 ballooning。这个设置是自动设置的，我们不推荐修改这些设置。

第7章 保存和恢复虚拟机

要释放系统资源，您可以关闭该系统中运行的虚拟机（VM）。然而，当您再次需要虚拟机时，您必须引导客户端操作系统（OS）并重启应用程序，这可能需要大量时间。要减少这个停机时间并让虚拟机工作负载更早开始运行，您可以使用保存和恢复功能来完全避免操作系统关闭和引导序列。

本节提供有关保存虚拟机的信息，以及在没有完全引导虚拟机的情况下将虚拟机恢复到同一状态的信息。

7.1. 保存和恢复虚拟机的工作方式

保存虚拟机（VM）会将其内存和设备状态保存到主机的磁盘中，并立即停止虚拟机进程。您可以保存处于运行状态或暂停状态的虚拟机，在恢复后，虚拟机将返回到那个状态。

这个过程释放了主机系统中的 RAM 和 CPU 资源以交换磁盘空间，这样可提高主机系统的性能。当虚拟机被恢复时，因为不需要引导客户机操作系统，也避免使用较长的启动周期。

要保存虚拟机，您可以使用命令行界面（CLI）。具体步骤请参阅[使用命令行界面保存虚拟机](#)。

要恢复虚拟机，您可以使用 [CLI](#) 或 [Web 控制台 GUI](#)。

7.2. 使用命令行界面保存虚拟机

您可以将虚拟机(VM)及其当前状态保存到主机的磁盘上。这很有用，例如当您需要将主机的资源用于某些目的时。然后，保存的虚拟机可以快速恢复到其以前的运行状态。

要使用命令行保存虚拟机，请按照以下流程操作。

先决条件

- 确定有足够的磁盘空间来保存虚拟机及其配置。请注意，虚拟机消耗的空间取决于分配给该虚拟机的 RAM 量。
- 确保虚拟机持久。
- 可选：根据需要备份虚拟机中的重要数据。

流程

- 使用 **virsh managedsave** 工具。

例如，以下命令可停止 *demo-guest1* 虚拟机并保存其配置。

```
# virsh managedsave demo-guest1
Domain 'demo-guest1' saved by libvirt
```

保存的虚拟机文件默认位于 `/var/lib/libvirt/qemu/save` 目录中，即 `demo-guest1.save`。

下次[启动](#)虚拟机时，它将自动从上述文件中恢复保存的状态。

验证

- 列出已启用保存的受管虚拟机。在以下示例中，列为 `saved` 的虚拟机启用了受管保存。

```
# virsh list --managed-save --all
Id  Name          State
```

```
-----  
- demo-guest1           saved  
- demo-guest2           shut off
```

列出具有受管保存镜像的虚拟机：

```
# virsh list --with-managed-save --all  
Id  Name          State  
----  
-  demo-guest1    shut off
```

请注意，要列出处于关机状态的保存的虚拟机，您必须使用命令的 **--all** 或 **inactive** 选项。

故障排除

- 如果保存的虚拟机文件变得损坏或不可读，恢复虚拟机将启动标准虚拟机引导。

其它资源

- virsh managedsave --help** 命令
- [使用命令行界面恢复保存的虚拟机](#)
- [使用 Web 控制台恢复保存的虚拟机](#)

7.3. 使用命令行界面启动虚拟机

您可以使用命令行界面(CLI)启动关闭的虚拟机(VM)或恢复保存的虚拟机。使用 CLI，您可以启动本地和远程虚拟机。

先决条件

- 已定义的一个不活跃地虚拟机。
- 虚拟机的名称。
- 对于远程虚拟机：
 - 虚拟机所在主机的 IP 地址。
 - 对主机的 root 访问权限。

流程

- 对于本地虚拟机，请使用 **virsh start** 工具。
例如，以下命令启动 *demo-guest1* 虚拟机。

```
# virsh start demo-guest1  
Domain 'demo-guest1' started
```

- 对于位于远程主机上的虚拟机，请使用 **virsh start** 工具以及与主机的 QEMU+SSH 连接。
例如，以下命令在 192.168.123.123 主机上启动 *demo-guest1* 虚拟机。

```
# virsh -c qemu+ssh://root@192.168.123.123/system start demo-guest1
```

```
root@192.168.123.123's password:
```

```
Domain 'demo-guest1' started
```

其它资源

- **virsh start --help** 命令
- [设置对远程虚拟化主机的简单访问](#)
- [当主机启动时自动启动虚拟机](#)

7.4. 使用 WEB 控制台启动虚拟机

如果虚拟机(VM) 处于关闭状态, 您可以使用 RHEL 8 web 控制台启动它。您还可以将虚拟机配置为在主机启动时自动启动。

先决条件

- Web 控制台 VM 插件 [已安装在您的系统上](#)。
- 已定义的一个不活跃地虚拟机。
- 虚拟机的名称。

流程

1. 在 **Virtual Machines** 界面中, 点击您要启动的虚拟机。
此时将打开一个新页面, 其中包含有关所选虚拟机的详细信息, 以及关闭和删除虚拟机的控制。
2. 点 **Run**。
虚拟机启动, 您可以[连接到其控制台或图形输出](#)。
3. 可选 : 要将虚拟机配置为在主机启动时自动启动, 请切换 **Overview** 部分中的 **Autostart** 复选框。
如果使用不由 libvirt 管理的网络接口, 您还必须对 systemd 配置进行额外的更改。否则, 受影响的虚拟机可能无法启动, 请参阅 [当主机启动时自动启动虚拟机](#)。

其它资源

- [在 web 控制台中关闭虚拟机](#)
- [使用 web 控制台重启虚拟机](#)

第 8 章 克隆虚拟机

要使用特定属性集合快速创建新虚拟机，您可以克隆现有虚拟机。

克隆会创建一个使用其自身磁盘镜像保存存储的新虚拟机，但大多数克隆的配置和源虚拟机的数据都是一样的。这样就可以准备多个虚拟机来为特定的任务进行优化，而无需单独优化每个虚拟机。

8.1. 克隆虚拟机的方式

克隆虚拟机会复制源虚拟机及其磁盘镜像的 XML 配置，并对配置进行修改以确保新虚拟机的唯一性。这包括更改虚拟机的名称，并确保它使用磁盘镜像克隆。存储在克隆的虚拟磁盘上的数据与源虚拟机是一致的。

这个过程比创建新虚拟机要快，并使用客户端操作系统安装它，并可用于快速生成带有特定配置和内容的虚拟机。

如果您计划为虚拟机创建多个克隆，首先请创建一个不包含以下内容的虚拟机模板：

- 唯一设置，如持久性网络 MAC 配置，这可阻止克隆正常工作。
- 敏感数据，如 SSH 密钥和密码文件。

具体说明请参阅 [创建虚拟机模板](#)。

其它资源

- [使用命令行界面克隆虚拟机](#)
- [使用 web 控制台克隆虚拟机](#)

8.2. 创建虚拟机模板

要创建可正常工作的多个虚拟机(VM)克隆，您可以对删除源虚拟机的唯一信息和配置，如 SSH 密钥或持久性网络 MAC 配置。这将创建一个虚拟机 模板，您可以使用它来轻松且安全地创建虚拟机克隆。

您可以使用 **[virt-sysprep](#)** 实用程序或 [根据您的要求手动创建](#) 虚拟机模板。

8.2.1. 使用 **virt-sysprep** 创建虚拟机模板

要从现有虚拟机(VM)创建克隆模板，您可以使用 **virt-sysprep** 工具。这会删除可能导致克隆正常工作的某些配置，如特定网络设置或系统注册元数据。因此，**virt-sysprep** 提高了虚拟机的克隆，并确保克隆更可靠。

先决条件

- **libguestfs-tools-c** 软件包（包含 **virt-sysprep** 工具）已安装在您的主机上：

```
# yum install libguestfs-tools-c
```

- 作为模板的源虚拟机将被关闭。
- 您知道源虚拟机的磁盘镜像所处的位置，并且您是虚拟机磁盘镜像文件的所有者。
请注意，在 libvirt [系统连接](#) 中创建的虚拟机的磁盘镜像位于 **/var/lib/libvirt/images** 目录中，默
认由 root 用户拥有：

```
# ls -la /var/lib/libvirt/images
-rw----- 1 root root 9665380352 Jul 23 14:50 a-really-important-vm.qcow2
-rw----- 1 root root 8591507456 Jul 26 2017 an-actual-vm-that-i-use.qcow2
-rw----- 1 root root 8591507456 Jul 26 2017 totally-not-a-fake-vm.qcow2
-rw----- 1 root root 10739318784 Sep 20 17:57 another-vm-example.qcow2
```

- 可选：源虚拟机磁盘中的任何重要数据都已备份。如果要保留源虚拟机，请首先克隆它，并将克隆转换为模板。[???](#)

流程

- 确保您已作为虚拟机磁盘镜像的所有者登录：

```
# whoami
root
```

- 可选：复制虚拟机的磁盘镜像。

```
# cp /var/lib/libvirt/images/a-really-important-vm.qcow2 /var/lib/libvirt/images/a-really-important-vm-original.qcow2
```

这用于验证虚拟机是否已成功转换为模板。

- 使用以下命令，将 `/var/lib/libvirt/images/a-really-important-vm.qcow2` 替换为源虚拟机磁盘镜像的路径。

```
# virt-sysprep -a /var/lib/libvirt/images/a-really-important-vm.qcow2
[ 0.0] Examining the guest ...
[ 7.3] Performing "abrt-data" ...
[ 7.3] Performing "backup-files" ...
[ 9.6] Performing "bash-history" ...
[ 9.6] Performing "blkid-tab" ...
[...]
```

验证

- 要确认进程成功，请将修改的磁盘镜像与原始镜像进行比较。以下示例显示了成功创建模板：

```
# virt-diff -a /var/lib/libvirt/images/a-really-important-vm-orig.qcow2 -A /var/lib/libvirt/images/a-really-important-vm.qcow2
-- 0644    1001 /etc/group-
-- 0000    797 /etc/gshadow-
= - 0444     33 /etc/machine-id
[...]
-- 0600    409 /home/username/.bash_history
- d 0700      6 /home/username/.ssh
-- 0600    868 /root/.bash_history
[...]
```

其它资源

- `virt-sysprep` 手册页中的 **OPERATIONS** 部分

- 使用命令行界面克隆虚拟机

8.2.2. 手动创建虚拟机模板

要从现有虚拟机(VM)创建模板，您可以手动重置或取消配置客户虚拟机来为克隆做好准备。

先决条件

- 确保您知道源虚拟机的磁盘镜像的位置，并且是虚拟机磁盘镜像文件的所有者。
请注意，在 libvirt 的 [系统连接](#) 中创建的虚拟机的磁盘镜像默认位于 **/var/lib/libvirt/images** 目录中，并由 root 用户拥有：

```
# ls -la /var/lib/libvirt/images
-rw----- 1 root root 9665380352 Jul 23 14:50 a-really-important-vm.qcow2
-rw----- 1 root root 8591507456 Jul 26 2017 an-actual-vm-that-i-use.qcow2
-rw----- 1 root root 8591507456 Jul 26 2017 totally-not-a-fake-vm.qcow2
-rw----- 1 root root 10739318784 Sep 20 17:57 another-vm-example.qcow2
```

- 确保虚拟机已关闭。
- 可选：虚拟机磁盘上的所有重要数据都已备份。如果要保留源虚拟机，请首先[克隆](#)它，并编辑克隆以创建模板。

流程

1. 配置虚拟机以进行克隆：
 - a. 在克隆上安装任何所需的软件。
 - b. 为操作系统配置任何非唯一的设置。
 - c. 配置任何非唯一的应用设置。
2. 删除网络配置：
 - a. 使用以下命令删除任何持久性的 udev 规则：

```
# rm -f /etc/udev/rules.d/70-persistent-net.rules
```



注意

如果没有删除 udev 规则，第一个 NIC 的名称可能是 **eth1** 而不是 **eth0**。

- b. 通过编辑 **/etc/sysconfig/network-scripts/ifcfg-eth[x]** 来从 ifcfg 脚本中删除唯一的网络详情，如下所示：
 - i. 删除 HWADDR 和 Static 行：



注意

如果 HWADDR 与新客户机的 MAC 地址不匹配，则**ifcfg** 将被忽略。

```
DEVICE=eth[x] BOOTPROTO=none ONBOOT=yes #NETWORK=10.0.1.0 <-
```

```
REMOVE #NETMASK=255.255.255.0 <- REMOVE #IPADDR=10.0.1.20 <-
REMOVE #HWADDR=xx:xx:xx:xx:xx <- REMOVE #USERCTL=no <- REMOVE #
Remove any other *unique or non-desired settings, such as UUID.*
```

ii. 配置 DHCP，但不包含 HWADDR 或任何其他唯一信息：

```
DEVICE=eth[x] BOOTPROTO=dhcp ONBOOT=yes
```

c. 如果系统中存在以下文件，请确保这些文件也包含同样的内容：

- `/etc/sysconfig/networking/devices/ifcfg-eth[x]`
- `/etc/sysconfig/networking/profiles/default/ifcfg-eth[x]`



注意

如果您已使用 **NetworkManager** 或虚拟机的任何特殊设置，请确保从 **ifcfg** 脚本中删除任何其他唯一的信息。

3. 删除注册详情：

- 对于在 Red Hat Network (RHN) 上注册的虚拟机：

```
# rm /etc/sysconfig/rhn/systemid
```

- 对于使用 Red Hat Subscription Manager (RHSM) 注册的虚拟机：

- 如果您不打算使用原始虚拟机：

```
# subscription-manager unsubscribe --all # subscription-manager unregister #
subscription-manager clean
```

- 如果您打算使用原始虚拟机：

```
# subscription-manager clean
```



注意

原始 RHSM 配置文件保留您 ID 代码的门户中。克隆后，使用以下命令重新激活虚拟机上的 RHSM 注册：

```
#subscription-manager register --consumerid=71rd64fx-6216-4409-
bf3a-e4b7c7bd8ac9
```

4. 删除其他唯一的详情：

- a. 删除 ssh 公钥/私钥对：

```
# rm -rf /etc/ssh/ssh_host_example
```

- b. 删除 LVM 设备的配置：

```
# rm /etc/lvm/devices/system.devices
```

c. 如果在多台计算机上运行，删除任何其他特定于应用程序的标识符或配置可能会导致冲突。

5. 删除 **gnome-initial-setup-done** 文件，将虚拟机配置为在下次引导时运行配置向导：

```
# rm ~/.config/gnome-initial-setup-done
```



注意

在下次引导时运行的向导依赖于已从虚拟机中删除的配置。另外，在克隆的第一次引导时，会建议您更改主机名。

8.3. 使用命令行界面克隆虚拟机

要使用特定属性集（如为测试目的）快速创建新虚拟机，您可以克隆现有的虚拟机。要使用 CLI 完成此操作，请遵循以下步骤。

先决条件

- 源虚拟机被关闭。
- 确保有足够的磁盘空间来存储克隆的磁盘镜像。
- 可选**：在创建多个虚拟机克隆时，从源虚拟机中删除唯一数据和设置，以确保克隆的虚拟机正常工作。具体说明请参阅 [创建虚拟机模板](#)。

流程

1. 使用 **virt-clone** 工具以及适合您环境和用例的选项。

使用案例示例

- 以下命令克隆一个名为 *doppelganger* 的本地虚拟机，并创建 *doppelganger-clone* 虚拟机。它还在与原始虚拟机磁盘镜像相同的位置，使用相同的数据创建 *doppelganger-clone.qcow2* 磁盘镜像：

```
# virt-clone --original doppelganger --auto-clone
Allocating 'doppelganger-clone.qcow2' | 50.0 GB 00:05:37
Clone 'doppelganger-clone' created successfully.
```

- 以下命令克隆一个名为 *geminus1* 的虚拟机，并创建一个名为 *geminus2* 的本地虚拟机，它只使用 *geminus1* 的多个磁盘中的两个。

```
# virt-clone --original geminus1 --name geminus2 --file /var/lib/libvirt/images/disk1.qcow2
--file /var/lib/libvirt/images/disk2.qcow2
Allocating 'disk1-clone.qcow2' | 78.0 GB 00:05:37
Allocating 'disk2-clone.qcow2' | 80.0 GB 00:05:37
Clone 'geminus2' created successfully.
```

- 要将虚拟机克隆到其他主机，请迁移虚拟机而无需在本地主机上取消它。例如，以下命令将之前创建的 `geminus2` 虚拟机克隆到 `10.0.0.1` 远程系统，包括本地磁盘。请注意，使用这些命令还需要 `10.0.0.1` 的 root 权限。

```
# virsh migrate --offline --persistent geminus2 qemu+ssh://root@10.0.0.1/system
root@10.0.0.1's password:

# scp /var/lib/libvirt/images/disk1-clone.qcow2
root@10.0.0.1:/user@remote_host.com:/var/lib/libvirt/images/

# scp /var/lib/libvirt/images/disk2-clone.qcow2
root@10.0.0.1:/user@remote_host.com:/var/lib/libvirt/images/
```

验证

验证虚拟机是否已成功克隆，且正在正常工作：

- 确认克隆已添加到主机上的虚拟机列表中。

```
# virsh list --all
Id  Name          State
-----
-  doppelganger    shut off
-  doppelganger-clone  shut off
```

- 启动克隆并观察它是否引导。

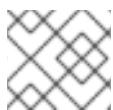
```
# virsh start doppelganger-clone
Domain 'doppelganger-clone' started
```

其它资源

- [virt-clone 手册页](#)
- [迁移虚拟机](#)

8.4. 使用 WEB 控制台克隆虚拟机

要快速创建带有特定属性集的新虚拟机(VM)，您可以克隆之前配置的虚拟机。下列说明解释了如何使用 Web 控制台来进行此操作。



注意

克隆虚拟机也会克隆与该虚拟机关联的磁盘。

先决条件

- Web 控制台 VM 插件 [已安装在您的系统上](#)。
- 确保要克隆的虚拟机已关闭。

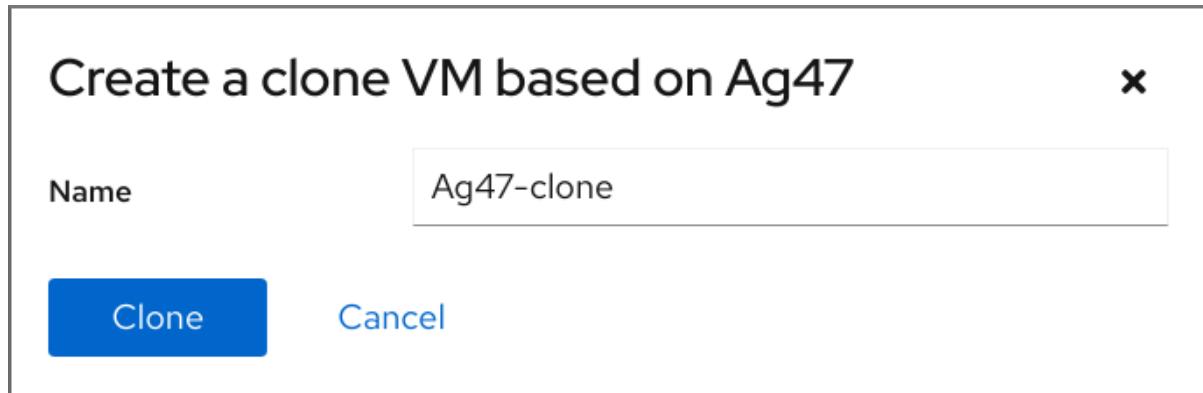
流程

- 在 web 控制台的虚拟机界面中，单击您要克隆的虚拟机的菜单按钮 。

此时会出现一个下拉菜单，控制各种虚拟机操作。

2. 单击 **Clone**。

此时将显示一个创建克隆虚拟机对话框。



3. 可选：输入虚拟机克隆的新名称。

4. 单击 **Clone**。

基于源虚拟机的新虚拟机被创建。

验证

- 确认克隆的虚拟机是否出现在主机上可用的虚拟机列表中。

第9章 迁移虚拟机

如果虚拟机所在的当前主机变得不适合或者无法再使用，或者要重新分发托管工作负载，您可以将该虚拟机迁移到另一个 KVM 主机中。

9.1. 迁移虚拟机的工作方式

虚拟机迁移的基本部分是将虚拟机的 XML 配置复制到不同的主机机器中。如果没有关闭迁移的虚拟机，迁移还会将虚拟机内存和任何虚拟设备的状态传送到目标主机机器中。要使虚拟机在目标主机上正常工作，虚拟机的磁盘镜像必须仍可用。

默认情况下，迁移的虚拟机在目标主机上是临时的，虚拟机在源主机上仍然被定义。

您可以使用 [实时或非实时迁移](#)对运行的虚拟机进行迁移。要迁移关闭虚拟机，必须使用[离线迁移](#)。详情请查看下表。

表 9.1. VM 迁移类型

迁移类型	描述	使用案例	存储要求
实时迁移	VM 将继续在源主机中运行，而 KVM 会将虚拟机的内存页面传送到目标主机。当迁移接近完成后，KVM 会非常简单地挂起虚拟机，并在目标主机上恢复它。	对于需要一直保持运行的虚拟机，这个方法非常有用。但是，如果虚拟机修改内存页面的速度比 KVM 可以传输它们的速度更快，比如 I/O 负载较重的虚拟机，则不能进行实时迁移，这需要使用非实时迁移。	VM 的磁盘镜像必须位于 共享网络 中，同时可访问源主机和目标主机。
非实时迁移	挂起虚拟机，将其配置及其内存复制到目标主机，并恢复虚拟机。	这个迁移方式需要虚拟机停机，但通常比实时迁移更可靠。建议有高 I/O 负载的虚拟机使用这个方法。	VM 的磁盘镜像必须位于 共享网络 中，同时可访问源主机和目标主机。
离线迁移	将虚拟机的配置移到目标主机	建议关闭虚拟机。	VM 的磁盘镜像不必在共享网络中可用，并可手动复制或移动到目标主机。

其它资源

- [迁移虚拟机的好处](#)
- [将虚拟机磁盘镜像与其他主机共享](#)

9.2. 迁移虚拟机的好处

迁移虚拟机对以下情况非常有用：

负载均衡

如果主机超载或者另一台主机使用不足，则可将虚拟机移动到使用率较低的主机中。

硬件独立

当您需要升级、添加或删除主机中的硬件设备时，您可以安全地将虚拟机重新定位到其他主机。这意味着，在改进硬件时虚拟机不需要停机。

节能

虚拟机可重新分发到其他主机，因此可关闭未载入的主机系统以便在低用量时节约能源并降低成本。

地理迁移

可将虚拟机移动到另一个物理位置，以减少延迟，或者因为其他原因需要。

9.3. 迁移虚拟机的限制

在 RHEL 8 中迁移虚拟机(VM)前，请确保您了解迁移的限制。

- 实时存储迁移无法在 RHEL 8 上执行，但您可以在虚拟机关闭时迁移存储。请注意，[红帽虚拟化](#)上提供了实时存储迁移。
- 将虚拟机从或迁移到 [libvirt 的会话连接](#) 是不可靠的，因此不建议这样做。
- 使用以下功能和配置的虚拟机在迁移时将无法正常工作，或者迁移失败。这些特性包括：
 - 设备透传
 - SR-IOV 设备分配
 - 介质设备，如 vGPU
- 使用非统一内存访问(NUMA)固定的主机之间的迁移只有在主机有类似的拓扑时才能正常工作。但是，迁移可能会对运行工作负载的性能造成负面影响。
- 源虚拟机和目标虚拟机上模拟的 CPU 必须相同，否则迁移可能会失败。虚拟机之间在以下 CPU 相关区域的任何区别都可能会引起迁移问题：
 - CPU 型号
 - 不支持在 Intel 64 主机和 AMD64 主机之间迁移，即使它们共享 x86-64 指令集。
 - 有关确保在迁移到具有不同 CPU 模型的主机后虚拟机可以正常工作的步骤，[请参阅验证虚拟机迁移的主机 CPU 兼容性](#)。
 - 固件设置
 - 微码版本
 - BIOS 版本
 - BIOS 设置
 - QEMU 版本
 - 内核版本
- 在一些情况下，实时迁移使用超过 1TB 内存的虚拟机可能无法正常工作。此类迁移的稳定性取决于以下方面：
 - 虚拟机的当前工作负载
 - 主机可用于迁移的网络带宽

- 部署可以支持的停机时间

对于涉及超过 1 TB 内存的虚拟机的实时迁移方案，客户应咨询红帽。

9.4. 验证虚拟机迁移的主机 CPU 兼容性

要使迁移的虚拟机(VM)在目标主机上正常工作，源和目标主机上的 CPU 必须兼容。要确保情况如此，请在开始迁移前计算一个常见 CPU 基准。



注意

本节中的说明使用了具有以下主机 CPU 的迁移示例：

- 源主机：Intel Core i7-8650U
- 目标主机：Intel Xeon CPU E5-2620 v2

先决条件

- 已在您的系统中[安装并启用虚拟化](#)。
- 您具有对源主机和迁移的目标主机的管理员访问权限。

流程

1. 在源主机上，获取其 CPU 功能并将其粘贴到新的 XML 文件中，如 **domCaps-CPUs.xml**。

```
# virsh domcapabilities | xmllint --xpath "//cpu[@name='host-model']" - > domCaps-CPUs.xml
```

2. 在 XML 文件中，将 **< mode> </mode>** 标签替换为 **< cpu> </cpu>**。

3. 可选：验证 **domCaps-CPUs.xml** 文件的内容类似如下：

```
# cat domCaps-CPUs.xml

<cpu>
  <model fallback="forbid">Skylake-Client-IBRS</model>
  <vendor>Intel</vendor>
  <feature policy="require" name="ss"/>
  <feature policy="require" name="pdcm"/>
  <feature policy="require" name="hypervisor"/>
  <feature policy="require" name="tsc_adjust"/>
  <feature policy="require" name="clflushopt"/>
  <feature policy="require" name="umip"/>
  <feature policy="require" name="md-clear"/>
  <feature policy="require" name="stibp"/>
  <feature policy="require" name="arch-capabilities"/>
  <feature policy="require" name="ssbd"/>
  <feature policy="require" name="xsaves"/>
  <feature policy="require" name="pdpe1gb"/>
  <feature policy="require" name="invtsc"/>
  <feature policy="require" name="ibpb"/>
  <feature policy="require" name="ibrs"/>
  <feature policy="require" name="amd-stibp"/>
```

```

<feature policy="require" name="amd-ssbd"/>
<feature policy="require" name="rsba"/>
<feature policy="require" name="skip-l1dfl-vmentry"/>
<feature policy="require" name="pschange-mc-no"/>
<feature policy="disable" name="hle"/>
<feature policy="disable" name="rtm"/>
</cpu>

```

4. 在目标主机上，使用以下命令获取其 CPU 功能：

```

# virsh domcapabilities | xmllint --xpath "//cpu	mode[@name='host-model']" -
<mode name="host-model" supported="yes">
<model fallback="forbid">Skylake-Client-IBRS</model>
<vendor>Intel</vendor>
<feature policy="require" name="ss"/>
<feature policy="require" name="vmx"/>
<feature policy="require" name="pdcm"/>
<feature policy="require" name="hypervisor"/>
<feature policy="require" name="tsc_adjust"/>
<feature policy="require" name="clflushopt"/>
<feature policy="require" name="umip"/>
<feature policy="require" name="md-clear"/>
<feature policy="require" name="stibp"/>
<feature policy="require" name="arch-capabilities"/>
<feature policy="require" name="ssbd"/>
<feature policy="require" name="xsaves"/>
<feature policy="require" name="pdpe1gb"/>
<feature policy="require" name="invtsc"/>
<feature policy="require" name="ibpb"/>
<feature policy="require" name="ibrs"/>
<feature policy="require" name="amd-stibp"/>
<feature policy="require" name="amd-ssbd"/>
<feature policy="require" name="rsba"/>
<feature policy="require" name="skip-l1dfl-vmentry"/>
<feature policy="require" name="pschange-mc-no"/>
<feature policy="disable" name="hle"/>
<feature policy="disable" name="rtm"/>
</mode>

```

5. 将从目标主机获取的 CPU 功能添加到源主机上的 **domCaps-CPUs.xml** 文件。同样，将 **<mode> </mode>** 标签替换为 **<cpu> </cpu>** 并保存文件。
6. 可选：验证 XML 文件现在包含这两个主机的 CPU 功能。

```

# cat domCaps-CPUs.xml

<cpu>
<model fallback="forbid">Skylake-Client-IBRS</model>
<vendor>Intel</vendor>
<feature policy="require" name="ss"/>
<feature policy="require" name="vmx"/>
<feature policy="require" name="pdcm"/>
<feature policy="require" name="hypervisor"/>
<feature policy="require" name="tsc_adjust"/>

```

```

<feature policy="require" name="clflushopt"/>
<feature policy="require" name="umip"/>
<feature policy="require" name="md-clear"/>
<feature policy="require" name="stibp"/>
<feature policy="require" name="arch-capabilities"/>
<feature policy="require" name="ssbd"/>
<feature policy="require" name="xsaves"/>
<feature policy="require" name="pdpe1gb"/>
<feature policy="require" name="invtsr"/>
<feature policy="require" name="ibpb"/>
<feature policy="require" name="ibrs"/>
<feature policy="require" name="amd-stibp"/>
<feature policy="require" name="amd-ssbd"/>
<feature policy="require" name="rsba"/>
<feature policy="require" name="skip-l1dfl-vmentry"/>
<feature policy="require" name="pschange-mc-no"/>
<feature policy="disable" name="hle"/>
<feature policy="disable" name="rtm"/>
</cpu>
<cpu>
    <model fallback="forbid">IvyBridge-IBRS</model>
    <vendor>Intel</vendor>
    <feature policy="require" name="ss"/>
    <feature policy="require" name="vmx"/>
    <feature policy="require" name="pdcm"/>
    <feature policy="require" name="pcid"/>
    <feature policy="require" name="hypervisor"/>
    <feature policy="require" name="arat"/>
    <feature policy="require" name="tsc_adjust"/>
    <feature policy="require" name="umip"/>
    <feature policy="require" name="md-clear"/>
    <feature policy="require" name="stibp"/>
    <feature policy="require" name="arch-capabilities"/>
    <feature policy="require" name="ssbd"/>
    <feature policy="require" name="xsaveopt"/>
    <feature policy="require" name="pdpe1gb"/>
    <feature policy="require" name="invtsr"/>
    <feature policy="require" name="ibpb"/>
    <feature policy="require" name="amd-ssbd"/>
    <feature policy="require" name="skip-l1dfl-vmentry"/>
    <feature policy="require" name="pschange-mc-no"/>
</cpu>

```

7. 使用 XML 文件计算您要迁移的虚拟机的 CPU 功能基准。

```

# virsh hypervisor-cpu-baseline domCaps-CPUs.xml

<cpu mode='custom' match='exact'>
    <model fallback='forbid'>IvyBridge-IBRS</model>
    <vendor>Intel</vendor>
    <feature policy='require' name='ss'/>
    <feature policy='require' name='vmx'/>
    <feature policy='require' name='pdcm'/>
    <feature policy='require' name='pcid'/>
    <feature policy='require' name='hypervisor'/>
    <feature policy='require' name='arat'/>

```

```

<feature policy='require' name='tsc_adjust'/>
<feature policy='require' name='umip'/>
<feature policy='require' name='md-clear'/>
<feature policy='require' name='stibp'/>
<feature policy='require' name='arch-capabilities'/>
<feature policy='require' name='ssbd'/>
<feature policy='require' name='xsaveopt'/>
<feature policy='require' name='pdpe1gb'/>
<feature policy='require' name='invtscl'/>
<feature policy='require' name='ibpb'/>
<feature policy='require' name='amd-ssbd'/>
<feature policy='require' name='skip-l1dfl-vmentry'/>
<feature policy='require' name='pschange-mc-no'/>
</cpu>

```

8. 打开您要迁移的虚拟机 XML 配置，并将 `<cpu>` 部分的内容替换为上一步中获取的设置。

```
# virsh edit VM-name
```

9. 如果虚拟机正在运行，请重新启动它。

```
# virsh reboot VM-name
```

后续步骤

- [将虚拟机磁盘镜像与其他主机共享](#)
- [使用命令行界面迁移虚拟机](#)
- [使用 web 控制台实时迁移虚拟机](#)

9.5. 将虚拟机磁盘镜像与其他主机共享

要在[支持的 KVM 主机](#)间执行虚拟机实时迁移，需要共享虚拟机存储。以下流程提供了使用 NFS 协议将本地存储的虚拟机镜像与源主机和目标主机共享的说明。

先决条件

- 旨在迁移的虚拟机被关闭。
- 可选：一个主机系统可用于托管不是源或目标主机的存储，但源和目标主机都可以通过网络访问它。这是共享存储的最佳解决方案，并被红帽推荐。
- 请确定 NFS 文件锁定没有被使用，因为在 KVM 中不支持它。
- 在源主机和目标主机上安装并启用 NFS。如果没有：
 - a. 安装 NFS 软件包：

```
# yum install nfs-utils
```

- b. 确保在防火墙中打开了 NFS 的端口，如 2049。

```
# firewall-cmd --permanent --add-service=nfs
```

```
# firewall-cmd --permanent --add-service=mountd
# firewall-cmd --permanent --add-service=rpc-bind
# firewall-cmd --permanent --add-port=2049/tcp
# firewall-cmd --permanent --add-port=2049/udp
# firewall-cmd --reload
```

- c. 启动 NFS 服务。

```
# systemctl start nfs-server
```

流程

- 连接到提供共享存储的主机。在这个示例中，是 **cargo-bay** 主机：

```
# ssh root@cargo-bay
root@cargo-bay's password:
Last login: Mon Sep 24 12:05:36 2019
root~#
```

- 创建存放磁盘镜像的目录，并将与迁移主机共享。

```
# mkdir /var/lib/libvirt/shared-images
```

- 将虚拟机的磁盘镜像从源主机复制到新创建的目录。例如，以下将 **wanderer1** 虚拟机的磁盘镜像复制到‘cargo-bay’ 主机上的 **/var/lib/libvirt/shared-images/** 目录中：

```
# scp /var/lib/libvirt/images/wanderer1.qcow2 root@cargo-bay:/var/lib/libvirt/shared-
images/wanderer1.qcow2
```

- 在您要用于共享存储的主机上，将共享目录添加到 **/etc/exports** 文件中。以下示例与 **source-example** 和 **dest-example** 主机共享 **/var/lib/libvirt/shared-images** 目录：

```
/var/lib/libvirt/shared-images source-example(rw,no_root_squash) dest-
example(rw,no_root_squash)
```

- 在源和目标主机上，将共享目录挂载到 **/var/lib/libvirt/images** 目录中：

```
# mount cargo-bay:/var/lib/libvirt/shared-images /var/lib/libvirt/images
```

验证

- 要验证进程是否成功，在源主机上启动虚拟机并观察它是否正确引导。

其它资源

- [导出 NFS 共享](#)

9.6. 使用命令行界面迁移虚拟机

如果虚拟机所在的当前主机变得不适合或者无法再使用，或者要重新分发托管工作负载，您可以将该虚拟机迁移到另一个 KVM 主机中。以下流程提供了各种迁移场景的说明和示例。

先决条件

- 源主机和目标主机都使用 KVM 管理程序。
- 源主机和目标主机可以通过网络相互访问。使用 **ping** 工具进行验证。
- 确保目标主机上打开了以下端口：
 - 使用 SSH 连接到目标主机需要端口 22。
 - 使用 TLS 连接到目标主机需要端口 16509。
 - 使用 TCP 连接到目标主机需要端口 16514。
 - QEMU 需要端口 49152-49215 来传输内存和磁盘迁移数据。
- 对于红帽支持的迁移，源主机和目标主机必须使用特定的操作系统和机器类型。要确保情况是这种情况，请查看 [支持的虚拟机迁移主机](#)。
- 虚拟机必须与目标主机的 CPU 功能兼容。要确保情况如此，请参阅[验证虚拟机迁移的主机 CPU 兼容性](#)。
- 要迁移的虚拟机的磁盘镜像位于源主机和目标主机都可访问的单独的网络位置。这在离线迁移中是可选的，但在迁移运行的虚拟机时是必需的。
有关设置这样的共享虚拟机存储的步骤，请参阅 [与其他主机 共享虚拟机磁盘镜像](#)。
- 迁移正在运行的虚拟机时，您的网络带宽必须高于虚拟机生成脏内存页面的速度。
要在开始实时迁移之前获取虚拟机的脏页面率，请执行以下操作：

- 监控虚拟机在短时间内的脏页面生成率。

```
# virsh domdirtyrate-calc vm-name 30
```

- 在监控完成后，获取其结果：

```
# virsh domstats vm-name --dirtyrate
Domain: 'vm-name'
dirtyrate.calc_status=2
dirtyrate.calc_start_time=200942
dirtyrate.calc_period=30
dirtyrate.megabytes_per_second=2
```

在本例中，VM 每秒会生成 2 MB 的脏内存页面。如果您不暂停虚拟机或降低其工作负载，那么在小于或等于 2 MB/秒带宽的网络中尝试实时迁移这样的虚拟机会导致实时迁移不会进行。

为确保实时迁移成功完成，红帽建议您的网络带宽要明显大于虚拟机的脏页面生成率。

- 当迁移公共网桥中现有虚拟机时，源和目标主机必须位于同一网络中。否则，迁移后 VM 网络将不会操作。
- 确保 **libvirtd** 服务启用并在运行。

```
# systemctl enable --now libvirtd.service
```

流程

1. 使用 **virsh migrate** 命令，以及适合您迁移要求的选项。

- 以下命令将 **wanderer1** 虚拟机从本地主机迁移到使用 SSH 隧道的 **dest-example** 主机的系统连接。虚拟机将在迁移过程中继续运行。

```
# virsh migrate --persistent --live wanderer1 qemu+ssh://dest-example/system
```

- 以下操作允许您对运行在本地主机上的 **wanderer2** 虚拟机的配置进行手动调整，然后将虚拟机迁移到 **dest-example** 主机。迁移的虚拟机将自动使用更新的配置。

```
# virsh dumpxml --migratable wanderer2 >wanderer2.xml
# vi wanderer2.xml
# virsh migrate --live --persistent --xml wanderer2.xml wanderer2 qemu+ssh://dest-example/system
```

当目标主机需要使用不同路径访问共享虚拟机存储或配置特定于目标主机的功能时，这个过程很有用。

- 以下命令暂停 **source-example** 主机上的 **wanderer3** 虚拟机，将它迁移到 **dest-example** 主机上，并指示它使用由 **wanderer3-alt.xml** 文件提供的调整的 XML 配置。迁移完成后，**libvirt** 会在目标主机上恢复虚拟机。

```
# virsh migrate wanderer3 qemu+ssh://source-example/system qemu+ssh://dest-example/system --xml wanderer3-alt.xml
```

迁移后，虚拟机在源主机上处于关闭状态，并在关闭后删除迁移的副本。

- 以下命令删除 **source-example** 主机上关闭的 **wanderer4** 虚拟机，并将其配置移到 **dest-example** 主机：

```
# virsh migrate --offline --persistent --undefinesource wanderer4
qemu+ssh://source-example/system qemu+ssh://dest-example/system
```

请注意，这种类型的迁移不需要将虚拟机的磁盘镜像移到共享存储中。但是，为了使虚拟机在目标主机上可用，您还需要迁移虚拟机的磁盘镜像。例如：

```
# scp root@source-example:/var/lib/libvirt/images/wanderer4.qcow2 root@dest-example:/var/lib/libvirt/images/wanderer4.qcow2
```

- 等待迁移完成。这个过程可能需要一些时间，具体要看网络带宽、系统负载和虚拟机的大小。如果 **virsh migrate** 没有使用 **--verbose** 选项，CLI 不会显示任何进度指示器，除了错误。迁移进行时，您可以使用 **virsh domjobinfo** 工具显示迁移统计信息。

验证

- 在目标主机上，列出可用虚拟机以验证虚拟机是否已迁移：

```
# virsh list
Id Name          State
-----
10 wanderer1    running
```

如果迁移仍在运行，这个命令会将虚拟机状态列为 **paused**。

故障排除

- 在某些情况下，目标主机与迁移虚拟机 XML 配置的某些值不兼容，比如网络名称或 CPU 类型。因此，虚拟机将无法在目标主机上引导。要解决这些问题，您可以使用 **virsh edit** 命令更新有问题的值。在更新值后，您必须重启虚拟机以使更改生效。
- 如果实时迁移需要很长时间才能完成，这可能是因为虚拟机负载很重，且有太多的内存页面改变使得实时迁移不可能实现。要解决这个问题，请挂起虚拟机，将迁移改为非实时迁移。

```
# virsh suspend wanderer1
```

其它资源

- virsh migrate --help** 命令
- virsh** 手册页

9.7. 使用 WEB 控制台实时迁移虚拟机

如果要迁移正在执行要求其持续运行任务的虚拟机(VM)，您可以将该虚拟机迁移到另一个 KVM 主机上，而不需要关闭它。这也称为实时迁移。下列说明解释了如何使用 Web 控制台来进行此操作。



警告

如果修改内存页面任务的速度比 KVM 可以传输页面的速度快 (I/O 负载较重的任务)，建议不要实时迁移虚拟机。

先决条件

- Web 控制台 VM 插件 [已安装在您的系统上](#)。
- 源和目标主机都在运行。
- 确保目标主机上打开了以下端口：
 - 使用 SSH 连接到目标主机需要端口 22。
 - 使用 TLS 连接到目标主机需要端口 16509。
 - 使用 TCP 连接到目标主机需要端口 16514。
 - QEMU 需要端口 49152-49215 来传输内存和磁盘迁移数据。
- 虚拟机必须与目标主机的 CPU 功能兼容。要确保情况如此，[请参阅验证虚拟机迁移的主机 CPU 兼容性](#)。
- 虚拟机磁盘镜像位于 [共享存储](#) 上，该存储可被源主机和目标主机访问。
- 迁移正在运行的虚拟机时，您的网络带宽必须高于虚拟机生成脏内存页面的速度。要在开始实时迁移前获得虚拟机的脏页面率，请在命令行界面中执行以下操作：

- 监控虚拟机在短时间内的脏页面生成率。

```
# virsh domdirtyrate-calc vm-name 30
```

- 在监控完成后，获取其结果：

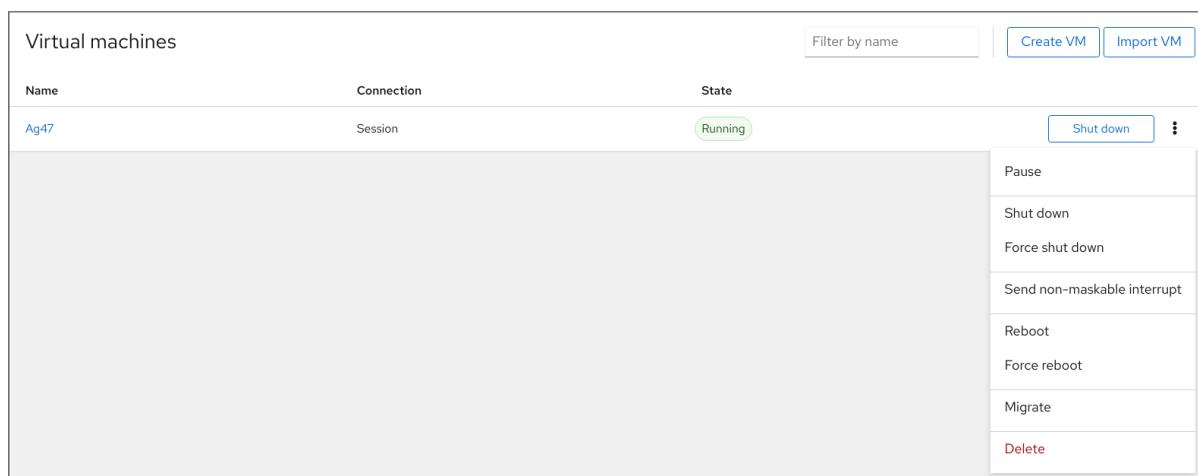
```
# virsh domstats vm-name --dirtyrate
Domain: 'vm-name'
dirtyrate.calc_status=2
dirtyrate.calc_start_time=200942
dirtyrate.calc_period=30
dirtyrate.megabytes_per_second=2
```

在本例中，VM 每秒会生成 2 MB 的脏内存页面。如果您不暂停虚拟机或降低其工作负载，那么在小于或等于 2 MB/秒带宽的网络中尝试实时迁移这样的虚拟机会导致实时迁移不会进行。

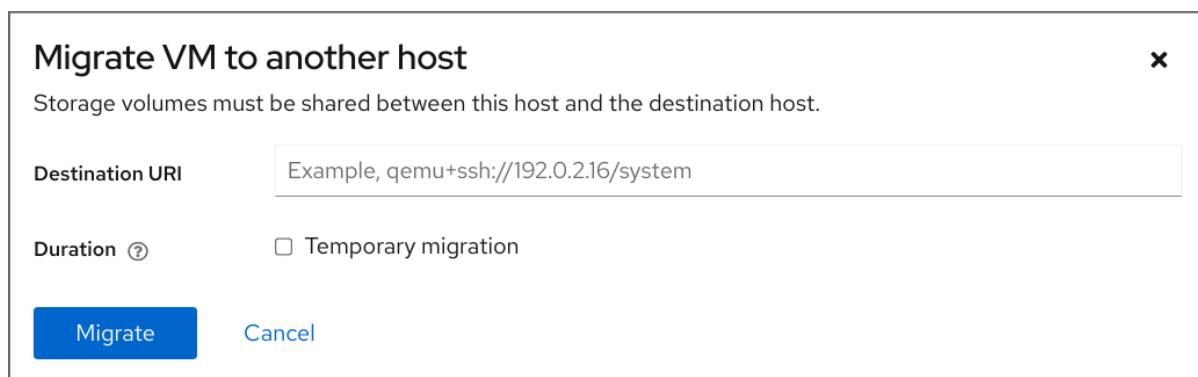
为确保实时迁移成功完成，红帽建议您的网络带宽要明显大于虚拟机的脏页面生成率。

流程

- 在 web 控制台的虚拟机界面中，单击您要迁移的虚拟机的菜单按钮 。此时会出现一个下拉菜单，控制各种虚拟机操作。



- 点击 **Migrate**
此时将显示将虚拟机迁移到另一主机的对话框。



- 输入目标主机的 URI。

4. 配置迁移的持续时间：

- **Permanent** - 如果您要永久迁移虚拟机，请勿选中该框。永久迁移会从源主机上完全删除虚拟机配置。
- **Temporary** - 临时迁移会将虚拟机的副本迁移到目标主机。虚拟机关闭时，此副本将从目标主机中删除。原始虚拟机会保留在源主机上。

5. 点击 **Migrate**

您的虚拟机被迁移到目标主机上。

验证

要验证虚拟机是否已成功迁移并正常工作：

- 确认虚拟机是否在目标主机上可用虚拟机的列表中。
- 启动迁移的虚拟机，并观察其是否启动。

9.8. 虚拟机迁移支持的主机

要使虚拟机迁移正常工作且受红帽支持，源和目标主机必须是特定的 RHEL 版本和机器类型。下表显示了支持的虚拟机迁移路径。

表 9.2. 实时迁移兼容性

迁移方法	发行类型	示例	支持状态
向前	主发行版本	7.6+ → 8.1	在支持的 RHEL 7 系统上：机器类型 i440fx 和 q35
向后	主发行版本	8.1 → 7.6+	在支持的 RHEL 8 系统上：机器类型 i440fx 和 q35
向前	次发行版本	8.0.1+ → 8.1+	在支持的 RHEL 7 系统上：RHEL 7.6.0 及更新版本中的机器类型 i440 和 q35 。 在支持的 RHEL 8 系统中：机器类型 q35 。
向后	次发行版本	8.1 → 8.0.1	在支持的 RHEL 7 系统中。完全支持机器类型 i440fx 和 q35 。 在支持的 RHEL 8 系统中：机器类型 q35 。

其它资源

- [Red Hat Enterprise Linux 生命周期](#)

9.9. 其它资源

- 使用 RHEL 7 和 RHEL 8 中的 **virt-v2v** 将虚拟机从其他 hypervisor 转换为 KVM

第 10 章 管理虚拟设备

管理虚拟机功能、特性和性能的最有效的方法之一是调整其虚拟设备。

以下章节提供了虚拟设备的 [概述](#)，以及如何使用 [CLI](#) 或 [Web 控制台](#) 来管理它们的说明。

10.1. 虚拟设备的工作原理

与物理机一样，虚拟机(VM)也要求专门的设备来为系统提供功能，如处理能力、内存、存储、网络或图形。物理系统通常将硬件设备用于这些目的。但是，由于虚拟机作为软件实现来工作，因此它们需要使用此类设备的软件抽象，称为 [虚拟设备](#)。

基础知识

附加到虚拟机的虚拟设备可在[创建虚拟机时](#)配置，也可以在现有虚拟机上管理。通常，只有在虚拟机处于关闭状态时才能从虚拟机挂载或卸载虚拟设备，但某些设备可在虚拟机运行时进行添加或删除。这个功能被称为设备热插 (hot plug) 和热拔 (hot unplug)。

创建新虚拟机时，[libvirt](#) 会自动创建和配置一组默认的基本虚拟设备，除非用户另有指定。它们基于主机系统架构和机器类型，通常包括：

- CPU
- 内存
- 键盘
- 网络接口控制器 (NIC)
- 各种设备控制器
- 一个视频卡
- 一个声卡

要在创建虚拟机后管理虚拟设备，请使用命令行界面(CLI)。但是，要管理 [虚拟存储设备](#) 和 [NIC](#)，您也可以使用 RHEL 8 web 控制台。

性能或灵活性

对于某些类型的设备，RHEL 8 支持多种实现，通常会在性能和灵活性之间权衡。

例如，用于虚拟磁盘的物理存储可以由各种格式的文件（如 [qcow2](#) 或 [raw](#)）来表示，并通过各种控制器呈现给虚拟机：

- 模拟控制器
- [virtio-scsi](#)
- [virtio-blk](#)

模拟控制器比 [virtio](#) 控制器慢，因为 [virtio](#) 设备是专为虚拟化设计的。另一方面，仿真控制器可以运行没有 [virtio](#) 设备驱动程序的操作系统。同样，[virtio-scsi](#) 为 SCSI 命令提供更完整的支持，并可以将更多磁盘附加到虚拟机。最后，[virtio-blk](#) 比 [virtio-scsi](#) 和仿真控制器提供更好的性能，但用例范围比较有限。例如，在使用 [virtio-blk](#) 时，无法将物理磁盘作为 LUN 设备附加到虚拟机。

有关虚拟设备类型的更多信息，请参阅 [虚拟设备的类型](#)。

其它资源

- 为虚拟机管理存储
- 使用 web 控制台管理虚拟机网络接口
- 管理 NVIDIA vGPU 设备

10.2. 虚拟设备类型

RHEL 8 中的虚拟化可以提供几种不同类型的虚拟设备，您可以附加到虚拟机(VM)：

模拟设备

模拟设备是广泛使用的物理设备的软件实现。为物理设备设计的驱动程序还与模拟设备兼容。因此可非常灵活地使用模拟设备。

但是，由于需要忠实地模拟特定类型的硬件，与相应的物理设备或更为优化的虚拟设备相比，模拟设备可能会遭受显著的性能损失。

支持以下模拟设备类型：

- 虚拟 CPU(vCPU)，有大量 CPU 型号可供选择。模拟的性能影响显著取决于主机 CPU 和模拟 vCPU 之间的差异。
- 模拟系统组件，如 PCI 总线控制器。
- 模拟存储控制器，如 SATA、SCSI 甚至 IDE。
- 模拟声音设备，如 ICH9、ICH6 或 AC97。
- 模拟图形卡，如 VGA 或 QXL 卡。
- 模拟网络设备，如 rtl8139。

半虚拟设备

半虚拟 (Paravirtualization) 提供了向虚拟机公开虚拟设备的速度。半虚拟设备公开专用于虚拟机使用的接口，因此可显著提高设备的性能。RHEL 8 使用 *virtio* API 作为 hypervisor 和虚拟机之间的一个层，来为虚拟机提供半虚拟化设备。这个方法的缺陷在于它需要在客户端操作系统中使用特定的设备驱动程序。

建议您尽可能为虚拟机使用半虚拟设备而不是模拟设备，特别是当它们运行大量 I/O 的应用程序时。半虚拟设备减少 I/O 延迟并增加 I/O 吞吐量，在某些情况下可使其非常接近裸机性能。其它半虚拟设备还会在不能使用的虚拟机中添加功能。

支持以下半虚拟设备类型：

- 半虚拟化网络设备(**virtio-net**)。
- 半虚拟化存储控制器：
 - **virtio-blk** - 提供块设备模拟。
 - **virtio-scsi** - 提供更完整的 SCSI 模拟。
- 半虚拟时钟。
- 半虚拟化串行设备(**virtio-serial**)。

- 气球 (balloon) 设备(**virtio-balloon**)用于在虚拟机及其主机之间动态分配内存。
- 半虚拟化随机数字生成器(**virtio-rng**)。
- 半虚拟化图形卡(**QXL**)。

物理共享设备

某些硬件平台可让虚拟机直接访问各种硬件设备和组件。这个过程被称为设备分配或者透传 (*passthrough*)。

以这种方式连接时，物理设备的某些方面可以直接用于虚拟机，就像它们可用于物理机一样。这为虚拟机中使用的设备提供了出众的性能。但是，物理连接到虚拟机的设备对主机不可用，也无法迁移。

然而，某些设备可以在多个虚拟机之间 共享。例如，一个物理设备在某些情况下可以提供多个介质设备，然后将其分配给不同的虚拟机。

支持以下 passthrough 设备类型：

- USB、PCI 和 SCSI 直通 - 直接向虚拟机公开通用的行业标准总线，以便其特定功能对虚拟客户机软件可用。
- 单根 I/O 虚拟化(SR-IOV)- 一种支持 PCI Express 资源的硬件强制隔离的规范。这使得将单个物理 PCI 资源划分成虚拟 PCI 功能变得安全而高效。它通常用于网络接口卡(NIC)。
- N_Port ID 虚拟化(NPIV)- 一种光纤通道技术，来与多个虚拟端口共享单个物理主机总线适配器 (HBA)。
- GPU 和 vGPU - 用于特定图形或计算工作负载的加速器。一些 GPU 可以直接连接到虚拟机，而某些类型也提供能够创建共享底层物理硬件的虚拟 GPU(vGPU)的能力。

10.3. 使用 CLI 管理附加到虚拟机的设备

要修改虚拟机的功能，您可以使用命令行界面(CLI)管理附加到虚拟机的设备。

您可以使用 CLI 来：

- [附加设备](#)
- [修改设备](#)
- [删除设备](#)

10.3.1. 将设备附加到虚拟机

您可以通过附加新的虚拟设备来向虚拟机(VM)添加特定的功能。

以下流程演示了如何使用命令行界面(CLI)来创建虚拟设备并将其附加到虚拟机(VM)。一些设备也可以 [使用 RHEL web 控制台](#) 附加到虚拟机。

例如，您可以通过将新虚拟磁盘设备附加到虚拟机来增加虚拟机的存储容量。这也被称为**内存热插拔**。



警告

RHEL 8 不支持从虚拟机中移除内存设备，也称为 **内存热插拔**，红帽强烈反对使用它。

先决条件

- 获取您要附加到虚拟机的设备所需的选项。要查看特定设备的可用选项，请使用 **virt-xml -device=?** 命令。例如：

```
# virt-xml --network=?
--network options:
[...]
address.unit
boot_order
clearxml
driver_name
[...]
```

流程

1. 要将设备附加到虚拟机，请使用 **virt-xml --add-device** 命令，包括该设备的定义和所需的选项：

- 例如，以下命令会在 **/var/lib/libvirt/images/** 目录中创建一个 20GB 的 *newdisk* qcow2 磁盘镜像，并在虚拟机启动时将其作为虚拟磁盘附加到运行的 *testguest* 虚拟机上：

```
# virt-xml testguest --add-device --disk
/var/lib/libvirt/images/newdisk.qcow2,format=qcow2,size=20
Domain 'testguest' defined successfully.
Changes will take effect after the domain is fully powered off.
```

- 在虚拟机运行时，下面的命令会将一个 USB 盘（在主机的 002 总线中作为设备 004）附加到 *testguest2* 虚拟机：

```
# virt-xml testguest2 --add-device --update --hostdev 002.004
Device hotplug successful.
Domain 'testguest2' defined successfully.
```

可以使用 **lssusb** 命令获取用于定义 USB 的总线-设备组合。

验证

要验证设备是否已被添加，请执行以下操作之一：

- 使用 **virsh dumpxml** 命令，并查看设备的 XML 定义是否已添加到虚拟机 XML 配置中的 **<devices>** 部分。

例如，以下输出显示了 *testguest* 虚拟机的配置，并确认已添加了 002.004 USB 闪存磁盘设备。

```
# virsh dumpxml testguest
[...]
```

```
<hostdev mode='subsystem' type='usb' managed='yes'>
  <source>
    <vendor id='0x4146'/>
    <product id='0x902e'/>
    <address bus='2' device='4'/>
  </source>
  <alias name='hostdev0'/>
  <address type='usb' bus='0' port='3'/>
</hostdev>
[...]
```

- 运行虚拟机并测试该设备是否存在并正常工作。

其它资源

- man virt-xml** 命令

10.3.2. 修改附加到虚拟机的设备

您可以通过编辑附加的虚拟设备的配置来更改虚拟机(VM)的功能。例如，如果想要优化虚拟机的性能，您可以更改其虚拟 CPU 型号来更好地匹配主机的 CPU。

以下流程提供了使用命令行界面(CLI)修改虚拟设备的常用说明。附加到虚拟机的一些设备，如磁盘和 NIC，也可以使用 [RHEL 8 web 控制台](#) 进行修改。

先决条件

- 获取您要附加到虚拟机的设备所需的选项。要查看特定设备的可用选项，请使用 **virt-xml -device=?** 命令。例如：

```
# virt-xml --network=?
--network options:
[...]
address.unit
boot_order
clearxml
driver_name
[...]
```

- 可选：使用 **virsh dumpxml vm-name** 备份虚拟机的 XML 配置，并将结果发送到文件中。例如，以下方法将 Motoko 虚拟机的配置备份为 **motoko.xml** 文件：

```
# virsh dumpxml Motoko > motoko.xml
# cat motoko.xml
<domain type='kvm' xmlns:qemu='http://libvirt.org/schemas/domain/qemu/1.0'>
  <name>Motoko</name>
  <uuid>ede29304-fe0c-4ca4-abcd-d246481acd18</uuid>
  [...]
</domain>
```

流程

- 使用 **virt-xml --edit** 命令，包括设备的定义以及所需的选项：

例如，以下可清除关闭的 *testguest* 虚拟机的 **<cpu>** 配置，并将其设置为 *host-model*：

```
# virt-xml testguest --edit --cpu host-model,clearxml=yes
Domain 'testguest' defined successfully.
```

验证

要校验设备已被修改，请执行以下任一操作：

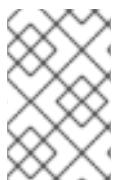
- 运行虚拟机并测试该设备是否存在并反映了所做的修改。
- 使用 **virsh dumpxml** 命令，并查看在虚拟机的 XML 配置中是否已修改了设备的 XML 定义。例如，以下输出显示了 *testguest* 虚拟机的配置，并确认 CPU 模式已被配置为 *host-model*。

```
# virsh dumpxml testguest
[...]
<cpu mode='host-model' check='partial'>
  <model fallback='allow'/>
</cpu>
[...]
```

故障排除

- 如果修改设备会导致虚拟机无法启动，请使用 **virsh** 定义工具通过重新加载之前备份的 XML 配置文件来恢复 XML 配置。

```
# virsh define testguest.xml
```



注意

对于对虚拟机 XML 配置的小更改，您可以使用 **virsh edit** 命令 - 例如 **virsh edit testguest**。但是，对于更广泛的更改，请勿使用此方法，因为它很可能会破坏配置，从而阻止虚拟机启动。

其它资源

- man virt-xml** 命令

10.3.3. 从虚拟机中删除设备

您可以通过移除虚拟设备来更改虚拟机(VM)的功能。例如，如果不再需要，您可以从其中一个虚拟机中删除虚拟磁盘设备。

以下流程演示了如何使用命令行界面(CLI)从虚拟机(VM)中移除虚拟设备。还可[使用 RHEL 8 web 控制台](#)从虚拟机中删除一些设备，如磁盘或 NIC。

先决条件

- 可选：使用 **virsh dumpxml vm-name** 备份虚拟机的 XML 配置，并将结果发送到文件中。例如，以下方法将 *Motoko* 虚拟机的配置备份为 **motoko.xml** 文件：

```
# virsh dumpxml Motoko > motoko.xml
# cat motoko.xml
<domain type='kvm' xmlns:qemu='http://libvirt.org/schemas/domain/qemu/1.0'>
  <name>Motoko</name>
```

```
<uuid>ede29304-fe0c-4ca4-abcd-d246481acd18</uuid>
[...]
</domain>
```

流程

- 使用 **virt-xml --remove-device** 命令，包括设备的定义。例如：

- 以下会在运行的 *testguest* 虚拟机关闭后，从其中删除标记为 *vdb* 的存储设备：

```
# virt-xml testguest --remove-device --disk target=vdb
Domain 'testguest' defined successfully.
Changes will take effect after the domain is fully powered off.
```

- 以下命令会立即从运行的 *testguest2* 虚拟机中删除 USB 闪存驱动器设备：

```
# virt-xml testguest2 --remove-device --update --hostdev type=usb
Device hotunplug successful.
Domain 'testguest2' defined successfully.
```

故障排除

- 如果移除设备会导致虚拟机无法启动，请使用 **virsh define** 工具，通过重新加载之前备份的 XML 配置文件来恢复 XML 配置。

```
# virsh define testguest.xml
```

其它资源

- [man virt-xml 命令](#)

10.4. 使用 WEB 控制台管理主机设备

要修改虚拟机的功能，您可以使用 RHEL 8 web 控制台管理附加到虚拟机的主机设备。

主机设备是附加到主机系统的物理设备。根据您的需求，您可以让虚拟机直接访问这些硬件设备和组件。

您可以使用 Web 控制台：

- [查看设备](#)
- [附加设备](#)
- [删除设备](#)

10.4.1. 使用 web 控制台查看附加到虚拟机的设备

在添加或修改附加到虚拟机(VM)的设备之前，您可能需要查看已附加到虚拟机的设备。以下流程提供了使用 Web 控制台查看这些设备的说明。

先决条件

- [Web 控制台 VM 插件 已安装在您的系统上。](#)

流程

- 在 **Virtual Machines** 界面中，点击您要查看其信息的虚拟机。
这会打开一个新页面，其中包含有关虚拟机的详细信息。

- 滚动到 **Host devices** 部分。

Host devices					Add host device
Type	Class	Model	Vendor	Source	
usb		Bluetooth wireless interface	Intel Corp.	Device 002 Bus 001	Remove

其它资源

- [管理虚拟设备](#)

10.4.2. 使用 web 控制台将设备附加到虚拟机

要为虚拟机添加特定的功能，您可以使用 web 控制台将主机设备附加到虚拟机。



注意

不能同时附加多个主机设备。一次只能附加一个设备。

如需更多信息，请参阅 [RHEL 8 已知问题](#)。

先决条件

- 如果要附加 PCI 设备，请确保 **hostdev** 元素的 **managed** 属性的状态被设为 **yes**。



注意

将 PCI 设备附加到虚拟机时，请不要省略 **hostdev** 元素的 **managed** 属性，或者将其设为 **no**。如果您这样做了，当把 PCI 设备传给虚拟机时，会无法自动将 PCI 设备与主机分离。当您关闭虚拟机时，也不能自动将它们重新附加到主机上。

因此，主机可能会变得无响应，或者意外关闭。

您可以在虚拟机的 XML 配置中找到 **managed** 属性的状态。以下示例打开了 Ag47 虚拟机的 XML 配置：

```
# virsh edit Ag47
```

- 备份虚拟机中的重要数据。
- 可选：备份虚拟机的 XML 配置。例如，备份 **Centurion** 虚拟机：

```
# virsh dumpxml Centurion > Centurion.xml
```

- Web 控制台 VM 插件 [已安装在您的系统上](#)。

流程

1. 在 **Virtual Machines** 接口中，点击您要给其附加主机设备的虚拟机。
这会打开一个新页面，其中包含有关所选虚拟机基本信息的 **概述**，以及访问虚拟机图形界面的 **控制台** 部分。
2. 滚动到 **Host devices**。
主机设备 部分显示有关附加到虚拟机的设备的信息，以及 **A添加** 或 **删除** 设备的选项。

Host devices					Add host device
Type	Class	Model	Vendor	Source	
usb		Bluetooth wireless interface	Intel Corp.	Device Bus	002 001
					Remove

3. 点击 **Add host device**。
此时会出现 **Add host device** 对话框。

Add host device

Type USB PCI

Device	Product	Vendor	Location
<input type="checkbox"/> Card Reader	Realtek Semiconductor Corp.	Device Bus	002 002
<input type="checkbox"/> 3.0 root hub	Linux Foundation	Device Bus	001 004
<input type="checkbox"/> Bluetooth wireless interface	Intel Corp.	Device Bus	002 001
<input type="checkbox"/> 2.0 root hub	Linux Foundation	Device Bus	001 003
<input type="checkbox"/> Integrated Camera (1280x720@30)	Chicony Electronics Co., Ltd	Device	003

[Add](#) [Cancel](#)

4. 选择您要附加到虚拟机的设备。

5. 点 **添加**

所选设备被附加到虚拟机。

验证

- 运行虚拟机并检查该设备是否出现在 **Host devices** 部分中。

10.4.3. 使用 web 控制台从虚拟机中删除设备

要释放资源，修改虚拟机的功能或两个都做，您可以使用 Web 控制台来修改虚拟机，并删除不再需要的主机设备。



警告

使用 Web 控制台删除连接的 USB 主机设备可能会失败，因为 USB 设备的设备和总线号关联错误。

如需更多信息，请参阅 [RHEL 8 已知问题](#)。

作为临时解决方案，使用 "virsh" 工具从虚拟机的 XML 配置中删除 USB 设备的 <hostdev> 部分。以下示例打开了 Ag47 虚拟机的 XML 配置：

```
# virsh edit Ag47
```

先决条件

- Web 控制台 VM 插件 [已安装在您的系统上](#)。
- 可选：使用 **virsh dumpxml vm-name** 备份虚拟机的 XML 配置，并将结果发送到文件中。例如，以下方法将 Motoko 虚拟机的配置备份为 **motoko.xml** 文件：

```
# virsh dumpxml Motoko > motoko.xml
# cat motoko.xml
<domain type='kvm' xmlns:qemu='http://libvirt.org/schemas/domain/qemu/1.0'>
  <name>Motoko</name>
  <uuid>ede29304-fe0c-4ca4-abcd-d246481acd18</uuid>
  [...]
</domain>
```

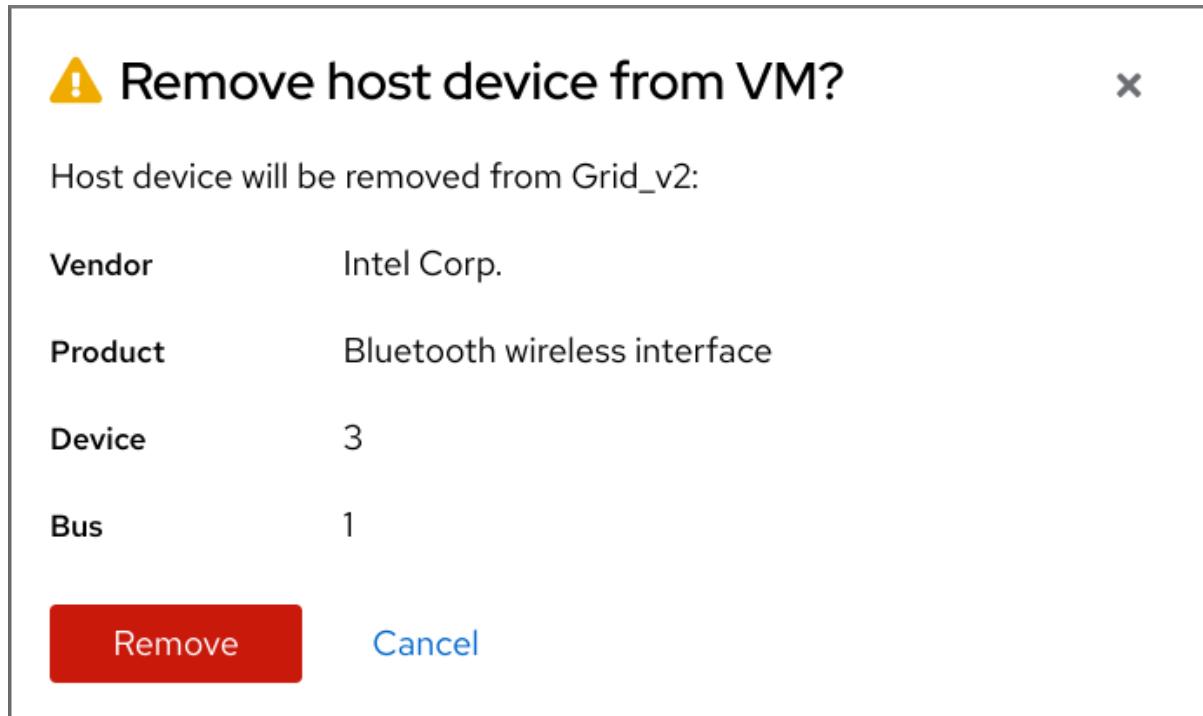
流程

- 在 **Virtual Machines** 接口中，点击您要从中删除主机设备的虚拟机。这会打开一个新页面，其中包含有关所选虚拟机基本信息的 **概述**，以及访问虚拟机图形界面的 **控制台** 部分。
- 滚动到 **Host devices**。
主机设备 部分显示有关附加到虚拟机的设备的信息，以及 **A添加** 或 **删除** 设备的选项。

Host devices					Add host device
Type	Class	Model	Vendor	Source	
usb		Bluetooth wireless interface	Intel Corp.	Device 002 Bus 001	Remove

3. 点击您要从虚拟机中删除设备旁边的 **Remove** 按钮。

此时会出现删除设备确认对话框。



4. 点击 **Remove**。

该设备从虚拟机中删除。

故障排除

- 如果删除主机设备导致虚拟机无法启动，请使用 **virsh define** 工具通过重新载入之前备份的 XML 配置文件来恢复 XML 配置。

```
# virsh define motoko.xml
```

10.5. 管理虚拟 USB 设备

使用虚拟机(VM)时，您可以访问并控制 USB 设备，如附加到主机系统的如闪存驱动器或 Web 相机。在这种情况下，主机系统会将设备的控制权传递给虚拟机。这也被称为 USB-passthrough。

以下部分提供有关使用命令行的信息：

- [将 USB 设备附加到虚拟机](#)
- [从虚拟机中删除 USB 设备](#)

10.5.1. 将 USB 设备附加到虚拟机

要将 USB 设备附加到虚拟机，您可以在虚拟机 XML 配置文件中包含 USB 设备信息。

先决条件

- 确定您要传递给虚拟机的设备已附加到主机。

流程

- 找到您要附加到虚拟机的 USB 总线和设备值。

例如：以下命令显示附加到该主机的 USB 设备列表。在这个示例中，使用的设备作为设备 005 总线附加到总线 001 中。

```
# lsusb
[...]
Bus 001 Device 003: ID 2567:0a2b Intel Corp.
Bus 001 Device 005: ID 0407:6252 Kingston River 2.0
[...]
```

- 使用 **virt-xml** 工具及 **--add-device** 参数。

例如，以下命令将 USB 闪存驱动器附加到 **Library** 虚拟机。

```
# virt-xml Library --add-device --hostdev 001.005
Domain 'Library' defined successfully.
```



注意

要将 USB 设备连接到正在运行的虚拟机，请将 **--update** 参数添加到上一命令中。

验证

- 运行虚拟机并测试该设备是否存在并正常工作。
- 使用 **virsh dumpxml** 命令查看设备的 XML 定义是否已添加到虚拟机 XML 配置文件中的 **<devices>** 部分。

```
# virsh dumpxml Library
[...]
<hostdev mode='subsystem' type='usb' managed='yes'>
  <source>
    <vendor id='0x0407' />
    <product id='0x6252' />
    <address bus='1' device='5' />
  </source>
  <alias name='hostdev0' />
  <address type='usb' bus='0' port='3' />
</hostdev>
[...]
```

其它资源

- man virt-xml** 命令
- [将设备附加到虚拟机](#)

10.5.2. 从虚拟机中删除 USB 设备

要从虚拟机中删除 USB 设备，您可以从虚拟机 XML 配置中删除 USB 设备信息。

流程

- 找到您要从虚拟机中删除的 USB 的 bus 和 device 值。

例如：以下命令显示附加到该主机的 USB 设备列表。在这个示例中，使用的设备作为设备 005 总线附加到总线 001 中。

```
# lsusb
[...]
Bus 001 Device 003: ID 2567:0a2b Intel Corp.
Bus 001 Device 005: ID 0407:6252 Kingston River 2.0
[...]
```

- 使用 **virt-xml** 工具及 **--remove-device** 参数。

例如，以下命令会从 **Library** 虚拟机中删除一个 USB 闪存驱动器，该驱动器作为总线 001 上的设备 005 附加到主机。

```
# virt-xml Library --remove-device --hostdev 001.005
Domain 'Library' defined successfully.
```



注意

要从正在运行的虚拟机中删除 USB 设备，请将 **--update** 参数添加到上一命令中。

验证

- 运行虚拟机并检查该设备是否已从设备列表中删除。

其它资源

- man virt-xml** 命令
- [将设备附加到虚拟机](#)

10.5.3. 将智能卡读卡器附加到虚拟机

如果智能卡读卡器已附加到主机，您也可以将其提供给该主机上的虚拟机(VM)。libvirt 提供了一个专用的虚拟设备，为客户虚拟机提供了一个智能卡接口。建议您只使用 **spicevmc** 设备类型，它利用 SPICE 远程显示协议将身份验证请求隧道传输到主机。

虽然可以将标准设备直通与智能卡读卡器一起使用，但此方法无法使设备在主机和客户机系统上可用。因此，在将智能卡读卡器附加到虚拟机时，您可以锁定主机系统。



重要

SPICE 远程显示协议已在 RHEL 8 中被弃用。由于将智能卡读卡器附加到虚拟机的唯一方法取决于 SPICE 协议，因此客户虚拟机中智能卡的使用也已在 RHEL 8 中被弃用。

在 RHEL 未来的主发行版本中，将智能卡读卡器附加到虚拟机的功能仅被第三方远程视觉解决方案支持。

先决条件

- 确定您要传给虚拟机的智能卡读卡器已附加到主机上。
- 确保智能卡读卡器类型 [在 RHEL 8 中被支持](#)。

流程

- 创建并将虚拟智能卡读卡器设备附加到虚拟机上。例如，要将智能卡读卡器附加到 **testguest** 虚拟机上：

```
# virt-xml testguest --add-device --smartcard mode=passthrough,type=spicevmc
Domain 'testguest' defined successfully.
Changes will take effect after the domain is fully powered off.
```



注意

要将虚拟智能卡读卡器设备附加到正在运行的虚拟机上，请在上一个命令中添加 **--update** 参数。

验证

1. 查看虚拟机的 XML 配置。

```
# virsh dumpxml testguest
```

2. 确保 XML 配置包含以下智能卡设备定义。

```
<smartcard mode='passthrough' type='spicevmc'/>
```

10.6. 管理虚拟光驱

当使用虚拟机时，您可以访问保存在主机中 ISO 镜像中的信息。要做到这一点，请将 ISO 镜像作为虚拟光驱附加到虚拟机，比如 CD 驱动器或者 DVD 驱动器。

以下部分提供有关使用命令行的信息：

- [将驱动器和 ISO 镜像附加到虚拟机](#)
- [替换虚拟光驱中的 ISO 镜像](#)
- [从虚拟光驱中删除 ISO 镜像](#)
- [从虚拟机中删除驱动器](#)

10.6.1. 为虚拟机附加光驱

要将 ISO 镜像作为虚拟光驱附加，请编辑虚拟机的 XML 配置文件并添加新驱动器。

先决条件

- 您必须将 ISO 镜像保存在本地主机中。
- 您必须知道 ISO 镜像的路径。

流程

- 使用 **virt-xml** 工具及 **--add-device** 参数。

例如，以下命令将存储在 **/MC/tank/** 目录中的 **Doc10** ISO 镜像附加到 **DN1** 虚拟机上。

```
# virt-xml DN1 --add-device --disk /MC/tank/Doc10.iso,device=cdrom
Domain 'DN1' defined successfully.
```

验证

- 运行虚拟机并测试该设备是否存在并正常工作。

其它资源

- [man virt-xml 命令](#)
- [将设备附加到虚拟机](#)

10.6.2. 使用虚拟光驱替换 ISO 镜像

要替换作为虚拟光驱附加到虚拟机(VM)的 ISO 镜像，请编辑虚拟机的 XML 配置文件，并指定替换。

先决条件

- 您必须将 ISO 镜像保存在本地主机中。
- 您必须知道 ISO 镜像的路径。

流程

- 定位 CD-ROM 附加到虚拟机的目标设备。您可以在虚拟机 XML 配置文件中找到这些信息。例如，以下命令显示 **DN1** 虚拟机的 XML 配置文件，其中 CD-ROM 的目标设备是 **sda**：

```
# virsh dumpxml DN1
...
<disk>
...
<source file='/MC/tank/Doc10.iso' />
<target dev='sda' bus='sata' />
...
</disk>
...
```

- 使用 **virt-xml** 工具及 **--edit** 参数。

例如，以下命令将附加到 **DN1** 虚拟机的目标 **sda** 处的 **Doc10** ISO 镜像替换为存储在 **/Dvrs/current/** 目录中的 **DrDN** ISO 镜像。

```
# virt-xml DN1 --edit target=sda --disk /Dvrs/current/DrDN.iso
Domain 'DN1' defined successfully.
```

验证

- 运行虚拟机并测试是否替换该设备并正常工作。

其它资源

- **man virt-xml** 命令

10.6.3. 从虚拟光驱中删除 ISO 镜像

要从附加到虚拟机(VM)的虚拟光驱中删除 ISO 镜像, 请编辑虚拟机的 XML 配置文件。

流程

1. 定位 CD-ROM 附加到虚拟机的目标设备。您可以在虚拟机 XML 配置文件中找到这些信息。例如, 以下命令显示 **DN1** 虚拟机的 XML 配置文件, 其中 CD-ROM 的目标设备是 **sda** :

```
# virsh dumpxml DN1
...
<disk>
...
<source file='/Dvrs/current/DrDN' />
<target dev='sda' bus='sata' />
...
</disk>
...
```

2. 使用 **virt-xml** 工具及 **--edit** 参数。

例如, 以下命令将从附加到 **DN1** 虚拟机的 CD 驱动器中删除 **DrDN** ISO 镜像。

```
# virt-xml DN1 --edit target=sda --disk path=
Domain 'DN1' defined successfully.
```

验证

- 运行虚拟机, 检查镜像已不再可用。

其它资源

- **man virt-xml** 命令

10.6.4. 从虚拟机中删除光驱

要删除附加到虚拟机的光驱, 编辑虚拟机的 XML 配置文件。

流程

1. 定位 CD-ROM 附加到虚拟机的目标设备。您可以在虚拟机 XML 配置文件中找到这些信息。例如, 以下命令显示 **DN1** 虚拟机的 XML 配置文件, 其中 CD-ROM 的目标设备是 **sda** :

```
# virsh dumpxml DN1
...
<disk type='file' device='cdrom' />
<driver name='qemu' type='raw' />
<target dev='sda' bus='sata' />
```

```

...
</disk>
...

```

2. 使用带 **--remove-device** 参数的**virt-xml** 工具。

例如，以下命令会从 DN1 虚拟机中删除作为目标 **sda** 附加的光驱。

```
# virt-xml DN1 --remove-device --disk target=sda
Domain 'DN1' defined successfully.
```

验证

- 确认该设备不再列在虚拟机 XML 配置文件中。

其它资源

- man virt-xml** 命令

10.7. 管理 SR-IOV 设备

模拟虚拟设备通常使用比硬件网络设备更多的 CPU 和内存。这可能会限制虚拟机的性能。但是，如果虚拟化主机上的任何设备都支持单根 I/O 虚拟化(SR-IOV)，您可以使用此功能来提高设备性能，并可能还提高虚拟机的整体性能。

10.7.1. 什么是 SR-IOV？

单根 I/O 虚拟化(SR-IOV)是一种规范，它允许单个 PCI Express(PCIe)设备向主机系统呈现多个独立的 PCI 设备，称为 **虚拟功能(VF)**。这样的每个设备：

- 提供与原始 PCI 设备相同的或类似的服务。
- 出现在主机 PCI 总线的不同地址上。
- 可使用 VFIO 分配功能分配到不同的虚拟机。

例如，单个具有 SR-IOV 的网络设备可以向多个虚拟机显示 VF。虽然所有 VF 都使用相同的物理卡、相同的网络连接和相同的网线，但每个虚拟机都直接控制其自己的硬件网络设备，并且不使用主机的额外资源。

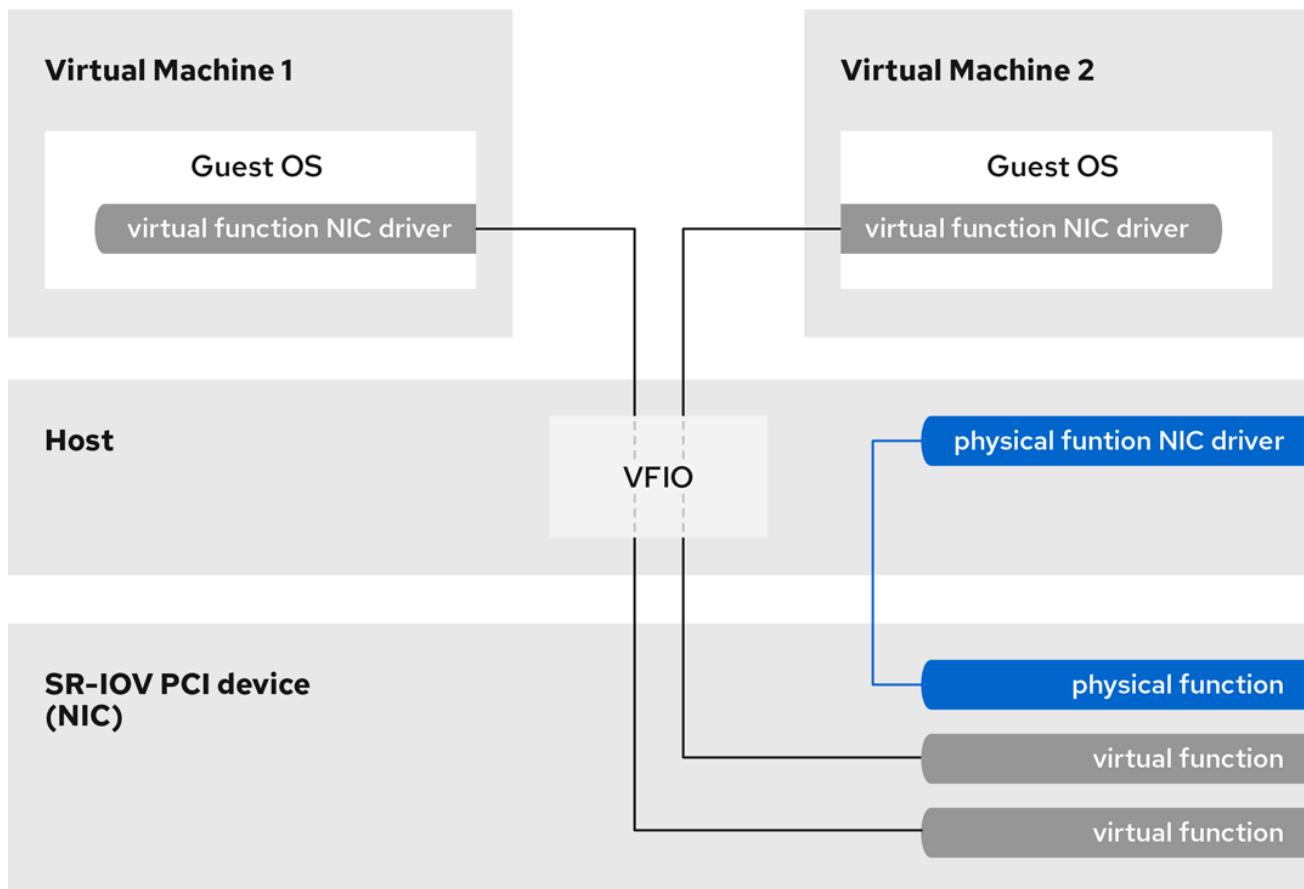
SR-IOV 的工作原理

SR-IOV 功能可能会因为引进了以下 PCI 功能：

- 物理功能(PF)** - 为主机提供设备（如网络）功能的 PCIe 功能，但也可以创建和管理一组 VF。每个具有 SR-IOV 功能的设备都有一个或多个 PF。
- 虚拟功能(VF)** - 充当独立设备的轻量级 PCIe 功能。每个 VF 都是从 PF 中派生的。一个设备可依赖于设备硬件的最大 VF 数。每个 VF 每次只能分配给一个虚拟机，但虚拟机可以分配多个 VF。

VM 将 VF 识别为虚拟设备。例如，由 SR-IOV 网络设备创建的 VF 显示为分配给虚拟机的网卡，其方式与主机系统上显示的物理网卡相同。

图 10.1. SR-IOV 架构



优点

使用 SR-IOV VF 而不是模拟设备的主要优点是：

- 提高的性能
- 减少主机 CPU 和内存资源使用量

例如，作为 vNIC 附加到虚拟机的 VF 性能几乎与物理 NIC 相同，并且优于半虚拟化或模拟的 NIC。特别是，当多个 VF 在单个主机上同时使用时，其性能优势可能非常显著。

缺点

- 要修改 PF 的配置，您必须首先将 PF 公开的 VF 数量改为零。因此，您还需要将这些 VF 提供的设备从分配给虚拟机的设备中删除。
- 附加了 VFIO 分配设备的虚拟机（包括 SR-IOV VF）无法迁移到另一台主机。在某些情况下，您可以通过将分配的设备与模拟的设备进行配对来临时解决这个限制。例如，您可以将分配的网络 VF 绑定到模拟的 vNIC 中，并在迁移前删除 VF。
- 另外，分配了 VFIO 的设备需要固定虚拟机内存，这会增加虚拟机的内存消耗，并防止在虚拟机上使用内存膨胀。

其它资源

- [SR-IOV 分配支持的设备](#)
- [在 IBM Z 中配置透传 PCI 设备](#)

10.7.2. 将 SR-IOV 网络设备附加到虚拟机

要将 SR-IOV 网络设备附加到 Intel 或 AMD 主机上的虚拟机(VM)，您必须从主机上支持 SR-IOV 的网络接口创建一个虚拟功能(VF)，并将 VF 作为设备分配给指定虚拟机。详情请查看以下步骤。

先决条件

- 您的主机的 CPU 和固件支持 I/O 内存管理单元(IOMMU)。
 - 如果使用 Intel CPU，它必须支持 Intel 的直接 I/O 虚拟化技术(VT-d)。
 - 如果使用 AMD CPU，则必须支持 AMD-Vi 功能。
- 主机系统使用访问控制服务(ACS)来为 PCIe 拓扑提供直接内存访问(DMA)隔离。与系统供应商一起验证这一点。
如需更多信息，请参阅[实施 SR-IOV 的硬件注意事项](#)。
- 物理网络设备支持 SR-IOV。要验证系统上的任何网络设备是否支持 SR-IOV，请使用 **lspci -v** 命令，并在输出中查找 **单根 I/O 虚拟化(SR-IOV)**。

```
# lspci -v
[...]
02:00.0 Ethernet controller: Intel Corporation 82576 Gigabit Network Connection (rev 01)
  Subsystem: Intel Corporation Gigabit ET Dual Port Server Adapter
  Flags: bus master, fast devsel, latency 0, IRQ 16, NUMA node 0
  Memory at fcba0000 (32-bit, non-prefetchable) [size=128K]
[...]
  Capabilities: [150] Alternative Routing-ID Interpretation (ARI)
  Capabilities: [160] Single Root I/O Virtualization (SR-IOV)
  Kernel driver in use: igb
  Kernel modules: igb
[...]
```

- 用于创建 VF 的主机网络接口正在运行。例如：要激活 `eth1` 接口并验证它正在运行：

```
# ip link set eth1 up
# ip link show eth1
8: eth1: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc mq state UP mode
  DEFAULT qlen 1000
    link/ether a0:36:9f:8f:3f:b8 brd ff:ff:ff:ff:ff:ff
    vf 0 MAC 00:00:00:00:00:00, spoof checking on, link-state auto
    vf 1 MAC 00:00:00:00:00:00, spoof checking on, link-state auto
    vf 2 MAC 00:00:00:00:00:00, spoof checking on, link-state auto
    vf 3 MAC 00:00:00:00:00:00, spoof checking on, link-state auto
```

- 要使 SR-IOV 设备分配正常工作，必须在主机 BIOS 和内核中启用 IOMMU 功能。要做到这一点：

- 在 Intel 主机上启用 VT-d：

- 使用 `intel_iommu=on` 和 `iommu=pt` 参数重新生成 GRUB 配置：

```
# grub2 --args="intel_iommu=on iommu=pt" --update-kernel=ALL
```

- 重启主机。

- 在 AMD 主机上启用 AMD-Vi：
 - i. 使用 **iommu=pt** 参数重新生成 GRUB 配置：

```
# grubby --args="iommu=pt" --update-kernel=ALL
```

- ii. 重启主机。

流程

1. 可选：确认您的网络设备可使用的最大 VF 数。要做到这一点，请使用以下命令，将 *eth1* 替换为您的 SR-IOV 兼容网络设备。

```
# cat /sys/class/net/eth1/device/sriov_totalvfs
7
```

2. 使用以下命令来创建虚拟功能(VF)：

```
# echo VF-number > /sys/class/net/network-interface/device/sriov_numvfs
```

在命令中，替换：

- 使用您要在其上创建 PF 的 VF 数替换 *VF-number*。
- 使用 VF 要创建的网络接口的名称替换 *network-interface*。

以下示例从 *eth1* 网络接口创建 2 个 VF：

```
# echo 2 > /sys/class/net/eth1/device/sriov_numvfs
```

3. 确定已添加了 VF：

```
# lspci | grep Ethernet
82:00.0 Ethernet controller: Intel Corporation 82599ES 10-Gigabit SFI/SFP+ Network
Connection (rev 01)
82:00.1 Ethernet controller: Intel Corporation 82599ES 10-Gigabit SFI/SFP+ Network
Connection (rev 01)
82:10.0 Ethernet controller: Intel Corporation 82599 Ethernet Controller Virtual Function (rev
01)
82:10.2 Ethernet controller: Intel Corporation 82599 Ethernet Controller Virtual Function (rev
01)
```

4. 通过为您用于创建 VF 的网络接口创建一个 udev 规则，使创建的 VF 持久化。例如，对于 *eth1* 接口，创建 **/etc/udev/rules.d/eth1.rules** 文件，并添加以下行：

```
ACTION=="add", SUBSYSTEM=="net", ENV{ID_NET_DRIVER}=="ixgbe",
ATTR{device/sriov_numvfs}="2"
```

这样可确保使用 **ixgbe** 驱动程序的两个 VF 在主机启动时可自动对 **eth1** 接口可用。如果不需要持久性 SR-IOV 设备，请跳过这一步。

**警告**

目前，当试图在 Broadcom NetXtreme II BCM57810 适配器上保留 VF 时，上述设置无法正常工作。另外，基于这些适配器将 VF 附加到 Windows 虚拟机当前还不可靠。

5. 将新添加的 VF 接口设备热插到正在运行的虚拟机中。

```
# virsh attach-interface testguest1 hostdev 0000:82:10.0 --managed --live --config
```

验证

- 如果过程成功，客户机操作系统会检测新的网络接口卡。

10.7.3. SR-IOV 分配支持的设备

并非所有设备都可用于 SR-IOV。以下设备已在 RHEL 8 中进行了测试，并验证了与 SR-IOV 兼容。

网络设备

- Intel 82599ES 10 千兆以太网控制器 - 使用 **ixgbe** 驱动程序
- Intel 以太网控制器 XL710 系列 - 使用 **i40e** 驱动程序
- Mellanox ConnectX-5 以太网适配器卡 - 使用 **mlx5_core** 驱动程序
- Intel 以太网络适配器 XXV710 - 使用 **i40e** 驱动程序
- Intel 82576 千兆以太网控制器 - 使用 **igb** 驱动程序
- Broadcom NetXtreme II BCM57810 - 使用 **bnx2x** 驱动程序

10.8. 将 DASD 设备附加到 IBM Z 中的虚拟机

使用 **vfio-ccw** 功能，您可以将直接访问存储设备(DASD)作为介质设备分配给 IBM Z 主机上的虚拟机(VM)。例如，虚拟机可以访问 z/OS 数据集，或向 z/OS 机器提供分配的 DASD。

先决条件

- 您的主机系统使用 IBM Z 硬件构架并支持 FICON 协议。
- 目标虚拟机使用 Linux 客户机操作系统。
- driverctl* 软件包已安装。

```
# yum install driverctl
```

- 在主机中载入了必要的内核模块。要验证，请使用：

```
# lsmod | grep vfio
```

输出应包含以下模块：

- **vfio_ccw**
- **vfio_mdev**
- **vfio_iommu_type1**
- 您有一个备用 DASD 设备供虚拟机独自使用，并且您知道设备的标识符。此流程使用 **0.0.002c** 作为示例。在执行这些命令时，请使用 DASD 设备的标识符替换 **0.0.002c**。

流程

1. 获取 DASD 设备的子通道标识符。

```
# lscss -d 0.0.002c
Device Subchan. DevType CU Type Use PIM PAM POM CHPIDs
-----
0.0.002c 0.0.29a8 3390/0c 3990/e9 yes f0 f0 ff 02111221 00000000
```

在本例中，子频道标识符被检测为 **0.0.29a8**。在此流程的以下命令中，将 **0.0.29a8** 替换为检测到的设备的子通道标识符。

2. 如果上一步中的 **lscss** 命令只显示标头输出，且没有设备信息，请执行以下步骤：

- a. 从 **cio_ignore** 列表中删除该设备。

```
# cio_ignore -r 0.0.002c
```

- b. 在客户机操作系统中，[编辑虚拟机的内核命令行](#)，将带有！标记的设备标识符添加到以 **cio_ignore=** 开头的行（如果它还没有存在）。

```
cio_ignore=all,!condev,!0.0.002c
```

- c. 在主机上重复第 1 步，以获取子通道标识符。

3. 将子通道绑定到 **vfio_ccw** 直通驱动程序。

```
# driverctl -b css set-override 0.0.29a8 vfio_ccw
```



注意

这会将 **0.0.29a8** 子通道永久绑定到 **vfio_ccw**，这意味着 DASD 在主机上将不可用。如果需要在主机上使用该设备，您必须首先删除到 'vfio_ccw' 的自动绑定，并将子通道重新绑定到默认驱动程序：

```
# driverctl -b css unset-override 0.0.29a8
```

4. 生成 UUID。

```
# uuidgen
30820a6f-b1a5-4503-91ca-0c10ba12345a
```

5. 定义并启动 DASD 介质设备。

```
# cat nodedev.xml
<device>
  <parent>css_0_0_29a8</parent>
  <capability type="mdev">
    <type id="vfio_ccw-io"/>
    <uuid>30820a6f-b1a5-4503-91ca-0c10ba12345a</uuid>
  </capability>
</device>

# virsh nodedev-define nodedev.xml
Node device 'mdev_30820a6f_b1a5_4503_91ca_0c10ba12345a_0_0_29a8' defined from
'nodedev.xml'

# virsh nodedev-start mdev_30820a6f_b1a5_4503_91ca_0c10ba12345a_0_0_29a8
Device mdev_30820a6f_b1a5_4503_91ca_0c10ba12345a_0_0_29a8 started
```

6. 如果虚拟机正在运行，请关闭虚拟机。

7. 将介质设备附加到虚拟机。为此，请使用 **virsh edit** 工具编辑虚拟机的 XML 配置，将以下部分添加到 XML 中，并将 **uuid** 值替换为您在上一步中生成的 UUID。

```
<hostdev mode='subsystem' type='mdev' model='vfio-ccw'>
  <source>
    <address uuid="30820a6f-b1a5-4503-91ca-0c10ba12345a"/>
  </source>
</hostdev>
```

8. 可选：将介质设备配置为在主机引导时自动启动。

```
# virsh nodedev-autostart mdev_30820a6f_b1a5_4503_91ca_0c10ba12345a_0_0_29a8
```

验证

1. 确保正确配置了介质设备。

```
# virsh nodedev-info mdev_30820a6f_b1a5_4503_91ca_0c10ba12345a_0_0_29a8
Name:      mdev_30820a6f_b1a5_4503_91ca_0c10ba12345a_0_0_29a8
Parent:    css_0_0_0121
Active:    yes
Persistent: yes
Autostart: yes
```

2. 获取 **libvirt** 分配给中介 DASD 设备的标识符。为此，可显示虚拟机的 XML 配置，并查找 **vfio-ccw** 设备。

```
# virsh dumpxml vm-name
<domain>
```

```
[...]
<hostdev mode='subsystem' type='mdev' managed='no' model='vfio-ccw'>
  <source>
    <address uuid='10620d2f-ed4d-437b-8aff-beda461541f9' />
  </source>
  <alias name='hostdev0' />
  <address type='ccw' cssid='0xfe' ssid='0x0' devno='0x0009' />
</hostdev>
[...]
</domain>
```

在本例中，分配的设备的标识符是 **0.0.0009**。

3. 启动虚拟机并登录到其客户端操作系统。
4. 在客户端操作系统中，确认 DASD 设备已被列出。例如：

```
# lscss | grep 0.0.0009
0.0.0009 0.0.0007 3390/0c 3990/e9    f0 f0 ff 12212231 00000000
```

5. 在客户端操作系统中，在线设置设备。例如：

```
# chccwdev -e 0.0009
Setting device 0.0.0009 online
Done
```

其它资源

- [关于cio_ignore 的 IBM 文档](#)
- [在运行时配置内核参数](#)

第 11 章 为虚拟机管理存储

虚拟机(VM)与物理机一样，需要用于数据、程序和系统文件的存储。作为虚拟机管理员，您可以将物理或基于网络的存储分配给虚拟机作为虚拟存储。您还可以修改存储是如何呈现给虚拟机的，无论底层硬件是什么。

以下章节提供有关不同类型的虚拟机存储的信息、它们是如何工作的，以及如何使用 CLI 或 Web 控制台管理它们。

11.1. 了解虚拟机存储

如果您对虚拟机(VM)存储不熟悉，或者不确定它是如何工作的，以下章节提供了有关虚拟机存储各个组件的概括介绍、它是如何工作的、管理基础知识以及红帽提供的支持的解决方案。

您可以找到以下相关信息：

- [存储池](#)
- [存储卷](#)
- [使用 libvirt 管理存储](#)
- [虚拟机存储概述](#)
- [支持和不支持的存储池类型](#)

11.1.1. 存储池简介

存储池是由 **libvirt** 管理的文件、目录或存储设备，来为虚拟机(VM)提供存储。您可以将存储池划分为存储卷，用于存储虚拟机镜像或作为额外存储附加到虚拟机。

此外，多个虚拟机可以共享相同的存储池，从而更好地分配存储资源。

- 存储池可以是持久的或临时的：
 - 主机系统重启后，持久性存储池会保留下。您可以使用 **virsh pool-define** 来创建持久性存储池。
 - 临时存储池仅在主机重启前存在。您可以使用 **virsh pool-create** 命令来创建临时性存储池。

存储池存储类型

存储池可以是本地的也可以基于网络的（共享）：

- **本地存储池**
本地存储池直接附加到主机服务器。它们包括本地设备中的本地目录、直接附加磁盘、物理分区以及逻辑卷管理（LVM）卷组。

本地存储池可用于开发、测试，以及不需要迁移或有大量虚拟机的小型部署。
- **联网的（共享）存储池**
联网的存储池包括使用标准协议通过网络共享的存储设备。

11.1.2. 存储卷简介

存储池划分为 **存储卷**。存储卷是物理分区、LVM 逻辑卷、基于文件的磁盘映像，以及由 **libvirt** 处理的其他存储类型的抽象。无论底层硬件是什么，存储卷都以本地存储设备（如磁盘）的形式出现在虚拟机中。

在主机上，存储卷通过其名称以及从中派生的存储池的标识符来引用。在 **virsh** 命令行上，它采用 **--pool storage_pool volume_name** 的形式。

例如：要在 *guest_images* 池中显示名为 *firstimage* 的卷信息。

```
# virsh vol-info --pool guest_images firstimage
Name:      firstimage
Type:      block
Capacity:   20.00 GB
Allocation: 20.00 GB
```

11.1.3. 使用 libvirt 进行存储管理

使用 **libvirt** 远程协议，您可以管理虚拟机存储的所有方面。这些操作也可以在远程主机上执行。因此，使用 **libvirt** 的管理应用程序（如 RHEL web 控制台）可用来执行配置虚拟机存储的所有所需的任务。

您可以使用 **libvirt** API 来查询存储池中卷的列表，或者获取有关该存储池中容量、分配和可用存储的信息。对于支持它的存储池，您还可以使用 **libvirt** API 来创建、克隆、调整大小和删除存储卷。另外，您可以使用 **libvirt** API 来上传数据到存储卷，从存储卷下载数据，或者从存储卷中删除数据。

11.1.4. 存储管理概述

为了说明用于管理存储的可用选项，以下示例介绍了使用 **mount -t nfs nfs.example.com:/path/to/share /path/to/data** 的 NFS 服务器的一个示例。

作为存储管理员：

- 您可以在虚拟化主机上定义 NFS 存储池来描述导出的服务器路径和客户端目标路径。因此，**libvirt** 可以在 **libvirt** 启动时自动挂载存储，或者在 **libvirt** 运行时根据需要自动挂载存储。
- 您可以根据名称简单地将存储池和存储卷添加到虚拟机。您不需要添加卷的目标路径。因此，即使目标客户端路径有变化，也不会影响虚拟机。
- 您可以将存储池配置为自动启动。执行此操作时，**libvirt** 会在**libvirt** 启动时自动将 NFS 共享磁盘挂载到指定的目录上。**libvirt** 将共享挂载到指定的目录上，类似于命令 **mount nfs.example.com:/path/to/share /vmdata**。
- 您可以使用 **libvirt** API 查询存储卷路径。这些存储卷基本上是 NFS 共享磁盘中的文件。然后，您可以将这些路径复制到虚拟机 XML 定义中的部分，该定义描述了虚拟机块设备的源存储。
- 对于 NFS，您可以使用一个使用 **libvirt** API 的应用程序来在存储池中创建和删除存储卷（NFS 共享中的文件），不超过池大小的限制（共享的存储容量）。
请注意，并非所有存储池类型都支持创建和删除卷。
- 当不再需要时，您可以停止存储池。停止存储池(**pool-destroy**) 可撤销启动操作，在这种情况下，即卸载 NFS 共享。销毁操作不会修改共享中的数据，即使该命令的名称看似要删除。如需更多信息，请参阅 **man virsh**。

11.1.5. 支持和不支持的存储池类型

支持的存储池类型

以下是 RHEL 支持的存储池类型列表：

- 基于目录的存储池
- 基于磁盘的存储池
- 基于分区的存储池
- glusterfs 存储池
- 基于 iSCSI 的存储池
- 基于 LVM 的存储池
- 基于 NFS 的存储池
- 使用 vHBA 设备基于 SCSI 的存储池
- 基于多路径的存储池
- 基于 RBD 的存储池

不支持的存储池类型

以下是 RHEL 不支持的 **libvirt** 存储池类型的列表：

- 基于 Sheepdog 的存储池
- 基于 Vstorage 的存储池
- 基于 ZFS 的存储池

11.2. 使用 CLI 管理虚拟机存储池

您可以使用 CLI 管理存储池的以下方面来为虚拟机分配存储：

- [查看存储池信息](#)
- [创建存储池](#)
 - [使用 CLI 创建基于目录的存储池](#)
 - [使用 CLI 创建基于磁盘的存储池](#)
 - [使用 CLI 创建基于文件系统的存储池](#)
 - [使用 CLI 创建基于 GlusterFS 的存储池](#)
 - [使用 CLI 创建基于 iSCSI 的存储池](#)
 - [使用 CLI 创建基于 LVM 的存储池](#)
 - [使用 CLI 创建基于 NFS 的存储池](#)
 - [使用 CLI 创建带有 vHBA 设备的基于 SCSI 的存储池](#)
- [删除存储池](#)

11.2.1. 使用 CLI 查看存储池信息

使用 CLI，您可以查看所有存储池的列表，其中包含有关存储池的有限或完整的详细信息。您还可以过滤列出的存储池。

流程

- 使用 **virsh pool-list** 命令来查看存储池信息。

```
# virsh pool-list --all --details
Name      State  Autostart Persistent Capacity Allocation Available
default   running yes     yes    48.97 GiB 23.93 GiB 25.03 GiB
Downloads running yes     yes   175.62 GiB 62.02 GiB 113.60 GiB
RHEL-Storage-Pool running yes     yes   214.62 GiB 93.02 GiB 168.60 GiB
```

其它资源

- virsh pool-list --help** 命令

11.2.2. 使用 CLI 创建基于目录的存储池

基于目录的存储池是基于现有挂载的文件系统中的目录。例如，当您想要将文件系统上的剩余空间用于其他用途时，这非常有用。您可以使用 **virsh** 工具来创建基于目录的存储池。

先决条件

- 确定您的管理程序支持目录存储池：

```
# virsh pool-capabilities | grep "'dir' supported='yes'"
```

如果命令显示任何输出结果，则代表支持目录池。

流程

1. 创建存储池

使用 **virsh pool-define-as** 命令来定义和创建目录类型的存储池。例如，要创建使用 `/guest_images` 目录的名为 `guest_images_dir` 的存储池：

```
# virsh pool-define-as guest_images_dir dir --target "/guest_images"
Pool guest_images_dir defined
```

如果您已经有要创建的存储池的 XML 配置，也可以根据 XML 定义池。详情请参阅 [基于目录的存储池参数](#)。

2. 创建存储池目标路径

使用 **virsh pool-build** 命令，来为事先格式化的文件系统存储池创建存储池目标路径，初始化存储源设备，并定义数据格式。

```
# virsh pool-build guest_images_dir
Pool guest_images_dir built

# ls -la /guest_images
```

```
total 8
drwx----- 2 root root 4096 May 31 19:38 .
dr-xr-xr-x. 25 root root 4096 May 31 19:38 ..
```

3. 验证是否已创建池

使用 **virsh pool-list** 命令来验证池是否已被创建。

```
# virsh pool-list --all
```

Name	State	Autostart
default	active	yes
guest_images_dir	inactive	no

4. 启动存储池

使用 **virsh pool-start** 命令来挂载存储池。

```
# virsh pool-start guest_images_dir
Pool guest_images_dir started
```



注意

virsh pool-start 命令仅适用于持久性存储池。临时存储池创建时会自动启动。

5. [可选]自动启动过程

默认情况下，使用 **virsh** 命令定义的存储池不会被设置为在每次虚拟化服务启动时自动启动。使用 **virsh pool-autostart** 命令将存储池配置为自动启动。

```
# virsh pool-autostart guest_images_dir
Pool guest_images_dir marked as autostarted
```

验证

- 使用 **virsh pool-info** 命令来验证存储池是否处于 *running* 状态。检查报告的大小是否与预期一样，以及是否正确配置了自动启动。

```
# virsh pool-info guest_images_dir
Name:      guest_images_dir
UUID:      c7466869-e82a-a66c-2187-dc9d6f0877d0
State:     running
Persistent: yes
Autostart: yes
Capacity:   458.39 GB
Allocation: 197.91 MB
Available:  458.20 GB
```

11.2.3. 使用 CLI 创建基于磁盘的存储池

在基于磁盘的存储池中，池是基于磁盘分区的。例如，当您希望整个磁盘分区专门用于虚拟机(VM)存储时，这非常有用。您可以使用 **virsh** 工具来创建基于磁盘的存储池。

先决条件

- 确定您的管理程序支持基于磁盘的存储池：

```
# virsh pool-capabilities | grep "'disk' supported='yes'"
```

如果命令显示任何输出结果，则代表支持基于磁盘的池。

- 准备一个设备，以此作为存储池的基础。因此，最好使用分区（例如 `/dev/sdb1`）或 LVM 卷。如果您提供了一个虚拟机，对整个磁盘或块设备（例如：`/dev/sdb`）具有写权限，则虚拟机可以对其进行分区，或者在其上创建自己的 LVM 组。这可能会导致主机上的系统错误。但是，如果您需要将整个块设备用于存储池，红帽建议保护设备上的任何重要分区不受 GRUB 的 `os-prober` 功能的影响。要做到这一点，编辑 `/etc/default/grub` 文件，并应用以下配置之一：

- 禁用 `os-prober`。

```
GRUB_DISABLE_OS_PROBER=true
```

- 防止 `os-prober` 发现特定的分区。例如：

```
GRUB_OS_PROBER_SKIP_LIST="5ef6313a-257c-4d43@/dev/sdb1"
```

- 在创建存储池前，备份所选存储设备上的任何数据。根据所使用的 `libvirt` 版本，给存储池指定专用的磁盘可能会重新格式化并擦除当前存储在磁盘设备上的所有数据。

流程

1. 创建存储池

使用 `virsh pool-define-as` 命令来定义和创建磁盘类型的存储池。以下示例创建一个名为 `guest_images_disk` 的存储池，它使用 `/dev/sdb` 设备，并挂载在 `/dev` 目录。

```
# virsh pool-define-as guest_images_disk disk --source-format=gpt --source-dev=/dev/sdb --target /dev
Pool guest_images_disk defined
```

如果您已经有要创建的存储池的 XML 配置，也可以根据 XML 定义池。详情请参阅 [基于磁盘的存储池参数](#)。

2. 创建存储池目标路径

使用 `virsh pool-build` 命令，为事先格式化的文件系统存储池创建存储池目标路径、初始化存储源设备，并定义数据格式。

```
# virsh pool-build guest_images_disk
Pool guest_images_disk built
```



注意

只有基于磁盘、基于文件系统和逻辑存储池才需要构建目标路径。如果 `libvirt` 检测到源存储设备的数据格式与所选存储池类型不同，则构建会失败，除非指定了 `overwrite` 选项。

3. 验证是否已创建池

使用 `virsh pool-list` 命令来验证池是否已被创建。

```
# virsh pool-list --all
```

Name	State	Autostart
default	active	yes
guest_images_disk	inactive	no

4. 启动存储池

使用 **virsh pool-start** 命令来挂载存储池。

```
# virsh pool-start guest_images_disk
Pool guest_images_disk started
```



注意

virsh pool-start 命令仅适用于持久性存储池。临时存储池创建时会自动启动。

5. [可选]自动启动过程

默认情况下，使用 **virsh** 命令定义的存储池不会被设置为在每次虚拟化服务启动时自动启动。使用 **virsh pool-autostart** 命令将存储池配置为自动启动。

```
# virsh pool-autostart guest_images_disk
Pool guest_images_disk marked as autostarted
```

验证

- 使用 **virsh pool-info** 命令来验证存储池是否处于 *running* 状态。检查报告的大小是否与预期一样，以及是否正确配置了自动启动。

```
# virsh pool-info guest_images_disk
Name:      guest_images_disk
UUID:      c7466869-e82a-a66c-2187-dc9d6f0877d0
State:     running
Persistent: yes
Autostart:  yes
Capacity:   458.39 GB
Allocation: 197.91 MB
Available:  458.20 GB
```

11.2.4. 使用 CLI 创建基于文件系统的存储池

当您要在未挂载的文件系统上创建存储池时，请使用基于文件系统的存储池。此存储池基于给定的文件系统挂载点。您可以使用 **virsh** 工具来创建基于文件系统的存储池。

先决条件

- 确保您的管理程序支持基于文件系统的存储池：

```
# virsh pool-capabilities | grep "'fs' supported='yes'"
```

如果该命令显示任何输出结果，则代表支持基于文件的池。

- 准备一个设备，以此作为存储池的基础。因此，最好使用分区（例如 `/dev/sdb1`）或 LVM 卷。如果您提供了一个虚拟机，对整个磁盘或块设备（例如：`/dev/sdb`）具有写权限，则虚拟机可以对其进行分区，或者在其上创建自己的 LVM 组。这可能会导致主机上的系统错误。
但是，如果您需要将整个块设备用于存储池，红帽建议保护设备上的任何重要分区不受 GRUB 的 `os-prober` 功能的影响。要做到这一点，编辑 `/etc/default/grub` 文件，并应用以下配置之一：

- 禁用 `os-prober`。

```
GRUB_DISABLE_OS_PROBER=true
```

- 防止 `os-prober` 发现特定的分区。例如：

```
GRUB_OS_PROBER_SKIP_LIST="5ef6313a-257c-4d43@/dev/sdb1"
```

流程

1. 创建存储池

使用 `virsh pool-define-as` 命令定义和创建文件系统类型的存储池。例如，要创建一个名为 `guest_images_fs` 的存储池，它使用 `/dev/sdc1` 分区，并挂载在 `/guest_images` 目录上：

```
# virsh pool-define-as guest_images_fs fs --source-dev /dev/sdc1 --target
/guest_images
Pool guest_images_fs defined
```

如果您已经有要创建的存储池的 XML 配置，也可以根据 XML 定义池。详情请参阅 [基于文件系统的存储池参数](#)。

2. 定义存储池目标路径

使用 `virsh pool-build` 命令，为事先格式化的文件系统存储池创建存储池目标路径、初始化存储源设备，并定义数据格式。

```
# virsh pool-build guest_images_fs
Pool guest_images_fs built

# ls -la /guest_images
total 8
drwx----- 2 root root 4096 May 31 19:38 .
dr-xr-xr-x. 25 root root 4096 May 31 19:38 ..
```

3. 验证是否已创建池

使用 `virsh pool-list` 命令来验证池是否已被创建。

```
# virsh pool-list --all

Name      State   Autostart
-----
default    active  yes
guest_images_fs  inactive no
```

4. 启动存储池

使用 `virsh pool-start` 命令来挂载存储池。

```
# virsh pool-start guest_images_fs
Pool guest_images_fs started
```



注意

virsh pool-start 命令仅适用于持久性存储池。临时存储池创建时会自动启动。

5. [可选]自动启动过程

默认情况下，使用 **virsh** 命令定义的存储池不会被设置为在每次虚拟化服务启动时自动启动。使用 **virsh pool-autostart** 命令将存储池配置为自动启动。

```
# virsh pool-autostart guest_images_fs
Pool guest_images_fs marked as autostarted
```

验证

1. 使用 **virsh pool-info** 命令来验证存储池是否处于 *running* 状态。检查报告的大小是否与预期一样，以及是否正确配置了自动启动。

```
# virsh pool-info guest_images_fs
Name:      guest_images_fs
UUID:      c7466869-e82a-a66c-2187-dc9d6f0877d0
State:     running
Persistent: yes
Autostart:  yes
Capacity:   458.39 GB
Allocation: 197.91 MB
Available:  458.20 GB
```

2. 验证在文件系统上的目标路径中是否存在 **lost+found** 目录，其表示该设备已挂载。

```
# mount | grep /guest_images
/dev/sdc1 on /guest_images type ext4 (rw)

# ls -la /guest_images
total 24
drwxr-xr-x. 3 root root 4096 May 31 19:47 .
dr-xr-xr-x. 25 root root 4096 May 31 19:38 ..
drwx----- 2 root root 16384 May 31 14:18 lost+found
```

11.2.5. 使用 CLI 创建基于 GlusterFS 的存储池

GlusterFS 是一个用户空间文件系统，其使用用户空间(FUSE)软件接口中的文件系统。如果要在 Gluster 服务器上拥有一个存储池，您可以使用 **virsh** 工具来创建基于 GlusterFS 的存储池。

先决条件

- 在主机上创建基于 GlusterFS 的存储池前，先准备 Gluster。
 - a. 使用以下命令列出 Gluster 服务器的状态来获取 Gluster 服务器的 IP 地址：

```
# gluster volume status
```

```
Status of volume: gluster-vol1
Gluster process          Port Online Pid
-----
Brick 222.111.222.111:/gluster-vol1    49155  Y  18634

Task Status of Volume gluster-vol1
-----
There are no active volume tasks
```

- b. 如果没有安装，请安装 **glusterfs-fuse** 软件包。
- c. 如果没有启用，请启用 **virt_use_fusefs** 布尔值。检查是否已启用。

```
# setsebool virt_use_fusefs on
# getsebool virt_use_fusefs
virt_use_fusefs --> on
```

- 确保您的管理程序支持基于 GlusterFS 的存储池：

```
# virsh pool-capabilities | grep "'gluster' supported='yes'"
```

如果命令显示任何输出，则支持基于 GlusterFS 的池。

流程

1. 创建存储池

使用 **virsh pool-define-as** 命令来定义和创建基于 GlusterFS 的存储池。例如，要创建一个名为 **guest_images_glusterfs** 的存储池，它使用名为 **gluster-vol1**、IP 为 **111.222.111.222** 的 Gluster 服务器，并挂载在 Gluster 服务器的根目录上：

```
# virsh pool-define-as --name guest_images_glusterfs --type gluster --source-host
111.222.111.222 --source-name gluster-vol1 --source-path /
Pool guest_images_glusterfs defined
```

如果您已经有要创建的存储池的 XML 配置，也可以根据 XML 定义池。详情请参阅 [基于 GlusterFS 的存储池参数](#)。

2. 验证是否已创建池

使用 **virsh pool-list** 命令来验证池是否已被创建。

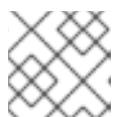
```
# virsh pool-list --all

Name      State   Autostart
-----
default   active   yes
guest_images_glusterfs inactive no
```

3. 启动存储池

使用 **virsh pool-start** 命令来挂载存储池。

```
# virsh pool-start guest_images_glusterfs
Pool guest_images_glusterfs started
```

**注意**

virsh pool-start 命令仅适用于持久性存储池。临时存储池创建时会自动启动。

4. [可选]自动启动过程

默认情况下，使用 **virsh** 命令定义的存储池不会被设置为在每次虚拟化服务启动时自动启动。使用 **virsh pool-autostart** 命令将存储池配置为自动启动。

```
# virsh pool-autostart guest_images_glusterfs
Pool guest_images_glusterfs marked as autostarted
```

验证

- 使用 **virsh pool-info** 命令来验证存储池是否处于 **running** 状态。检查报告的大小是否与预期一样，以及是否正确配置了自动启动。

```
# virsh pool-info guest_images_glusterfs
Name:      guest_images_glusterfs
UUID:      c7466869-e82a-a66c-2187-dc9d6f0877d0
State:     running
Persistent: yes
Autostart:  yes
Capacity:   458.39 GB
Allocation: 197.91 MB
Available:  458.20 GB
```

11.2.6. 使用 CLI 创建基于 iSCSI 的存储池

互联网小型计算机系统接口(iSCSI)是基于 IP 的存储网络标准，用于连接数据存储设施。如果要在 iSCSI 服务器上拥有一个存储池，您可以使用 **virsh** 工具来创建基于 iSCSI 的存储池。

先决条件

- 确定您的管理程序支持基于 iSCSI 的存储池：

```
# virsh pool-capabilities | grep "'iscsi' supported='yes'"
```

如果该命令显示任何输出结果，则代表支持基于 iSCSI 的池。

流程**1. 创建存储池**

使用 **virsh pool-define-as** 命令来定义和创建 iSCSI 类型的存储池。例如，要创建一个名为 **guest_images_iscsi** 的存储池，它使用 **server1.example.com** 上的 **iqn.2010-05.com.example.server1:iscsirhel7guest** IQN，并挂载在 **/dev/disk/by-path** 路径上：

```
# virsh pool-define-as --name guest_images_iscsi --type iscsi --source-host
server1.example.com --source-dev iqn.2010-05.com.example.server1:iscsirhel7guest --
target /dev/disk/by-path
Pool guest_images_iscsi defined
```

如果您已经有要创建的存储池的 XML 配置，也可以根据 XML 定义池。详情请查看 [基于 iSCSI 的存储池参数](#)。

2. 验证是否已创建池

使用 **virsh pool-list** 命令来验证池是否已被创建。

```
# virsh pool-list --all
Name      State   Autostart
-----
default    active  yes
guest_images_iscsi inactive no
```

3. 启动存储池

使用 **virsh pool-start** 命令来挂载存储池。

```
# virsh pool-start guest_images_iscsi
Pool guest_images_iscsi started
```



注意

virsh pool-start 命令仅适用于持久性存储池。临时存储池创建时会自动启动。

4. [可选]自动启动过程

默认情况下，使用 **virsh** 命令定义的存储池不会被设置为在每次虚拟化服务启动时自动启动。使用 **virsh pool-autostart** 命令将存储池配置为自动启动。

```
# virsh pool-autostart guest_images_iscsi
Pool guest_images_iscsi marked as autostarted
```

验证

- 使用 **virsh pool-info** 命令来验证存储池是否处于 *running* 状态。检查报告的大小是否与预期一样，以及是否正确配置了自动启动。

```
# virsh pool-info guest_images_iscsi
Name:      guest_images_iscsi
UUID:      c7466869-e82a-a66c-2187-dc9d6f0877d0
State:     running
Persistent: yes
Autostart:  yes
Capacity:   458.39 GB
Allocation: 197.91 MB
Available:  458.20 GB
```

11.2.7. 使用 CLI 创建基于 LVM 的存储池

如果要拥有一个存储池，其是 LVM 卷组的一部分，您可以使用 **virsh** 工具来创建基于 LVM 的存储池。

建议

在创建基于 LVM 的存储池前请注意以下几点：

- 基于 LVM 的存储池不能为 LVM 提供完整的灵活性。
- libvirt** 支持精简逻辑卷，但不提供精简存储池的功能。

- 基于 LVM 的存储池是卷组。您可以使用 **virsh** 工具创建卷组，但使用这种方式，您在创建的卷组中只能有一个设备。要创建具有多个设备的卷组，请使用 LVM 工具，请参阅 [如何在 Linux 中使用 LVM 创建卷组](#)。
- 有关卷组的详情，请参考 *Red Hat Enterprise Linux Logical Volume Manager Administration Guide*。
- 基于 LVM 的存储池需要一个完整磁盘分区。如果您使用 **virsh** 命令激活了一个新分区或设备，则分区将被格式化，所有数据将被清除。在这些过程中，如果您使用主机的现有卷组，则不会删除任何内容。

先决条件

- 确定您的管理程序支持基于 LVM 的存储池：

```
# virsh pool-capabilities | grep "'logical' supported='yes'"
```

如果命令显示任何输出结果，则支持基于 LVM 的池。

流程

1. 创建存储池

使用 **virsh pool-define-as** 命令来定义和创建 LVM 类型的存储池。例如，以下命令创建名为 **guest_images_lvm** 的存储池，该池使用 **lvm_vg** 卷组，并挂载在 **/dev/lvm_vg** 目录上：

```
# virsh pool-define-as guest_images_lvm logical --source-name lvm_vg --target /dev/lvm_vg
Pool guest_images_lvm defined
```

如果您已经有要创建的存储池的 XML 配置，也可以根据 XML 定义池。详情请查看 [基于 LVM 的存储池参数](#)。

2. 验证是否已创建池

使用 **virsh pool-list** 命令来验证池是否已被创建。

```
# virsh pool-list --all
Name      State   Autostart
-----
default    active  yes
guest_images_lvm  inactive no
```

3. 启动存储池

使用 **virsh pool-start** 命令来挂载存储池。

```
# virsh pool-start guest_images_lvm
Pool guest_images_lvm started
```



注意

virsh pool-start 命令仅适用于持久性存储池。临时存储池创建时会自动启动。

4. [可选]自动启动过程

默认情况下，使用 **virsh** 命令定义的存储池会被设置为在每次虚拟化服务启动时自动启动。使用 **virsh pool-autostart** 命令将存储池配置为自动启动。

```
# virsh pool-autostart guest_images_lvm
Pool guest_images_lvm marked as autostarted
```

验证

- 使用 **virsh pool-info** 命令来验证存储池是否处于 *running* 状态。检查报告的大小是否与预期一样，以及是否正确配置了自动启动。

```
# virsh pool-info guest_images_lvm
Name:      guest_images_lvm
UUID:      c7466869-e82a-a66c-2187-dc9d6f0877d0
State:     running
Persistent: yes
Autostart: yes
Capacity:   458.39 GB
Allocation: 197.91 MB
Available:  458.20 GB
```

11.2.8. 使用 CLI 创建基于 NFS 的存储池

如果要在网络文件系统(NFS)服务器上拥有一个存储池，您可以使用 **virsh** 工具来创建基于 NFS 的存储池。

先决条件

- 确定您的管理程序支持基于 NFS 的存储池：

```
# virsh pool-capabilities | grep "<value>nfs</value>"
```

如果该命令显示任何输出结果，则代表支持基于 NFS 的池。

流程

1. 创建存储池

使用 **virsh pool-define-as** 命令来定义和创建 NFS 类型的存储池。例如，要创建名为 **guest_images_netfs** 的存储池，其使用 IP 为 **111.222.111.222** 的 NFS 服务器，使用目标目录 **/var/lib/libvirt/images/nfspool** 挂载在服务器目录**/home/net_mount** 上：

```
# virsh pool-define-as --name guest_images_netfs --type netfs --source-
host='111.222.111.222' --source-path='/home/net_mount' --source-format='nfs' --
target='/var/lib/libvirt/images/nfspool'
```

如果您已经有要创建的存储池的 XML 配置，也可以根据 XML 定义池。详情请查看 [基于 NFS 的存储池参数](#)。

2. 验证是否已创建池

使用 **virsh pool-list** 命令来验证池是否已被创建。

```
# virsh pool-list --all
```

Name	State	Autostart
default	active	yes
guest_images_netfs	inactive	no

3. 启动存储池

使用 **virsh pool-start** 命令来挂载存储池。

```
# virsh pool-start guest_images_netfs
Pool guest_images_netfs started
```



注意

virsh pool-start 命令仅适用于持久性存储池。临时存储池创建时会自动启动。

4. [可选]自动启动过程

默认情况下，使用 **virsh** 命令定义的存储池不会被设置为在每次虚拟化服务启动时自动启动。使用 **virsh pool-autostart** 命令将存储池配置为自动启动。

```
# virsh pool-autostart guest_images_netfs
Pool guest_images_netfs marked as autostarted
```

验证

- 使用 **virsh pool-info** 命令来验证存储池是否处于 **running** 状态。检查报告的大小是否与预期一样，以及是否正确配置了自动启动。

```
# virsh pool-info guest_images_netfs
Name:      guest_images_netfs
UUID:      c7466869-e82a-a66c-2187-dc9d6f0877d0
State:     running
Persistent: yes
Autostart: yes
Capacity:   458.39 GB
Allocation: 197.91 MB
Available:  458.20 GB
```

11.2.9. 通过 CLI，创建带有 vHBA 设备的基于 SCSI 的存储池

如果您希望在小型计算机系统接口(SCSI)设备上拥有一个存储池，您的主机必须能够使用虚拟主机总线适配器(vHBA)连接到 SCSI 设备。然后，您可以使用 **virsh** 工具来创建基于 SCSI 的存储池。

先决条件

- 确定您的管理程序支持基于 SCSI 的存储池：

```
# virsh pool-capabilities | grep "'scsi' supported='yes'"
```

如果该命令显示任何输出结果，则代表支持基于 SCSI 的池。

- 在使用 vHBA 设备创建基于 SCSI 的存储池前，先创建一个 vHBA。如需更多信息，请参阅 [创建 vHBA](#)。

流程

1. 创建存储池

使用 **virsh pool-define-as** 命令，来定义和创建使用 vHBA 的 SCSI 存储池。例如，以下命令会创建一个名为 **guest_images_vhba** 的存储池，它使用由 **scsi_host3** 父适配器识别的 vHBA、全球范围的端口号 **5001a4ace3ee047d**，以及全局范围的节点号 **5001a4a93526d0a1**。存储池挂载在 **/dev/disk/** 目录上：

```
# virsh pool-define-as guest_images_vhba scsi --adapter-parent scsi_host3 --adapter-wwnn 5001a4a93526d0a1 --adapter-wwpn 5001a4ace3ee047d --target /dev/disk/
Pool guest_images_vhba defined
```

如果您已经有要创建的存储池的 XML 配置，也可以根据 XML 定义池。详情请查看[使用 vHBA 设备的基于 SCSI 的存储池的参数](#)。

2. 验证是否已创建池

使用 **virsh pool-list** 命令来验证池是否已被创建。

```
# virsh pool-list --all
```

Name	State	Autostart
default	active	yes
guest_images_vhba	inactive	no

3. 启动存储池

使用 **virsh pool-start** 命令来挂载存储池。

```
# virsh pool-start guest_images_vhba
Pool guest_images_vhba started
```



注意

virsh pool-start 命令仅适用于持久性存储池。临时存储池创建时会自动启动。

4. [可选]自动启动过程

默认情况下，使用 **virsh** 命令定义的存储池不会被设置为在每次虚拟化服务启动时自动启动。使用 **virsh pool-autostart** 命令将存储池配置为自动启动。

```
# virsh pool-autostart guest_images_vhba
Pool guest_images_vhba marked as autostarted
```

验证

- 使用 **virsh pool-info** 命令来验证存储池是否处于 **running** 状态。检查报告的大小是否与预期一样，以及是否正确配置了自动启动。

```
# virsh pool-info guest_images_vhba
Name:      guest_images_vhba
UUID:      c7466869-e82a-a66c-2187-dc9d6f0877d0
State:     running
Persistent: yes
Autostart:  yes
```

```
Capacity: 458.39 GB
Allocation: 197.91 MB
Available: 458.20 GB
```

11.2.10. 使用 CLI 删除存储池

要从主机系统中删除存储池，您必须停止池，并删除其 XML 定义。

流程

- 使用 **virsh pool-list** 命令列出定义的存储池。

```
# virsh pool-list --all
Name      State   Autostart
-----
default    active   yes
Downloads  active   yes
RHEL-Storage-Pool active yes
```

- 使用 **virsh pool-destroy** 命令停止您要删除的存储池。

```
# virsh pool-destroy Downloads
Pool Downloads destroyed
```

- 可选：对于某些类型的存储池，您可以使用 **virsh pool-delete** 命令删除存储池所在的目录。请注意，为此，目录必须为空。

```
# virsh pool-delete Downloads
Pool Downloads deleted
```

- 使用 **virsh pool-undefine** 命令删除存储池的定义。

```
# virsh pool-undefine Downloads
Pool Downloads has been undefined
```

验证

- 确认删除了存储池。

```
# virsh pool-list --all
Name      State   Autostart
-----
default    active   yes
rhel-Storage-Pool active yes
```

11.3. 使用 WEB 控制台管理虚拟机存储池

使用 RHEL web 控制台，您可以管理存储池来为虚拟机分配存储。

您可以使用 Web 控制台：

- 查看存储池信息。

- 创建存储池：
 - [创建基于目录的存储池。](#)
 - [创建基于 NFS 的存储池。](#)
 - [创建基于 iSCSI 的存储池。](#)
 - [创建基于 LVM 的存储池。](#)
 - [使用 vHBA 设备创建基于 SCSI 的存储池。](#)
- [删除存储池。](#)
- [取消激活存储池。](#)

11.3.1. 使用 Web 控制台查看存储池信息

使用 Web 控制台，您可以查看关于系统上可用存储池的详细信息。存储池可用于为您的虚拟机创建磁盘映像。

先决条件

- Web 控制台 VM 插件 [已安装在您的系统上。](#)

流程

1. 点 **Virtual Machines** 接口顶部的 **Storage Pools**。
此时会出现 **存储池** 窗口，显示配置的存储池列表。

Name	Size	Connection	State	Actions
default	20.22 / 99.98 GiB	System	active	Deactivate ⋮
gnome-boxes	18.67 / 99.98 GiB	Session	inactive	Activate ⋮

该信息包括：

- **Name** - 存储池的名称。
 - **Size** - 当前分配和存储池的总容量。
 - **connection** - 用于访问存储池的连接。
 - **State** - 存储池的状态。
2. 点击您要查看其信息的存储池旁的箭头。
行会展开，以显示概述窗格，其中含有所选存储池的详细信息。



该信息包括：

- **Target path** - 目录支持的存储池类型的源，如 **dir** 或 **netfs**。
- **Persistent** - 表示存储池是否具有持久性配置。
- **Autostart** - 表示在系统启动时存储池是否自动启动。
- **类型** - 存储池的类型。

3. 要查看与存储池相关联的存储卷的列表，请单击 **Storage Volumes**。

这时将出现存储卷窗格，显示已配置的存储卷的列表。

Name	Used by	Size
<input type="checkbox"/> volume1		0 / 1GB
<input type="checkbox"/> volume2		0 / 1GB

该信息包括：

- **Name** - 存储卷的名称。
- **Used by** - 当前使用存储卷的虚拟机。
- **Size** - 卷的大小。

其它资源

- 使用 web 控制台查看虚拟机信息

11.3.2. 使用 Web 控制台创建基于目录的存储池

基于目录的存储池是基于现有挂载的文件系统中的目录。例如，当您想要将文件系统上的剩余空间用于其他用途时，这非常有用。

先决条件

- Web 控制台 VM 插件 [已安装在您的系统上](#)。

流程

1. 在 RHEL web 控制台中，点 **Virtual Machines** 选项卡中的 **Storage pool**。此时会出现 **存储池** 窗口，显示配置的存储池列表（若有）。

Name	Size	Connection	State	
default	20.22 / 99.98 GiB	System	active	Deactivate ⋮
gnome-boxes	18.67 / 99.98 GiB	Session	inactive	Activate ⋮

2. 单击 **Create storage pool**。

此时会出现 **Create 存储池** 对话框。

3. 输入存储池的名称。
4. 在 **Type** 下拉菜单中选择 **Filesystem directory**。

Name	Storage pool name
Type	Filesystem directory
Target path	Path on host's filesystem
Startup	<input checked="" type="checkbox"/> Start pool when host boots
Create Cancel	



注意

如果您没有在下拉菜单中选择 **Filesystem** 目录 选项，则您的管理程序不支持基于目录的存储池。

5. 输入以下信息：
 - **Target path** - 目录支持的存储池类型的源，如 **dir** 或 **netfs**。
 - 「启动」 - 主机引导时是否启动存储池。 6. 点击 **Create**。
- 创建存储池时，**Create Storage Pool**对话框将关闭，新的存储池会出现在存储池列表中。

其它资源

- [了解存储池](#)
- [使用 Web 控制台查看存储池信息](#)

11.3.3. 使用 Web 控制台创建基于 NFS 的存储池

基于 NFS 的存储池是基于服务器上托管的文件系统。

先决条件

- Web 控制台 VM 插件 已安装在您的系统上。

流程

1. 在 RHEL web 控制台中，点 **Virtual Machines** 选项卡中的 **Storage pool**。此时会出现 **存储池** 窗口，显示配置的存储池列表（若有）。

Name	Size	Connection	State	
default	20.22 / 99.98 GiB	System	active	Deactivate
gnome-boxes	18.67 / 99.98 GiB	Session	inactive	Activate

2. 单击 **Create storage pool**。此时会出现 **Create 存储池** 对话框。
3. 输入存储池的名称。
4. 在 **Type** 下拉菜单中选择 **Network File System**。

Name	Storage pool name
Type	Network file system
Target path	Path on host's filesystem
Host	Host name
Source path	The directory on the server being exported
Startup	<input checked="" type="checkbox"/> Start pool when host boots
Create Cancel	



注意

如果您在下拉菜单中选择 **网络文件系统** 选项，则您的管理程序不支持基于 NFS 的存储池。

5. 输入其他信息：

- **目标路径** - 指定目标的路径。这将是用于存储池的路径。
- **主机** - 挂载点所在的网络服务器的主机名。这可以是主机名或 IP 地址。
- **源路径** - 网络服务器中使用的目录。
- **启动」** - 主机引导时是否启动存储池。

6. 点击 **Create**。

已创建存储池。这会关闭 **Create storage pool** 对话框，新的存储池会出现在存储池列表中。

其它资源

- [了解存储池](#)
- [使用 Web 控制台查看存储池信息](#)

11.3.4. 使用 Web 控制台创建基于 iSCSI 的存储池

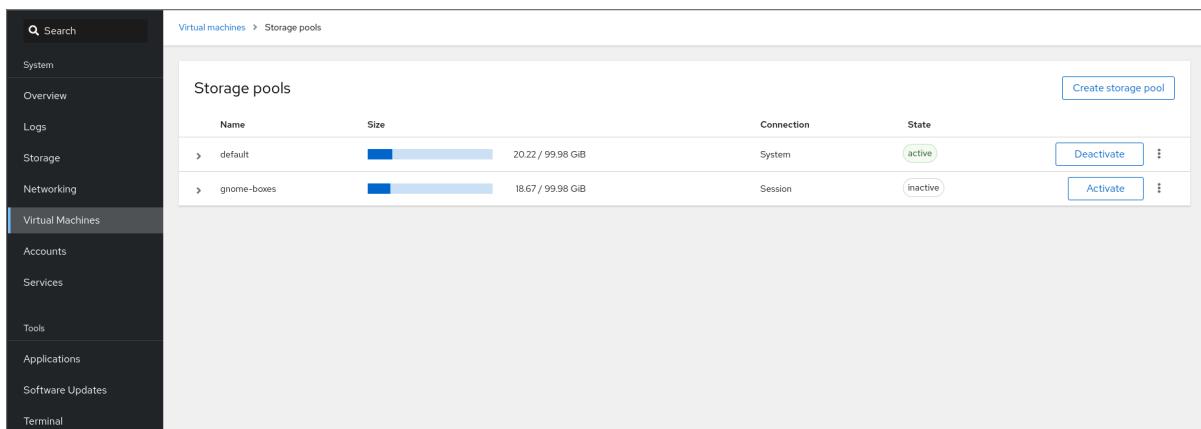
基于 iSCSI 的存储池是基于互联网小型计算机系统接口(iSCSI)，这是一种基于 IP 的存储网络标准，用于连接数据存储设施。

先决条件

- Web 控制台 VM 插件 [已安装在您的系统上](#)。

流程

1. 在 RHEL web 控制台中，点 **Virtual Machines** 选项卡中的 **Storage pool**。
此时会出现 **存储池** 窗口，显示配置的存储池列表（若有）。



2. 单击 **Create storage pool**。
此时会出现 **Create 存储池** 对话框。
3. 输入存储池的名称。
4. 在 **Type** 下拉菜单中选择 **iSCSI 目标**。

Create storage pool

Name	Storage pool name
Type	iSCSI target
Target path	Path on host's filesystem
Host	Host name
Source path	iSCSI target IQN
Startup	<input checked="" type="checkbox"/> Start pool when host boots
<input style="background-color: #0070C0; color: white; font-weight: bold; padding: 5px 10px; margin-right: 10px;" type="button" value="Create"/> <input style="font-weight: bold; padding: 5px 10px;" type="button" value="Cancel"/>	

5. 输入其他信息：

- **目标路径** - 指定目标的路径。这将是用于存储池的路径。
- **主机** - iSCSI 服务器的主机名或 IP 地址。
- **源路径** - iSCSI 目标的唯一 iSCSI 限定名称(IQN)。
- **启动** - 主机引导时是否启动存储池。

6. 点击 **Create**。

已创建存储池。这会关闭 **Create storage pool** 对话框，新的存储池会出现在存储池列表中。

其它资源

- [了解存储池](#)
- [使用 Web 控制台查看存储池信息](#)

11.3.5. 使用 Web 控制台创建基于磁盘的存储池

基于磁盘的存储池使用整个磁盘分区。



警告

- 根据所使用的 libvirt 版本，在存储池中指定一个磁盘可能会重新格式化并清除当前存储在磁盘设备上的所有数据。强烈建议您在创建存储池前备份存储设备中的数据。
- 当整个磁盘或块设备传递给虚拟机时，虚拟机可能会对其分区或者创建自己的 LVM 组。这可能导致主机机器检测到这些分区或者 LVM 组并导致错误。在手动创建分区或 LVM 组并将其传递给虚拟机时，也可以发生这些错误。

要避免这些错误，请改为使用基于文件的存储池。

先决条件

- Web 控制台 VM 插件 [已安装在您的系统上](#)。

流程

- 在 RHEL web 控制台中，点 **Virtual Machines** 选项卡中的 **Storage pool**。此时会出现 **存储池** 窗口，显示配置的存储池列表（若有）。

Name	Size	Connection	State	Action
default	20.22 / 99.98 GiB	System	active	Deactivate
gnome-boxes	18.67 / 99.98 GiB	Session	inactive	Activate

- 单击 **Create storage pool**。此时会出现 **Create 存储池** 对话框。
- 输入存储池的名称。
- 在 **Type** 下拉菜单中选择 **物理磁盘设备**。

Create storage pool

Name	Storage pool name		
Type	Physical disk device		
Target path	Path on host's filesystem		
Source path	Physical disk device on host	Format	dos
Startup	<input checked="" type="checkbox"/> Start pool when host boots		
Create		Cancel	



注意

如果您在下拉菜单中选择 **物理磁盘设备** 选项，则您的管理程序不支持基于磁盘的存储池。

5. 输入其他信息：

- **目标路径** - 指定目标设备的路径。这将是用于存储池的路径。
- **源路径** - 指定存储设备的路径。例如：`/dev/sdb`。
- **格式** - 分区表的类型。
- **启动」** - 主机引导时是否启动存储池。

6. 点击 **Create**。

已创建存储池。这会关闭 **Create storage pool** 对话框，新的存储池会出现在存储池列表中。

其它资源

- [了解存储池](#)
- [使用 Web 控制台查看存储池信息](#)

11.3.6. 使用 Web 控制台创建基于 LVM 的存储池

基于 LVM 的存储池是基于卷组，您可以使用逻辑卷管理器(LVM)进行管理。卷组是多个物理卷的组合，它可创建单个存储结构。



注意

- 基于 LVM 的存储池不能为 LVM 提供完整的灵活性。
- **libvirt** 支持精简逻辑卷，但不提供精简存储池的功能。
- 基于 LVM 的存储池需要一个完整磁盘分区。如果您使用 **virsh** 命令激活了一个新分区或设备，则分区将被格式化，所有数据将被清除。在这些过程中，如果您使用主机的现有卷组，则不会删除任何内容。
- 要创建具有多个设备的卷组，请使用 LVM 工具，请参阅 [如何在 Linux 中使用 LVM 创建卷组](#)。
有关卷组的详情，请参考 *Red Hat Enterprise Linux Logical Volume Manager Administration Guide*。

先决条件

- Web 控制台 VM 插件 [已安装在您的系统上](#)。

流程

1. 在 RHEL web 控制台中，点 **Virtual Machines** 选项卡中的 **Storage pool**。
此时会出现 **存储池** 窗口，显示配置的存储池列表（若有）。

Name	Size	Connection	State	
default	20.22 / 99.98 GiB	System	active	Deactivate
gnome-boxes	18.67 / 99.98 GiB	Session	inactive	Activate

2. 单击 **Create storage pool**。
此时会出现 **Create 存储池** 对话框。
3. 输入存储池的名称。
4. 在 **Type** 下拉菜单中选择 **LVM 卷组**。

Create storage pool

Name	Storage pool name
Type	LVM volume group
Source volume group	Volume group name
Startup	<input checked="" type="checkbox"/> Start pool when host boots
<input type="button" value="Create"/> <input type="button" value="Cancel"/>	



注意

如果您在下拉菜单中选择 **LVM 卷组** 选项，则您的管理程序不支持基于 LVM 的存储池。

5. 输入其他信息：

- 源卷组 - 要使用的 LVM 卷组名称。
- 启动」 - 主机引导时是否启动存储池。

6. 点击 **Create**。

已创建存储池。这会关闭 **Create storage pool** 对话框，新的存储池会出现在存储池列表中。

其它资源

- [了解存储池](#)
- [使用 Web 控制台查看存储池信息](#)

11.3.7. 通过 web 控制台，创建带有 vHBA 设备的基于 SCSI 的存储池

基于 SCSI 的存储池基于小型计算机系统接口(SCSI)设备。在此配置中，您的主机必须能够使用虚拟主机总线适配器(vHBA)连接到 SCSI 设备。

先决条件

- 创建 vHBA。如需更多信息，请参阅 [创建 vHBA](#)。
- Web 控制台 VM 插件 [已安装在您的系统上](#)。

流程

1. 在 RHEL web 控制台中，点 **Virtual Machines** 选项卡中的 **Storage pool**。此时会出现 **存储池** 窗口，显示配置的存储池列表（若有）。

Name	Size	Connection	State	Action
default	20.22 / 99.98 GiB	System	active	Deactivate
gnome-boxes	18.67 / 99.98 GiB	Session	inactive	Activate

2. 单击 **Create storage pool**。

此时会出现 **Create 存储池** 对话框。

3. 输入存储池的名称。
4. 在 **Type** 下拉菜单中选择 **iSCSI 直接目标**。

Name	Storage pool name
Type	iSCSI direct target
Host	Host name
Source path	iSCSI target IQN
Initiator	iSCSI initiator IQN
Startup	<input checked="" type="checkbox"/> Start pool when host boots

Create **Cancel**



注意

如果您在下拉菜单中选择 **iSCSI 直接目标** 选项，则您的管理程序不支持基于 SCSI 的存储池。

5. 输入其他信息：
 - **主机** - 挂载点所在的网络服务器的主机名。这可以是主机名或 IP 地址。
 - **源路径** - iSCSI 目标的唯一 iSCSI 限定名称(IQN)。
 - **Initiator** - iSCSI initiator 的唯一 iSCSI 限定名称(IQN)，即 vHBA。
 - **启动」** - 主机引导时是否启动存储池。 6. 点击 **Create**。
- 已创建存储池。这会关闭 **Create storage pool** 对话框，新的存储池会出现在存储池列表中。

其它资源

- [了解存储池](#)
- [使用 Web 控制台查看存储池信息](#)

11.3.8. 使用 Web 控制台删除存储池

您可以移除存储池，从而释放主机或网络上的资源，以提高系统性能。删除存储池也会释放资源，然后可供其他虚拟机(VM)使用。



重要

除非明确指定，否则删除存储池不会同时删除该池中的存储卷。

要临时取消激活存储池而不是删除它，请参阅使用 [Web 控制台取消激活存储池](#)

先决条件

- Web 控制台 VM 插件 [已安装在您的系统上](#)。
- [从虚拟机中分离磁盘](#)。
- 如果要删除关联的存储卷以及池，[激活这个池](#)。

流程

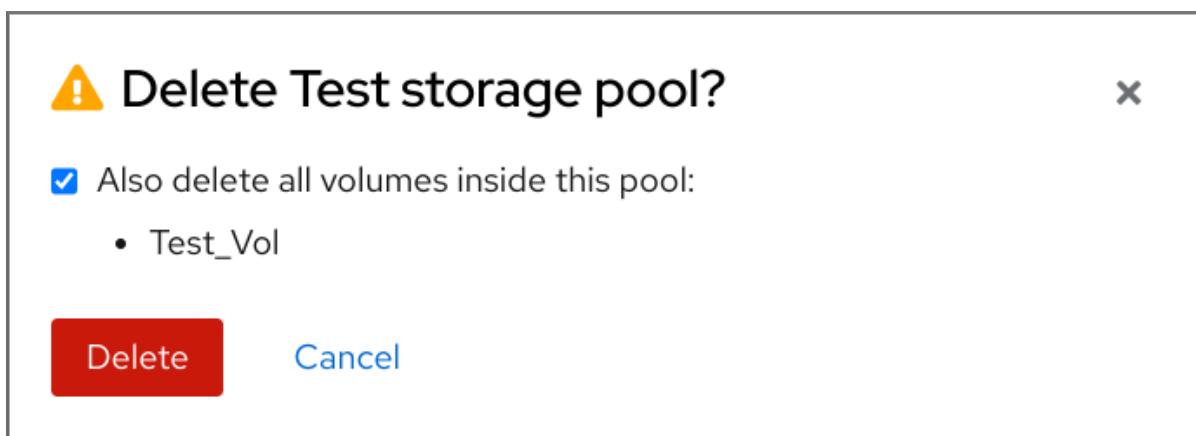
1. 点 **Virtual Machines** 选项卡上的 **Storage Pools**。

此时会出现**存储池**窗口，显示配置的存储池列表。

Name	Size	Connection	State	Action
default	20.22 / 99.98 GiB	System	active	Deactivate
gnome-boxes	18.67 / 99.98 GiB	Session	inactive	Activate

2. 点击您要删除的存储池的 Menu 按钮 **:**，然后点 **Delete**。

此时会出现确认对话框。



3. 可选：要删除池中的存储卷，请在对话框中选择对应的复选框。

4. 点击 **Delete**。

存储池已删除。如果您已在上一步中选择了复选框，则相关的存储卷也会被删除。

其它资源

- [了解存储池](#)
- [使用 Web 控制台查看存储池信息](#)

11.3.9. 使用 Web 控制台取消激活存储池

如果您不想永久删除存储池，您可以临时取消激活它。

当您取消激活存储池时，无法在那个池中创建新卷。但是，在池中拥有卷的任何虚拟机(VM)将继续运行。这在很多方面都很有用，例如，您可以限制池中可以创建的卷数量以提高系统性能。

要使用 RHEL web 控制台取消激活存储池，请参阅以下步骤。

先决条件

- Web 控制台 VM 插件 [已安装在您的系统上](#)。

流程

1. 点击 Virtual Machines 选项卡顶部的 **Storage Pools**。此时会出现存储池窗口，显示配置的存储池列表。

Name	Size	Connection	State	Action
default	20.22 / 99.98 GiB	System	active	<button>Deactivate</button>
gnome-boxes	18.67 / 99.98 GiB	Session	inactive	<button>Activate</button>

2. 点存储池行中的 取消激活。

存储池将停用。

其它资源

- [了解存储池](#)
- [使用 Web 控制台查看存储池信息](#)

11.4. 创建存储池的参数

根据您需要的存储池类型，您可以修改其 XML 配置文件，并定义特定的存储池类型。本节提供了关于创建各种类型的存储池所需的 XML 参数以及示例。

11.4.1. 基于目录的存储池参数

当您想使用 XML 配置文件创建或修改基于目录的存储池时，您必须包含某些必要的参数。有关这些参数的更多信息，请参阅下表。

您可以使用 **virsh pool-define** 命令来根据指定文件中的 XML 配置创建存储池。例如：

```
# virsh pool-define ~/guest_images.xml
Pool defined from guest_images_dir
```

参数

下表提供了基于目录的存储池 XML 文件所需的参数列表。

表 11.1. 基于目录的存储池参数

描述	XML
存储池的类型	<code><pool type='dir'></code>
存储池的名称	<code><name>name</name></code>
指定目标的路径。这将是用于存储池的路径。	<code><target> <path>target_path</path> </target></code>

示例

以下是一个基于 `/guest_images` 目录的存储池的 XML 文件示例：

```
<pool type='dir'>
<name>dirpool</name>
<target>
<path>/guest_images</path>
</target>
</pool>
```

其它资源

- 使用 CLI 创建基于目录的存储池

11.4.2. 基于磁盘的存储池参数

当您想使用 XML 配置文件创建或修改基于磁盘的存储池时，您必须包含某些必要的参数。有关这些参数的更多信息，请参阅下表。

您可以使用 **virsh pool-define** 命令来根据指定文件中的 XML 配置创建存储池。例如：

```
# virsh pool-define ~/guest_images.xml
Pool defined from guest_images_disk
```

参数

下表提供了基于磁盘存储池的 XML 文件所需的参数列表。

表 11.2. 基于磁盘的存储池参数

描述	XML
存储池的类型	<pool type='disk'>
存储池的名称	<name> <i>name</i> </name>
指定存储设备的路径。例如： /dev/sdb 。	<source> <path> <i>source_path</i> </path> </source>
指定目标设备的路径。这将是用于存储池的路径。	<target> <path> <i>target_path</i> </path> </target>

示例

以下是基于磁盘存储池的 XML 文件示例：

```
<pool type='disk'>
  <name>phy_disk</name>
  <source>
    <device path='/dev/sdb' />
    <format type='gpt' />
  </source>
  <target>
    <path>/dev</path>
  </target>
</pool>
```

其它资源

- 使用 CLI 创建基于磁盘的存储池

11.4.3. 基于文件系统的存储池参数

当您想要使用 XML 配置文件创建或修改基于文件系统的存储池时，您必须包含某些必要的参数。有关这些参数的更多信息，请参阅下表。

您可以使用 **virsh pool-define** 命令来根据指定文件中的 XML 配置创建存储池。例如：

```
# virsh pool-define ~/guest_images.xml
Pool defined from guest_images_fs
```

参数

下表提供了基于文件系统的存储池 XML 文件所需的参数列表。

表 11.3. 基于文件系统的存储池参数

描述	XML
存储池的类型	<code><pool type='fs'></code>
存储池的名称	<code><name><i>name</i></name></code>
指定分区的路径。例如： <code>/dev/sdc1</code>	<code><source> <device path=<i>device_path</i> /></code>
文件系统类型，例如 <code>ext4</code> 。	<code><format type=<i>fs_type</i> /> </source></code>
指定目标的路径。这将是用于存储池的路径。	<code><target> <path><i>path-to-pool</i></path> </target></code>

示例

以下是基于 `/dev/sdc1` 分区的存储池的一个 XML 文件示例：

```
<pool type='fs'>
  <name>guest_images_fs</name>
  <source>
    <device path='/dev/sdc1' />
    <format type='auto' />
  </source>
  <target>
    <path>/guest_images</path>
  </target>
</pool>
```

其它资源

- [使用 CLI 创建基于文件系统的存储池](#)

11.4.4. glusterfs 基于存储池参数

当您想要使用 XML 配置文件创建或修改基于 GlusterFS 的存储池时，您必须包含某些必要的参数。有关这些参数的更多信息，请参阅下表。

您可以使用 **virsh pool-define** 命令来根据指定文件中的 XML 配置创建存储池。例如：

```
# virsh pool-define ~/guest_images.xml
Pool defined from guest_images_glusterfs
```

参数

下表提供了基于 GlusterFS 的存储池 XML 文件所需的参数列表。

表 11.4. glusterfs 基于存储池参数

描述	XML
存储池的类型	<code><pool type='gluster'></code>
存储池的名称	<code><name>name</name></code>
Gluster 服务器的主机名或 IP 地址	<code><source> <name=gluster-name /></code>
用于存储池的 Gluster 服务器上的路径。	<code><dir path=gluster-path /> </source></code>

示例

以下是基于 111.222.111.222 Gluster 文件系统的存储池 XML 文件示例：

```
<pool type='gluster'>
<name>Gluster_pool</name>
<source>
<host name='111.222.111.222' />
<dir path='/' />
<name>gluster-vol1</name>
</source>
</pool>
```

其它资源

- [使用 CLI 创建基于 GlusterFS 的存储池](#)

11.4.5. 基于 iSCSI 的存储池参数

当您想使用 XML 配置文件创建或修改基于 iSCSI 的存储池时，您必须包含某些必要的参数。有关这些参数的更多信息，请参阅下表。

您可以使用 **virsh pool-define** 命令来根据指定文件中的 XML 配置创建存储池。例如：

```
# virsh pool-define ~/guest_images.xml
Pool defined from guest_images_iscsi
```

参数

下表提供了基于 iSCSI 存储池的 XML 文件所需的参数列表。

表 11.5. 基于 iSCSI 的存储池参数

描述	XML
存储池的类型	<code><pool type='iscsi'></code>
存储池的名称	<code><name><i>name</i></name></code>
主机的名称	<code><source> <host name=<i>hostname</i> /></code>
iSCSI IQN	<code><device path= <i>iSCSI_IQN</i> /> </source></code>
指定目标的路径。这将是用于存储池的路径。	<code><target> <path>/dev/disk/by-path</path> </target></code>
[可选] iSCSI initiator 的 IQN。只有 ACL 将 LUN 限制为特定发起方时才需要。	<code><initiator> <iqn name='<i>initiator0</i>' /> </initiator></code>



注意

可以使用 `virsh find-storage-pool-sources-as iscsi` 命令来确定 iSCSI 启动器的 IQN。

示例

以下是基于指定 iSCSI 设备的存储池的 XML 文件示例：

```
<pool type='iscsi'>
  <name>iSCSI_pool</name>
  <source>
    <host name='server1.example.com' />
    <device path='iqn.2010-05.com.example.server1:iscsirhel7guest' />
  </source>
  <target>
    <path>/dev/disk/by-path</path>
  </target>
</pool>
```

其它资源

- [使用 CLI 创建基于 iSCSI 的存储池](#)

11.4.6. 基于 LVM 的存储池参数

当您想使用 XML 配置文件创建或修改基于 LVM 的存储池时，您必须包含某些必要的参数。有关这些参数的更多信息，请参阅下表。

您可以使用 **virsh pool-define** 命令来根据指定文件中的 XML 配置创建存储池。例如：

```
# virsh pool-define ~/guest_images.xml
Pool defined from guest_images_logical
```

参数

下表提供了基于 LVM 的存储池 XML 文件所需的参数列表。

表 11.6. 基于 LVM 的存储池参数

描述	XML
存储池的类型	<code><pool type='logical'></code>
存储池的名称	<code><name>name</name></code>
存储池设备的路径	<code><source> <device path='device_path' /></code>
卷组名称	<code><name>VG-name</name></code>
虚拟组格式	<code><format type='lvm2' /> </source></code>
目标路径	<code><target> <path=target_path /> </target></code>



注意

如果逻辑卷由多个磁盘分区组成，则可能会列出多个源设备。例如：

```
<source>
  <device path='/dev/sda1' />
  <device path='/dev/sdb3' />
  <device path='/dev/sdc2' />
  ...
</source>
```

示例

以下是基于指定 LVM 的存储池的 XML 文件示例：

```
<pool type='logical'>
  <name>guest_images_lvm</name>
  <source>
    <device path='/dev/sdc' />
    <name>libvirt_lvm</name>
```

```

<format type='lvm2' />
</source>
<target>
  <path>/dev/libvirt_lvm</path>
</target>
</pool>

```

其它资源

- 使用 CLI 创建基于 LVM 的存储池

11.4.7. 基于 NFS 的存储池参数

当您想使用 XML 配置文件创建或修改基于 NFS 的存储池时，您必须包含某些必要的参数。有关这些参数的更多信息，请参阅下表。

您可以使用 **virsh pool-define** 命令来根据指定文件中的 XML 配置创建存储池。例如：

```
# virsh pool-define ~/guest_images.xml
Pool defined from guest_images_netsfs
```

参数

下表提供了基于 NFS 的存储池 XML 文件所需的参数列表。

表 11.7. 基于 NFS 的存储池参数

描述	XML
存储池的类型	<pool type='netsfs'>
存储池的名称	<name> <i>name</i> </name>
挂载点所在的网络服务器的主机名。这可以是主机名或 IP 地址。	<source> <host name= <i>hostname</i> />
存储池的格式	下面是其中之一： <format type='nfs' /> <format type='glusterfs' /> <format type='cifs' />
网络服务器上使用的目录	<dir path= <i>source_path</i> /> </source>
指定目标的路径。这将是用于存储池的路径。	<target> <path> <i>target_path</i> </path> </target>

示例

以下是基于 **file_server** NFS 服务器的 **/home/net_mount** 目录的存储池的一个 XML 文件示例：

```
<pool type='netfs'>
<name>nfspool</name>
<source>
<host name='file_server' />
<format type='nfs' />
<dir path='/home/net_mount' />
</source>
<target>
<path>/var/lib/libvirt/images/nfspool</path>
</target>
</pool>
```

其它资源

- 使用 CLI 创建基于 NFS 的存储池

11.4.8. 使用 vHBA 设备的基于 SCSI 的存储池的参数

要为使用虚拟主机适配器总线(vHBA)设备的基于 SCSI 的存储池创建或修改 XML 配置文件，您必须在 XML 配置文件中包括某些必要的参数。有关所需参数的更多信息，请参阅下表。

您可以使用 **virsh pool-define** 命令来根据指定文件中的 XML 配置创建存储池。例如：

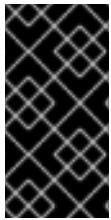
```
# virsh pool-define ~/guest_images.xml
Pool defined from guest_images_vhba
```

参数

下表提供了使用 vHBA 的基于 SCSI 的存储池 XML 文件所需的参数列表。

表 11.8. 使用 vHBA 设备的基于 SCSI 的存储池的参数

描述	XML
存储池的类型	<pool type='scsi'>
存储池的名称	<name> <i>name</i> </name>
vHBA 的标识符。 parent 属性是可选的。	<source> <adapter type='fc_host' [parent= <i>parent_scsi_device</i>] wwn=' <i>WWN</i> ' wwpn=' <i>WWPN</i> ' /> </source>
目标路径。这将是用于存储池的路径。	<target> <path= <i>target_path</i> /> </target>



重要

当 `<path>` 字段是 `/dev/` 时，`libvirt` 会为卷设备路径生成一个唯一的短设备路径。例如：`/dev/sdc`。否则会使用物理主机路径。例如：`/dev/disk/by-path/pci-0000:10:00.0-fc-0x5006016044602198-lun-0`。唯一的短设备路径允许多个存储池在多个虚拟机(VM)中列出相同的卷。如果多个虚拟机使用了物理主机路径，则可能会出现重复的设备类型警告。



注意

可以在 `<adapter>` 字段中使用 `parent` 属性来通过不同的路径标识 NPIV LUN 的物理 HBA 父级。此字段 `scsi_hostN` 与 `vports` 和 `max_vports` 属性相结合，以完成父识别。`parent`、`parent_wwnn`、`parent_wwpn` 或 `parent_fabric_wwn` 属性提供了不同程度的保证，确保在主机重启后使用相同的 HBA。

- 如果没有指定 `parent`, `libvirt` 将使用支持 NPIV 的第一个 `scsi_hostN` 适配器。
- 如果只指定了 `parent`, 则在配置中添加了额外的 SCSI 主机适配器时可能会出现问题。
- 如果指定了 `parent_wwnn` 或 `parent_wwpn`, 则在主机重启后使用相同的 HBA。
- 如果使用 `parent_fabric_wwn`, 则在主机重新启动后, 将选择同一结构上的 HBA , 而不考虑所用的 `scsi_hostN`。

示例

以下是使用 vHBA 的基于 SCSI 的存储池的 XML 文件示例。

- 它是 HBA 中唯一存储池的存储池：

```
<pool type='scsi'>
  <name>vhbapool_host3</name>
  <source>
    <adapter type='fc_host' wwnn='5001a4a93526d0a1' wwpn='5001a4ace3ee047d' />
  </source>
  <target>
    <path>/dev/disk/by-path</path>
  </target>
</pool>
```

- 存储池是使用单个 vHBA 的多个存储池中的一个，并使用 `parent` 属性来识别 SCSI 主机设备：

```
<pool type='scsi'>
  <name>vhbapool_host3</name>
  <source>
    <adapter type='fc_host' parent='scsi_host3' wwnn='5001a4a93526d0a1'
      wwpn='5001a4ace3ee047d' />
  </source>
  <target>
    <path>/dev/disk/by-path</path>
  </target>
</pool>
```

其它资源

- 通过 CLI，创建带有 vHBA 设备的基于 SCSI 的存储池

11.5. 使用 CLI 管理虚拟机存储卷

您可以使用 CLI 管理存储卷的以下方面来为虚拟机分配存储：

- 查看存储卷信息
- 创建存储卷
- 删除存储卷

11.5.1. 使用 CLI 查看存储卷信息

使用命令行，您可以查看主机上所有可用的存储池的列表，以及指定存储池的详细信息。

流程

- 使用 **virsh vol-list** 命令列出指定存储池中的存储卷。

```
# virsh vol-list --pool RHEL-Storage-Pool --details
Name          Path                Type  Capacity Allocation
.bash_history /home/VirtualMachines/.bash_history    file  18.70 KiB  20.00 KiB
.bash_logout   /home/VirtualMachines/.bash_logout     file   18.00 B   4.00 KiB
.bash_profile  /home/VirtualMachines/.bash_profile    file  193.00 B   4.00 KiB
.bashrc        /home/VirtualMachines/.bashrc         file   1.29 KiB   4.00 KiB
.git-prompt.sh /home/VirtualMachines/.git-prompt.sh   file  15.84 KiB  16.00 KiB
.gitconfig     /home/VirtualMachines/.gitconfig       file  167.00 B   4.00 KiB
RHEL_Volume.qcow2 /home/VirtualMachines/RHEL8_Volume.qcow2 file  60.00 GiB
13.93 GiB
```

- 使用 **virsh vol-info** 命令列出指定存储池中的存储卷。

```
# vol-info --pool RHEL-Storage-Pool --vol RHEL_Volume.qcow2
Name:      RHEL_Volume.qcow2
Type:      file
Capacity:  60.00 GiB
Allocation: 13.93 GiB
```

11.5.2. 使用 CLI 创建并分配存储卷

要获取磁盘镜像，并将其作为虚拟磁盘附加到虚拟机(VM)，请创建一个存储卷，并将其 XML 配置分配到虚拟机。

先决条件

- 主机上存在带有未分配空间的存储池。
 - 要验证，列出主机上的存储池：

```
# virsh pool-list --details
```

Name	State	Autostart	Persistent	Capacity	Allocation	Available
------	-------	-----------	------------	----------	------------	-----------

default	running	yes	yes	48.97 GiB	36.34 GiB	12.63 GiB
Downloads	running	yes	yes	175.92 GiB	121.20 GiB	54.72 GiB
VM-disks	running	yes	yes	175.92 GiB	121.20 GiB	54.72 GiB

- 如果您没有现有的存储池，请创建一个。如需更多信息，请参阅[为虚拟机管理存储](#)。

流程

- 使用 **virsh vol-create-as** 命令创建存储卷。例如，要基于 **guest-images-fs** 存储池创建一个 20 GB 的 qcow2 卷：

```
# virsh vol-create-as --pool guest-images-fs --name vm-disk1 --capacity 20 --format qcow2
```

重要：特定的存储池类型不支持 **virsh vol-create-as** 命令，而是需要特定的进程来创建存储卷：

- 基于 **GlusterFS** - 使用 **qemu-img** 命令来创建存储卷。
- 基于 **iSCSI** - 事先在 iSCSI 服务器中准备 iSCSI LUN。
- 基于多路径 - 使用 **multipathd** 命令来准备或管理多路径。
- 基于 **vHBA** - 事先准备光纤通道卡。

- 创建一个 XML 文件，并在其中添加以下几行。此文件将用于将存储卷作为磁盘添加到虚拟机。

```
<disk type='volume' device='disk'>
  <driver name='qemu' type='qcow2'/>
  <source pool='guest-images-fs' volume='vm-disk1' />
  <target dev='hdk' bus='ide' />
</disk>
```

本例指定一个在上一步中创建的 **vm-disk1** 卷的虚拟磁盘，并将卷设为 **ide** 总线上的磁盘 **hdk**。根据您的环境修改对应的参数。

重要：对于特定的存储池类型，您必须使用不同的 XML 格式来描述存储卷磁盘。

- 对于 **基于 GlusterFS 的池**：

```
<disk type='network' device='disk'>
  <driver name='qemu' type='raw' />
  <source protocol='gluster' name='Volume1/Image'>
    <host name='example.org' port='6000' />
  </source>
  <target dev='vda' bus='virtio' />
  <address type='pci' domain='0x0000' bus='0x00' slot='0x03' function='0x0' />
</disk>
```

- 对于 **基于多路径的池**：

```
<disk type='block' device='disk'>
  <driver name='qemu' type='raw' />
  <source dev='/dev/mapper/mpatha' />
```

```
<target dev='sda' bus='scsi' />
</disk>
```

- 对于 基于 RBD 存储的池：

```
<disk type='network' device='disk'>
<driver name='qemu' type='raw' />
<source protocol='rbd' name='pool/image'>
<host name='mon1.example.org' port='6321' />
</source>
<target dev='vdc' bus='virtio' />
</disk>
```

- 使用 XML 文件为虚拟机分配存储卷。例如，要将 `~/vm-disk1.xml` 中定义的磁盘分配给 **testguest1** 虚拟机：

```
# attach-device --config testguest1 ~/vm-disk1.xml
```

验证

- 在虚拟机的客户机操作系统中，确认磁盘映像已作为未格式化和未分配的磁盘提供。

11.5.3. 使用 CLI 删除存储卷

要从主机系统中删除存储卷，您必须停止池并删除它的 XML 定义。

先决条件

- 任何使用您要删除的存储卷的虚拟机都会被关闭。

流程

- 使用 **virsh vol-list** 命令列出指定存储池中的存储卷。

```
# virsh vol-list --pool RHEL-SP
Name          Path
-----
.bash_history  /home/VirtualMachines/.bash_history
.bash_logout   /home/VirtualMachines/.bash_logout
.bash_profile  /home/VirtualMachines/.bash_profile
.bashrc        /home/VirtualMachines/.bashrc
.git-prompt.sh /home/VirtualMachines/.git-prompt.sh
.gitconfig     /home/VirtualMachines/.gitconfig
vm-disk1       /home/VirtualMachines/vm-disk1
```

- 可选：使用 **virsh vol-wipe** 命令来擦除存储卷。例如，要擦除与存储池 **RHEL-SP** 关联的名为 **vm-disk1** 的存储卷：

```
# virsh vol-wipe --pool RHEL-SP vm-disk1
Vol vm-disk1 wiped
```

- 使用 **virsh vol-delete** 命令来删除存储卷。例如，要删除与存储池 **RHEL-SP** 关联的名为 **vm-disk1** 的存储卷：

```
# virsh vol-delete --pool RHEL-SP vm-disk1
Vol vm-disk1 deleted
```

验证

- 再次使用 **virsh vol-list** 命令来验证存储卷是否已被删除。

```
# virsh vol-list --pool RHEL-SP
Name          Path
-----
.bash_history  /home/VirtualMachines/.bash_history
.bash_logout   /home/VirtualMachines/.bash_logout
.bash_profile  /home/VirtualMachines/.bash_profile
.bashrc        /home/VirtualMachines/.bashrc
.git-prompt.sh /home/VirtualMachines/.git-prompt.sh
.gitconfig     /home/VirtualMachines/.gitconfig
```

11.6. 使用 WEB 控制台管理虚拟机存储卷

使用 RHEL，您可以管理为虚拟机(VM)分配存储的存储卷。

您可以使用 RHEL web 控制台进行：

- [创建存储卷。](#)
- [删除存储卷。](#)

11.6.1. 使用 Web 控制台创建存储卷

要创建可正常工作的虚拟机(VM)，您需要将本地存储设备分配给虚拟机的，该设备可存储虚拟机镜像及虚拟机相关的数据。您可以在存储池中创建存储卷，并将其分配为作为存储磁盘的虚拟机。

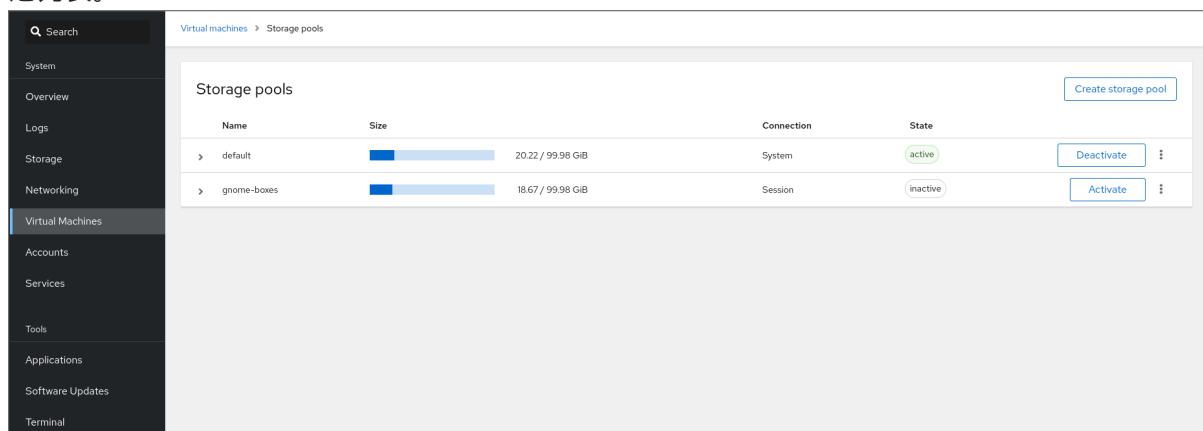
要使用 Web 控制台创建存储卷，请参阅以下步骤。

先决条件

- [Web 控制台 VM 插件已安装在您的系统上。](#)

流程

- 点击 Virtual Machines 选项卡顶部的 **Storage Pools**。此时会出现存储池窗口，显示配置的存储池列表。



2. 在 **Storage Pools** 窗口中，点击您要创建存储卷的存储池。
行会展开，以显示包含所选存储池基本信息的 **Overview** 窗格。

Target path	/var/lib/libvirt/images
Persistent	yes
Autostart	yes
Type	dir

3. 点展开行中的 **Overview** 选项卡旁的 **Storage Volumes**。
Storage Volume 选项卡会出现有关现有存储卷的基本信息。

Name	Used by	Size
Grid_v1.qcow2	Grid_v1	0 / 10 GB
Grid_v1.qcow2		0 / 10 GB
Grid_v2.qcow2	Grid_v2	6.8 / 10 GB

4. 点创建卷。
此时会出现 **Create storage volume** 对话框。

5. 在 Create Storage Volume 对话框中输入以下信息：

- **Name** - 存储卷的名称。
- **size** - MiB 或 GiB 存储卷的大小。
- **格式** - 存储卷的格式。支持的类型为 **qcow2** 和 **raw**。

6. 点击 **Create**。
存储卷创建后，创建存储卷对话框关闭，新的存储卷会出现在存储卷列表中。

其它资源

- [了解存储卷](#)
- [使用 web 控制台向虚拟机添加新磁盘](#)

11.6.2. 使用 Web 控制台删除存储卷

您可以删除存储卷来释放存储池中的空间，或删除与失效的虚拟机(VM)相关联的存储条目。

要使用 RHEL web 控制台删除存储卷，请参阅以下流程。

先决条件

- Web 控制台 VM 插件 已安装在您的系统上。
- 任何使用您要删除的存储卷的虚拟机都会被关闭。

流程

1. 点击 Virtual Machines 选项卡顶部的 **Storage Pools**。此时会出现存储池窗口，显示配置的存储池列表。

Name	Size	Connection	State	Action
default	20.22 / 99.98 GiB	System	active	Deactivate
gnome-boxes	18.67 / 99.98 GiB	Session	inactive	Activate

2. 在 **Storage Pools** 窗口中，单击您要从中删除存储卷的存储池。行会展开，以显示包含所选存储池基本信息的 Overview 窗格。

Target path	/var/lib/libvirt/images
Persistent	yes
Autostart	yes
Type	dir

3. 点展开行中的 Overview 选项卡旁的 **Storage Volumes**。Storage Volume 选项卡会出现有关现有存储卷的基本信息。

Name	Used by	Size
Grid_v1-1.qcow2	Grid_v1	0 / 10 GB
Grid_v1.qcow2	Grid_v1	0 / 10 GB
Grid_v2.qcow2	Grid_v2	6.8 / 10 GB

4. 选择您要删除的存储卷。

Name	Used by	Size
Grid_v1-1.qcow2	Grid_v1	0 / 10 GB
Grid_v1.qcow2		0 / 10 GB
Grid_v2.qcow2	Grid_v2	6.8 / 10 GB
<input checked="" type="checkbox"/> v2		0 / 15 GB

5. 单击 Delete 1 Volume

其它资源

- 了解存储卷

11.7. 使用 WEB 控制台管理虚拟机存储磁盘

使用 RHEL，您可以管理附加到虚拟机的存储磁盘。

您可以使用 RHEL web 控制台进行：

- 查看虚拟机磁盘信息。
- 向虚拟机添加新磁盘。
- 将磁盘附加到虚拟机。
- 从虚拟机中分离磁盘。

11.7.1. 在 web 控制台中查看虚拟机磁盘信息

使用 web 控制台，您可以查看分配给所选虚拟机(VM)的磁盘的详细信息。

先决条件

- Web 控制台 VM 插件 已安装在您的系统上。

流程

1. 点击您要查看其信息的虚拟机。

此时将打开一个新页面，其中包含关于所选虚拟机基本信息的概述部分，以及用于访问虚拟机的图形界面的控制台部分。

2. 滚动到 磁盘。

磁盘部分显示分配给虚拟机的磁盘以及 添加、删除 或 编辑 磁盘的选项。

Device	Used	Capacity	Bus	Access	Source		
disk	8.9 GiB	10 GiB	virtio	Writable	File	/var/lib/libvirt/images/Grid_v2.qcow2	<button>Remove</button> <button>Edit</button>
disk	0 GiB	15 GiB	virtio	Writable	Pool	default	<button>Remove</button> <button>Edit</button>
				Volume	v2		

该信息包括：

- **Device** - 磁盘的设备类型。
- **Used** - 当前分配的磁盘数量。
- **Capacity** - 存储卷的最大大小。
- **Bus** - 模拟的磁盘设备的类型。
- **Access** - 磁盘是否是 **可写** 或 **只读**。对于 **raw** 磁盘，您还可以将访问权限设为 **可写和共享**。
- **Source** - 磁盘设备或者文件。

其它资源

- [使用 web 控制台查看虚拟机信息](#)

11.7.2. 使用 web 控制台向虚拟机添加新磁盘

您可以使用 RHEL 8 web 控制台来创建新存储卷，并将其附加到虚拟机，来将新磁盘添加给虚拟机 (VM)。

先决条件

- Web 控制台 VM 插件 [已安装在您的系统上](#)。

流程

1. 在 **Virtual Machines** 界面中，单击要为其创建并附加新磁盘的虚拟机。
此时将打开一个新页面，其中包含关于所选虚拟机基本信息的概述部分，以及用于访问虚拟机的图形界面的控制台部分。
2. 滚动到 **磁盘**。
磁盘部分显示分配给虚拟机的磁盘以及 **添加**、**删除** 或 **编辑** 磁盘的选项。

Disks						Add disk
Device	Used	Capacity	Bus	Access	Source	
disk	8.9 GiB	10 GiB	virtio	Writeable	File	/var/lib/libvirt/images/Grid_v2.qcow2
					Volume	v2
						Remove Edit
disk	0 GiB	15 GiB	virtio	Writeable	Pool	default
						Remove Edit

3. 点 **Add Disk**。
此时会出现 Add Disk 对话框。

Add disk

Source	<input checked="" type="radio"/> Create new <input type="radio"/> Use existing <input type="radio"/> Custom path		
Pool	default		
Name	New volume name		
Size	1 GiB	Format	qcow2
Persistence	<input type="checkbox"/> Always attach		
Show additional options			
Add		Cancel	

4. 选择 *Create New* 选项。

5. 配置新磁盘。

- **Pool** - 选择创建虚拟磁盘的存储池。
- **Name** - 为要创建的虚拟磁盘输入一个名称。
- **Size** - 输入大小并选择要创建的虚拟磁盘的单元（MiB 或 GiB）。
- **Format** - 选择要创建的虚拟磁盘的格式。支持的类型为 **qcow2** 和 **raw**。
- **Persistence** - 如果选中，虚拟磁盘是永久的。如果没有选择，虚拟磁盘就是临时的。



注意

临时磁盘只能添加到正在运行的虚拟机中。

- **其它选项** - 为虚拟磁盘设置附加配置。

- **Cache** - 选择缓存机制。
- **Bus** - 选择要模拟的磁盘设备的类型。

6. 点添加。

虚拟磁盘已创建并连接到虚拟机。

其它资源

- 在 web 控制台中查看虚拟机磁盘信息
- 使用 web 控制台将现有磁盘附加到虚拟机
- 使用 web 控制台从虚拟机中分离磁盘

11.7.3. 使用 web 控制台将现有磁盘附加到虚拟机

使用 web 控制台，您可以将现有的存储卷作为磁盘附加到虚拟机(VM)。

先决条件

- Web 控制台 VM 插件 已安装在您的系统上。

流程

1. 在 **Virtual Machines** 界面中，单击要为其创建并附加新磁盘的虚拟机。
此时将打开一个新页面，其中包含关于所选虚拟机基本信息的概述部分，以及用于访问虚拟机的图形界面的控制台部分。
2. 滚动到 **磁盘**。
磁盘部分显示分配给虚拟机的磁盘以及 **添加、删除 或 编辑** 磁盘的选项。

Disks						
Device	Used	Capacity	Bus	Access	Source	
disk	8.9 GiB	10 GiB	virtio	Writeable	File	/var/lib/libvirt/images/Grid_v2.qcow2
disk	0 GiB	15 GiB	virtio	Writeable	Pool	default
				Volume	v2	

3. 点 **Add Disk**。
此时会出现 Add Disk 对话框。

Add disk

Source Create new Use existing Custom path

Pool default

Name New volume name

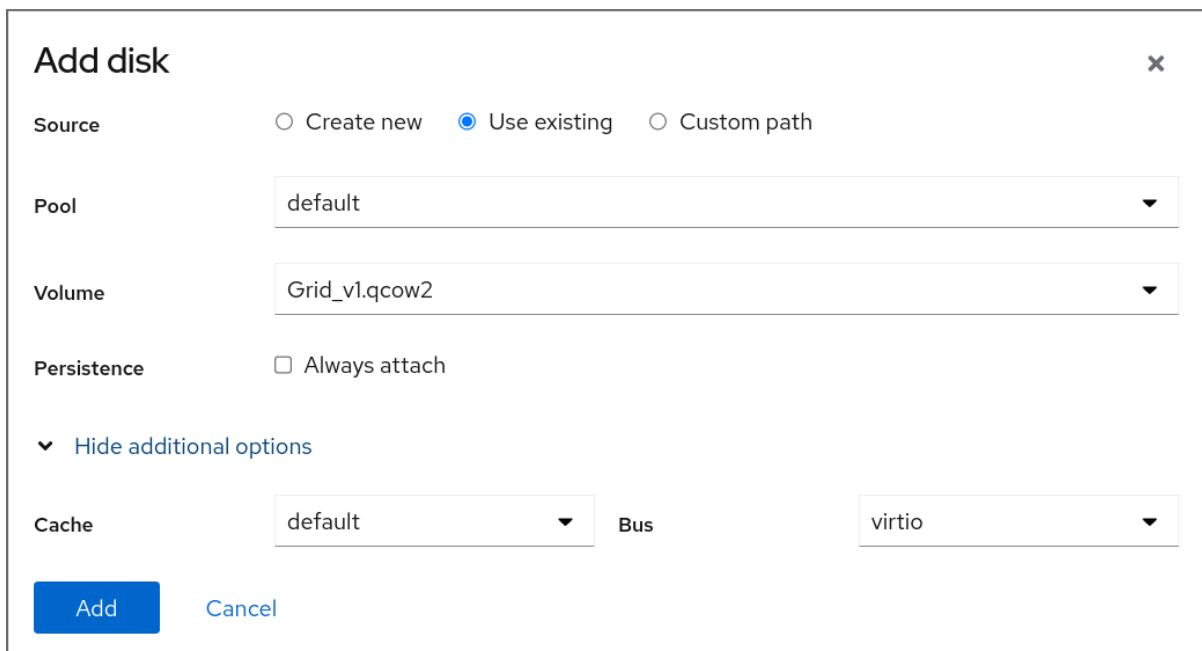
Size 1 GiB Format qcow2

Persistence Always attach

› Show additional options

Add **Cancel**

4. 点**使用现有**按钮。
正确的配置字段会出现在 Add Disk 对话框中。



5. 为虚拟机配置磁盘。

- **Pool** - 选择要从中附加虚拟磁盘的存储池。
- **Volume** - 选择将被附加的存储卷。
- **Persistence** - 虚拟机运行时可用。选中 **Always attach** 复选框，使虚拟磁盘持久可用。清除复选框，使虚拟磁盘为临时磁盘。
- **其它选项** - 为虚拟磁盘设置附加配置。
 - **Cache** - 选择缓存机制。
 - **Bus** - 选择要模拟的磁盘设备的类型。

6. 点 **添加**

所选虚拟磁盘附加到虚拟机。

其它资源

- 在 web 控制台中 [查看虚拟机磁盘信息](#)
- 使用 web 控制台向虚拟机添加新磁盘
- 使用 web 控制台从虚拟机中分离磁盘

11.7.4. 使用 web 控制台将磁盘从虚拟机分离

使用 web 控制台，您可以从虚拟机(VM)分离磁盘。

先决条件

- Web 控制台 VM 插件 [已安装在您的系统上](#)。

流程

1. 在 **Virtual Machines** 界面中，单击您要从中分离磁盘的虚拟机。

此时将打开一个新页面，其中包含关于所选虚拟机基本信息的概述部分，以及用于访问虚拟机的图形界面的控制台部分。

2. 滚动到 磁盘。

磁盘部分显示分配给虚拟机的磁盘以及 **添加、删除 或 编辑** 磁盘的选项。

Disks							Add disk
Device	Used	Capacity	Bus	Access	Source		
disk	8.9 GiB	10 GiB	virtio	Writable	File	/var/lib/libvirt/images/Grid_v2.qcow2	Remove Edit
disk	0 GiB	15 GiB	virtio	Writable	Pool	default	Remove Edit
				Volume	v2		

3. 点您要从虚拟机中分离的磁盘旁的删除按钮。此时会出现 **Remove Disk** 确认对话框。

4. 在确认对话框中，单击 **Remove**。

虚拟磁盘与虚拟机分离。

其它资源

- [在 web 控制台中查看虚拟机磁盘信息](#)
- [使用 web 控制台向虚拟机添加新磁盘](#)
- [使用 web 控制台将现有磁盘附加到虚拟机](#)

11.8. 使用 LIBVIRT SECRET 保护 iSCSI 存储池

可以使用 **virsh** 配置用户名和密码参数，以保护 iSCSI 存储池。您可以在定义池之前或之后配置它，但必须启动池才能使验证设置生效。

以下提供了使用 **libvirt secret** 保护基于 iSCSI 的存储池的说明。



注意

如果在创建 iSCSI 目标时定义了 **user_ID** 和 **密码**，则需要这个流程。

先决条件

- 确保您已创建了一个基于 iSCSI 的存储池。如需更多信息，请参阅[使用 CLI 创建基于 iSCSI 的存储池](#)。

流程

1. 创建具有质询握手身份验证协议(CHAP)用户名的 libvirt secret 文件。例如：

```
<secret ephemeral='no' private='yes'>
  <description>Passphrase for the iSCSI example.com server</description>
  <usage type='iscsi'>
    <target>iscsirhel7secret</target>
  </usage>
</secret>
```

2. 使用 **virsh secret-define** 命令定义 libvirt secret。

```
# virsh secret-define secret.xml
```

3. 使用 **virsh secret-list** 命令验证 UUID。

```
# virsh secret-list
UUID          Usage
-----
2d7891af-20be-4e5e-af83-190e8a922360  iscsi iscsirhel7secret
```

4. 使用 **virsh secret-set-value** 命令，将 secret 分配给上一步输出中的 UUID。这样可保证 CHAP 用户名和密码位于由 libvirt 控制的 secret 列表中。例如：

```
# virsh secret-set-value --interactive 2d7891af-20be-4e5e-af83-190e8a922360
Enter new value for secret:
Secret value set
```

5. 使用 **virsh edit** 命令在存储池的 XML 文件中添加一个身份验证条目，并添加 **<auth>** 元素，指定 **authentication typ**、**username** 和 **secret usage**。

例如：

```
<pool type='iscsi'>
  <name>iscsirhel7pool</name>
  <source>
    <host name='192.168.122.1' />
    <device path='iqn.2010-05.com.example.server1:iscsirhel7guest' />
    <auth type='chap' username='redhat'>
      <secret usage='iscsirhel7secret' />
    </auth>
  </source>
  <target>
    <path>/dev/disk/by-path</path>
  </target>
</pool>
```

注意

<auth> 子元素位于虚拟机的 **<pool>** 和 **<disk>** XML 元素的不同位置。对于 **<pool>**，**<auth>** 在 **<source>** 元素中指定，这描述了在哪里查找池源，因为身份验证是某些池源（iSCSI 和 RBD）的属性。对于子元素的 **<disk>**，对 iSCSI 或 RBD 磁盘的身份验证是磁盘的一个属性。另外，磁盘的 **<auth>** 子元素与存储池的 **<auth>** 子元素不同。

```
<auth username='redhat'>
  <secret type='iscsi' usage='iscsirhel7secret' />
</auth>
```

6. 要激活更改，激活存储池。如果池已启动，停止并重启存储池：

```
# virsh pool-destroy iscsirhel7pool
# virsh pool-start iscsirhel7pool
```

11.9. 创建 VHBA

虚拟主机总线适配器(vHBA)设备将主机系统连接到 SCSI 设备，对于创建基于 SCSI 的存储池是必需的。

您可以通过在 XML 配置文件中定义 vHBA 设备来创建它。

流程

- 使用 **virsh nodedev-list --cap vports** 命令，在您的主机系统上定位 HBA。
以下示例显示了支持 vHBA 的两个 HBA 的主机：

```
# virsh nodedev-list --cap vports
scsi_host3
scsi_host4
```

- 使用 **virsh nodedev-dumpxml HBA_device** 命令查看 HBA 的详细信息。

```
# virsh nodedev-dumpxml scsi_host3
```

命令的输出列出了用于创建 vHBA 的 **<name>**、**<wwnn>** 和 **<wwpn>** 字段。**<max_vports>** 显示支持的 vHBA 的最大数。例如：

```
<device>
<name>scsi_host3</name>
<path>/sys/devices/pci0000:00/0000:00:04.0/0000:10:00.0/host3</path>
<parent>pci_0000_10_00_0</parent>
<capability type='scsi_host'>
<host>3</host>
<unique_id>0</unique_id>
<capability type='fc_host'>
<wwnn>20000000c9848140</wwnn>
<wwpn>10000000c9848140</wwpn>
<fabric_wwn>2002000573de9a81</fabric_wwn>
</capability>
<capability type='vport_ops'>
<max_vports>127</max_vports>
<vports>0</vports>
</capability>
</capability>
</device>
```

在本例中，**<max_vports>** 值显示在 HBA 配置中总共有 127 个虚拟端口可用。**<vports>** 值显示当前正在使用的虚拟端口数。这些值在创建 vHBA 后更新。

- 为 vHBA 主机创建类似如下的 XML 文件。在这些示例中，该文件名为 **vhba_host3.xml**。
本例使用 **scsi_host3** 来描述父 vHBA。

```
<device>
<parent>scsi_host3</parent>
<capability type='scsi_host'>
<capability type='fc_host'>
</capability>
</capability>
</device>
```

这个示例使用 WWNN/WWPN 对描述父 vHBA。

```
<device>
  <name>vhba</name>
  <parent wwnn='20000000c9848140' wwpn='10000000c9848140' />
  <capability type='scsi_host'>
    <capability type='fc_host'>
      </capability>
    </capability>
  </device>
```



注意

WWNN 和 WWPN 值必须与上一步中 HBA 详情中的值匹配。

<parent> 字段指定与这个 vHBA 设备关联的 HBA 设备。下一步中使用 **<device>** 标签中的详细信息，来为主机创建新的 vHBA 设备。有关 **nodedev** XML 格式的更多信息，请参阅 [libvirt 上游页面](#)。



注意

virsh 命令不提供定义 **parent_wwnn**、**parent_wwpn** 或 **parent_fabric_wwn** 属性的方式。

4. 使用 **virsh nodedev-create** 命令，根据上一步中创建的 XML 文件来创建 VHBA。

```
# virsh nodedev-create vhba_host3
Node device scsi_host5 created from vhba_host3.xml
```

验证

- 使用 **virsh nodedev-dumpxml** 命令验证新的 vHBA 详情(scsi_host5):

```
# virsh nodedev-dumpxml scsi_host5
<device>
  <name>scsi_host5</name>
  <path>/sys/devices/pci0000:00/0000:00:04.0/0000:10:00.0/host3/vport-3:0-0/host5</path>
  <parent>scsi_host3</parent>
  <capability type='scsi_host'>
    <host>5</host>
    <unique_id>2</unique_id>
    <capability type='fc_host'>
      <wwnn>5001a4a93526d0a1</wwnn>
      <wwpn>5001a4ace3ee047d</wwpn>
      <fabric_wwn>2002000573de9a81</fabric_wwn>
    </capability>
  </capability>
</device>
```

其它资源

- 通过 CLI，创建带有 vHBA 设备的基于 SCSI 的存储池

第 12 章 在虚拟机中管理 GPU 设备

要提高 RHEL 8 主机上虚拟机(VM)的图形性能，您可以给虚拟机分配一个主机 GPU。

- 您可以将 GPU 从主机分离，并将 GPU 的完全控制直接传给虚拟机。
- 您可以从物理 GPU 创建多个中介设备，并将这些设备作为虚拟 GPU(vGPU)分配给多个客户机。目前仅支持所选的 NVIDIA GPU，且只能为单个客户机分配一个中介设备。

12.1. 为虚拟机分配 GPU

要访问并控制附加到主机系统的 GPU，您必须将主机系统配置为将 GPU 的直接控制传给虚拟机(VM)。



注意

如果您要查找有关分配虚拟 GPU 的信息，[请参阅管理 NVIDIA vGPU 设备](#)。

先决条件

- 您必须在主机机器内核中启用 IOMMU 支持。

- 在 Intel 主机上，您必须启用 VT-d：

1. 使用 **intel_iommu=on** 和 **iommu=pt** 参数重新生成 GRUB 配置：

```
# grubby --args="intel_iommu=on iommu_pt" --update-kernel DEFAULT
```

2. 重启主机。

- 在 AMD 主机上，您必须启用 AMD-Vi。

请注意，在 AMD 主机上，默认启用 IOMMU，您可以添加 **iommu=pt** 将其切换到直通模式：

1. 使用 **iommu=pt** 参数重新生成 GRUB 配置：

```
# grubby --args="iommu=pt" --update-kernel DEFAULT
```



注意

pt 选项只为直通模式中使用的设备启用 IOMMU，并提供更好的主机性能。但是，并非所有硬件都支持这个选项。无论是否启用了这个选项，您仍然可以分配设备。

2. 重启主机。

流程

1. 防止驱动程序绑定到 GPU。

- a. 识别 GPU 连接到的 PCI 总线地址。

```
# lspci -Dnn | grep VGA
```

```
0000:02:00.0 VGA compatible controller [0300]: NVIDIA Corporation GK106GL [Quadro K4000] [10de:11fa] (rev a1)
```

b. 防止主机的图形驱动程序使用 GPU。为此，请使用 GPU 的带有 pci-stub 驱动程序的 PCI ID。

例如，以下命令可防止驱动程序绑定到附加在 10de:11fa 总线的 GPU：

```
# grubby --args="pci-stub.ids=10de:11fa" --update-kernel DEFAULT
```

c. 重启主机。

2. 可选：如果因为支持的限制而无法将某些 GPU 功能（如音频）直通给虚拟机，您可以修改 IOMMU 组中端点的驱动程序绑定，使其只能直通必要的 GPU 功能。

a. 将 GPU 设置转换为 XML，并记下要防止附加到主机驱动程序的端点的 PCI 地址。

为此，通过向地址添加 **pci_** 前缀，将 GPU 的 PCI 总线地址转换为 libvirt 兼容的格式，并将分隔符转换为下划线。

例如，以下命令显示附加在 0000:02:00.0 总线地址的 GPU 的 XML 配置。

```
# virsh nodedev-dumpxml pci_0000_02_00_0
```

```
<device>
<name>pci_0000_02_00_0</name>
<path>/sys/devices/pci0000:00/0000:00:03.0/0000:02:00.0</path>
<parent>pci_0000_00_03_0</parent>
<driver>
<name>pci-stub</name>
</driver>
<capability type='pci'>
<domain>0</domain>
<bus>2</bus>
<slot>0</slot>
<function>0</function>
<product id='0x11fa'>GK106GL [Quadro K4000]</product>
<vendor id='0x10de'>NVIDIA Corporation</vendor>
<iommuGroup number='13'>
<address domain='0x0000' bus='0x02' slot='0x00' function='0x0' />
<address domain='0x0000' bus='0x02' slot='0x00' function='0x1' />
</iommuGroup>
<pci-express>
<link validity='cap' port='0' speed='8' width='16' />
<link validity='sta' speed='2.5' width='16' />
</pci-express>
</capability>
</device>
```

b. 防止端点附加到主机驱动程序。

在本例中，要将 GPU 分配给虚拟机，防止与音频功能对应的端点 **<address domain='0x0000' bus='0x02' slot='0x00' function='0x1' />** 附加到主机音频驱动程序，请将端点附加到 VFIO-PCI。

```
# driverctl set-override 0000:02:00.1 vfio-pci
```

3. 将 GPU 附加到虚拟机

a. 使用 PCI 总线地址为 GPU 创建 XML 配置文件。

例如，您可以使用 GPU 总线地址中的参数创建以下 XML 文件，GPU-Assign.xml。

```
<hostdev mode='subsystem' type='pci' managed='yes'>
<driver name='vfio'/>
<source>
<address domain='0x0000' bus='0x02' slot='0x00' function='0x0'/>
</source>
</hostdev>
```

b. 将文件保存到主机系统上。

c. 将文件与虚拟机的 XML 配置合并。

例如，以下命令将 GPU XML 文件 GPU-Assign.xml 与 **System1** 虚拟机的 XML 配置文件合并：

```
# virsh attach-device System1 --file /home/GPU-Assign.xml --persistent
Device attached successfully.
```



注意

GPU 是作为辅助图形设备附加到虚拟机的。不支持将 GPU 分配为主图形设备，红帽不推荐删除虚拟机 XML 配置中的主模拟图形设备。

验证

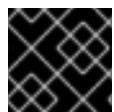
- 该设备会出现在虚拟机 XML 配置的 <devices > 部分下。如需更多信息，请参阅 [示例虚拟机 XML 配置](#)。

已知问题

- 可附加到虚拟机的 GPU 数量受分配的 PCI 设备的最大数量限制，在 RHEL 8 中，当前为 64。但是，将多个 GPU 附加到虚拟机可能会导致客户端上内存映射 I/O(MMIO)出现问题，这可能会导致 GPU 无法对虚拟机使用。
要临时解决这个问题，请设置较大的 64 位 MMIO 空间并配置 vCPU 物理地址位，以使扩展 64 位 MMIO 空间可以寻址。
- 将 NVIDIA GPU 设备附加到使用 RHEL 8 客户机操作系统的虚拟机当前禁用该虚拟机上的 Wayland 会话，并改为加载 Xorg 会话。这是因为 NVIDIA 驱动程序和 Wayland 之间的不兼容。

12.2. 管理 NVIDIA vGPU 设备

vGPU 功能使得可以将物理 NVIDIA GPU 设备划分为多个虚拟设备，称为 **中介设备**。然后可将这些 mediated devices 分配给多个虚拟机（VM）作为虚拟 GPU。因此，这些虚拟机可以共享单个物理 GPU 的性能。



重要

将物理 GPU 分配给使用或不使用中介设备的虚拟机，使主机无法使用 GPU。

12.2.1. 设置 NVIDIA vGPU 设备

要设置 NVIDIA vGPU 功能，您需要为 GPU 设备下载 NVIDIA vGPU 驱动程序，创建介质设备，并将其分配给预期的虚拟机。具体步骤请查看以下说明。

先决条件

- 您的 GPU 支持 vGPU 介质设备。有关支持创建 vGPU 的 NVIDIA GPU 的最新列表，请参阅 [NVIDIA vGPU 软件文档](#)。
- 如果您不知道主机正在使用哪个 GPU，请安装 `lshw` 软件包，并使用 `lshw -C display` 命令。以下示例显示系统使用与 vGPU 兼容的 NVIDIA Tesla P4 GPU。

```
# lshw -C display

*-display
      description: 3D controller
      product: GP104GL [Tesla P4]
      vendor: NVIDIA Corporation
      physical id: 0
      bus info: pci@0000:01:00.0
      version: a1
      width: 64 bits
      clock: 33MHz
      capabilities: pm msi pciexpress cap_list
      configuration: driver=vfio-pci latency=0
      resources: irq:16 memory:f6000000-f6ffffff memory:e0000000-efffffff
                  memory:f0000000-f1ffffffff
```

流程

1. 下载 NVIDIA vGPU 驱动程序并在您的系统中安装它们。具体步骤请查看 [NVIDIA 文档](#)。
2. 如果 NVIDIA 软件安装程序没有创建 `/etc/modprobe.d/nvidia-installer-disable-nouveau.conf` 文件，请在 `/etc/modprobe.d/` 中创建一个任何名称的 `conf` 文件，并在文件中添加以下行：

```
blacklist nouveau
options nouveau modeset=0
```

3. 为当前内核重新生成初始 ramdisk，然后重新启动。

```
# dracut --force
# reboot
```

4. 检查内核是否已加载了 `nvidia_vgpu_vfio` 模块，`nvidia-vgpu-mgr.service` 服务是否正在运行。

```
# lsmod | grep nvidia_vgpu_vfio
nvidia_vgpu_vfio 45011 0
nvidia 14333621 10 nvidia_vgpu_vfio
mdev 20414 2 vfio_mdev,nvidia_vgpu_vfio
vfio 32695 3 vfio_mdev,nvidia_vgpu_vfio,vfio_iommu_type1

# systemctl status nvidia-vgpu-mgr.service
nvidia-vgpu-mgr.service - NVIDIA vGPU Manager Daemon
   Loaded: loaded (/usr/lib/systemd/system/nvidia-vgpu-mgr.service; enabled; vendor preset: disabled)
     Active: active (running) since Fri 2018-03-16 10:17:36 CET; 5h 8min ago
       Main PID: 1553 (nvidia-vgpu-mgr)
          [...]
```

另外，如果基于 NVIDIA Ampere GPU 设备创建 vGPU，请确保为物理 GPU 启用虚拟功能。具体步骤请查看 [NVIDIA 文档](#)。

5. 生成设备 UUID。

```
# uuidgen
30820a6f-b1a5-4503-91ca-0c10ba58692a
```

6. 根据检测到的 GPU 硬件，使用 mediated 设备配置准备 XML 文件。例如，以下命令会在 0000:01:00.0 PCI 总线上运行的 NVIDIA Tesla P4 卡中配置 nvidia-63 vGPU 类型的介质设备，并使用上一步中生成的 UUID。

```
<device>
<parent>pci_0000_01_00_0</parent>
<capability type="mdev">
<type id="nvidia-63"/>
<uuid>30820a6f-b1a5-4503-91ca-0c10ba58692a</uuid>
</capability>
</device>
```

7. 根据您准备的 XML 文件定义 vGPU 介质设备。例如：

```
# virsh nodedev-define vgpu-test.xml
Node device mdev_30820a6f_b1a5_4503_91ca_0c10ba58692a_0000_01_00_0 created
from vgpu-test.xml
```

8. 可选：验证中介设备是否被列为不活动状态。

```
# virsh nodedev-list --cap mdev --inactive
mdev_30820a6f_b1a5_4503_91ca_0c10ba58692a_0000_01_00_0
```

9. 启动您创建的 vGPU 介质设备。

```
# virsh nodedev-start mdev_30820a6f_b1a5_4503_91ca_0c10ba58692a_0000_01_00_0
Device mdev_30820a6f_b1a5_4503_91ca_0c10ba58692a_0000_01_00_0 started
```

10. 可选：确保 中介设备被列为活动状态。

```
# virsh nodedev-list --cap mdev
mdev_30820a6f_b1a5_4503_91ca_0c10ba58692a_0000_01_00_0
```

11. 将 vGPU 设备设置为在主机重启后自动启动

```
# virsh nodedev-autostart
mdev_30820a6f_b1a5_4503_91ca_0c10ba58692a_0000_01_00_0
Device mdev_d196754e_d8ed_4f43_bf22_684ed698b08b_0000_9b_00_0 marked as
autostarted
```

12. 将 mediated 设备附加到要共享 vGPU 资源的虚拟机。要做到这一点，请将以下行以及之前生成的 UUID 添加到虚拟机 XML 配置的 <devices/> 部分。

```
<hostdev mode='subsystem' type='mdev' managed='no' model='vfio-pci' display='on'>
<source>
```

```

<address uuid='30820a6f-b1a5-4503-91ca-0c10ba58692a' />
</source>
</hostdev>
```

请注意，每个 UUID 每次只能分配给一个虚拟机。另外，如果虚拟机没有 QEMU 视频设备，如 **virtio-vga**，在 **<hostdev>** 行中添加 **ramfb='on'** 参数。

- 要使 vGPU 中介设备的全部功能在分配的虚拟机上可用，请在虚拟机上设置 NVIDIA vGPU 虚拟机软件许可。有关详情和说明，请参阅 [NVIDIA Virtual GPU Software License Server User Guide](#)。

验证

- 查询您创建的 vGPU 的功能，并确保它列为 active 和 persistent。

```
# virsh nodedev-info mdev_30820a6f_b1a5_4503_91ca_0c10ba58692a_0000_01_00_0
Name:      virsh nodedev-autostart
mdev_30820a6f_b1a5_4503_91ca_0c10ba58692a_0000_01_00_0
Parent:    pci_0000_01_00_0
Active:    yes
Persistent: yes
Autostart: yes
```

- 启动虚拟机并验证客户端操作系统是否检测到 mediated device 作为 NVIDIA GPU。例如，如果虚拟机使用 Linux：

```
# lspci -d 10de: -k
07:00.0 VGA compatible controller: NVIDIA Corporation GV100GL [Tesla V100 SXM2 32GB]
(rev a1)
    Subsystem: NVIDIA Corporation Device 12ce
    Kernel driver in use: nvidia
    Kernel modules: nouveau, nvidia_drm, nvidia
```

已知问题

- 将 NVIDIA vGPU 介质设备分配给使用 RHEL 8 客户机操作系统的虚拟机当前禁用该虚拟机上的 Wayland 会话，并改为加载 Xorg 会话。这是因为 NVIDIA 驱动程序和 Wayland 之间的不兼容。

其它资源

- [NVIDIA vGPU 软件文档](#)
- [man virsh 命令](#)

12.2.2. 删除 NVIDIA vGPU 设备

要更改分配的 vGPU 介质设备 的配置，您需要从分配的虚拟机中删除现有设备。具体步骤请查看以下操作：

先决条件

- 要从中删除该设备的虚拟机关闭。

流程

1. 获取您要删除的介质设备的 ID。

```
# virsh nodedev-list --cap mdev
mdev_30820a6f_b1a5_4503_91ca_0c10ba58692a_0000_01_00_0
```

2. 停止 vGPU mediated 设备的运行实例。

```
# virsh nodedev-destroy
mdev_30820a6f_b1a5_4503_91ca_0c10ba58692a_0000_01_00_0
Destroyed node device 'mdev_30820a6f_b1a5_4503_91ca_0c10ba58692a_0000_01_00_0'
```

3. 可选：确保中介设备已被停用。

```
# virsh nodedev-info mdev_30820a6f_b1a5_4503_91ca_0c10ba58692a_0000_01_00_0
Name:      virsh nodedev-autostart
mdev_30820a6f_b1a5_4503_91ca_0c10ba58692a_0000_01_00_0
Parent:    pci_0000_01_00_0
Active:    no
Persistent: yes
Autostart: yes
```

4. 从虚拟机 XML 配置中删除该设备。为此，使用 **virsh edit** 工具编辑虚拟机的 XML 配置，并删除 mdev 的配置段。这个片段类似如下：

```
<hostdev mode='subsystem' type='mdev' managed='no' model='vfio-pci'>
  <source>
    <address uuid='30820a6f-b1a5-4503-91ca-0c10ba58692a' />
  </source>
</hostdev>
```

请注意，停止和分离 mediated 设备不会删除它，而是将其保留为 **defined**。因此，您可以 [重启](#) 并将设备 [附加](#) 到其他虚拟机。

5. 可选：要删除已停用的中介设备，请删除其定义。

```
# virsh nodedev-undefine
mdev_30820a6f_b1a5_4503_91ca_0c10ba58692a_0000_01_00_0
Undefined node device 'mdev_30820a6f_b1a5_4503_91ca_0c10ba58692a_0000_01_00_0'
```

验证

- 如果您只停止和分离该设备，请确保介质设备被列为 inactive。

```
# virsh nodedev-list --cap mdev --inactive
mdev_30820a6f_b1a5_4503_91ca_0c10ba58692a_0000_01_00_0
```

- 如果您也删除了该设备，请确保以下命令不会显示它。

```
# virsh nodedev-list --cap mdev
```

其它资源

- [man virsh 命令](#)

12.2.3. 获取有关您系统的 NVIDIA vGPU 信息

要评估可用的 vGPU 功能，您可以获取系统上关于中介设备的其它信息，例如：

- 可创建给定类型的 mediated 设备
- 您的系统中已经配置了哪些介质设备。

流程

- 要查看您主机上可以支持 vGPU 介质设备的可用 GPU 设备，请使用 **virsh nodedev-list --cap mdev_types** 命令。例如，下面显示了有两个 NVIDIA Quadro RTX6000 设备的系统。

```
# virsh nodedev-list --cap mdev_types
pci_0000_5b_00_0
pci_0000_9b_00_0
```

- 要显示特定 GPU 设备支持的 vGPU 类型以及其他元数据，请使用 **virsh nodedev-dumpxml** 命令。

```
# virsh nodedev-dumpxml pci_0000_9b_00_0
<device>
  <name>pci_0000_9b_00_0</name>
  <path>/sys/devices/pci0000:0a/0000:0a:00.0/0000:0b:00.0</path>
  <parent>pci_0000_9a_00_0</parent>
  <driver>
    <name>nvidia</name>
  </driver>
  <capability type='pci'>
    <class>0x030000</class>
    <domain>0</domain>
    <bus>155</bus>
    <slot>0</slot>
    <function>0</function>
    <product id='0x1e30'>TU102GL [Quadro RTX 6000/8000]</product>
    <vendor id='0x10de'>NVIDIA Corporation</vendor>
    <capability type='mdev_types'>
      <type id='nvidia-346'>
        <name>GRID RTX6000-12C</name>
        <deviceAPI>vfio-pci</deviceAPI>
        <availableInstances>2</availableInstances>
      </type>
      <type id='nvidia-439'>
        <name>GRID RTX6000-3A</name>
        <deviceAPI>vfio-pci</deviceAPI>
        <availableInstances>8</availableInstances>
      </type>
      [...]
      <type id='nvidia-440'>
        <name>GRID RTX6000-4A</name>
        <deviceAPI>vfio-pci</deviceAPI>
        <availableInstances>6</availableInstances>
      </type>
      <type id='nvidia-261'>
        <name>GRID RTX6000-8Q</name>
        <deviceAPI>vfio-pci</deviceAPI>
```

```
<availableInstances>3</availableInstances>
</type>
</capability>
<iommuGroup number='216'>
  <address domain='0x0000' bus='0x9b' slot='0x00' function='0x3' />
  <address domain='0x0000' bus='0x9b' slot='0x00' function='0x1' />
  <address domain='0x0000' bus='0x9b' slot='0x00' function='0x2' />
  <address domain='0x0000' bus='0x9b' slot='0x00' function='0x0' />
</iommuGroup>
<numa node='2' />
<pci-express>
  <link validity='cap' port='0' speed='8' width='16' />
  <link validity='sta' speed='2.5' width='8' />
</pci-express>
</capability>
</device>
```

其它资源

- **man virsh** 命令

12.2.4. 用于 NVIDIA vGPU 的远程桌面流服务

在启用了 NVIDIA vGPU 或 NVIDIA GPU 直通的 RHEL 8 hypervisor 上支持以下远程桌面流服务：

- HP ZCentral Remote Boost/Teradici
- NICE DCV
- Mechdyne TGX

有关支持详情请查看适当的供应商支持列表。

12.2.5. 其它资源

- [NVIDIA vGPU 软件文档](#)

第13章 配置虚拟机网络连接

要使虚拟机(VM)通过网络连接到您的主机、连接到您主机上的其他虚拟机，并且连接到外部网络上的某些位置，必须相应地配置虚拟机网络。为了提供虚拟机网络，RHEL 8 hypervisor 和新创建的虚拟机都有一个默认的网络配置，可对其做进一步修改。例如：

- 您可以使主机上的虚拟机通过主机外的位置被发现和连接，就如同虚拟机与主机位于同一网络中一样。
- 您可以部分或完全将虚拟机与入站网络流量隔离，以提高其安全性，并将虚拟机出现任何问题而影响主机的风险降至最低。

以下章节解释了各种类型的虚拟机网络配置，并提供了设置所选虚拟机网络配置的说明。

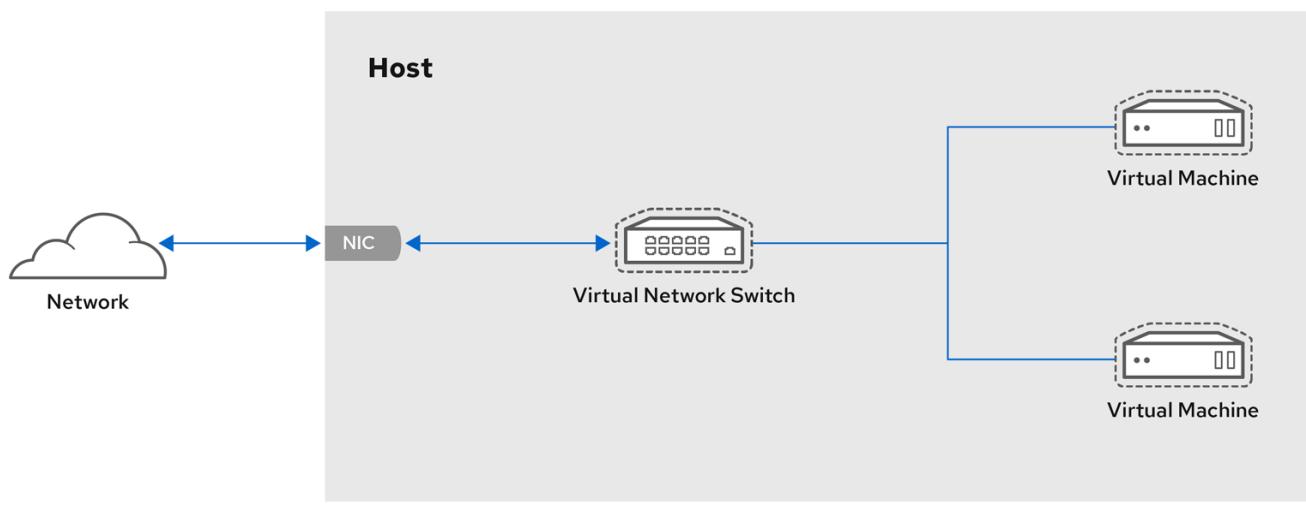
13.1. 了解虚拟网络

虚拟机(VM)连接到网络上的其他设备和位置必须由主机硬件来提供。以下小节解释了虚拟机网络连接的机制，并描述了默认的虚拟机网络设置。

13.1.1. 虚拟网络的工作方式

虚拟网络使用了虚拟网络交换机的概念。虚拟网络交换机是在主机机器中运行的软件构造。VM 通过虚拟网络交换机连接到网络。根据虚拟交换机的配置，虚拟机可以使用由 hypervisor 管理的现有虚拟网络或不同的网络连接方法。

下图展示了将两个虚拟机连接到网络的虚拟网络交换机：



RHEL_52_1219

从客户机操作系统的角度来看，虚拟网络连接与物理网络连接是一样的。主机计算机将虚拟网络交换机视为网络接口。首次安装和启动 **libvirt** 服务时，它会创建 **virbr0**，虚拟机的默认网络接口。

要查看有关此接口的信息，请在主机上使用 **ip** 工具。

```
$ ip addr show virbr0
3: virbr0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc noqueue state
UNKNOWN link/ether 1b:c4:94:cf:fd:17 brd ff:ff:ff:ff:ff:ff
inet 192.168.122.1/24 brd 192.168.122.255 scope global virbr0
```

默认情况下，单个主机上的所有虚拟机都连接到名为 **default** 的同一 **NAT** 类型 虚拟网络，该网络使用 **virbr0** 接口。详情请参阅 [虚拟网络默认配置](#)。

对于虚拟机的基本的出站网络访问，通常不需要额外的网络设置，因为默认网络已与 **libvirt-daemon-config-network** 软件包一起安装了，并在 **libvirtd** 服务启动时自动启动。

如果需要其他虚拟机网络功能，您可以创建额外的虚拟网络和网络接口，并配置虚拟机以使用它们。除了默认的 NAT 外，这些网络和接口也可以配置为使用以下一种模式：

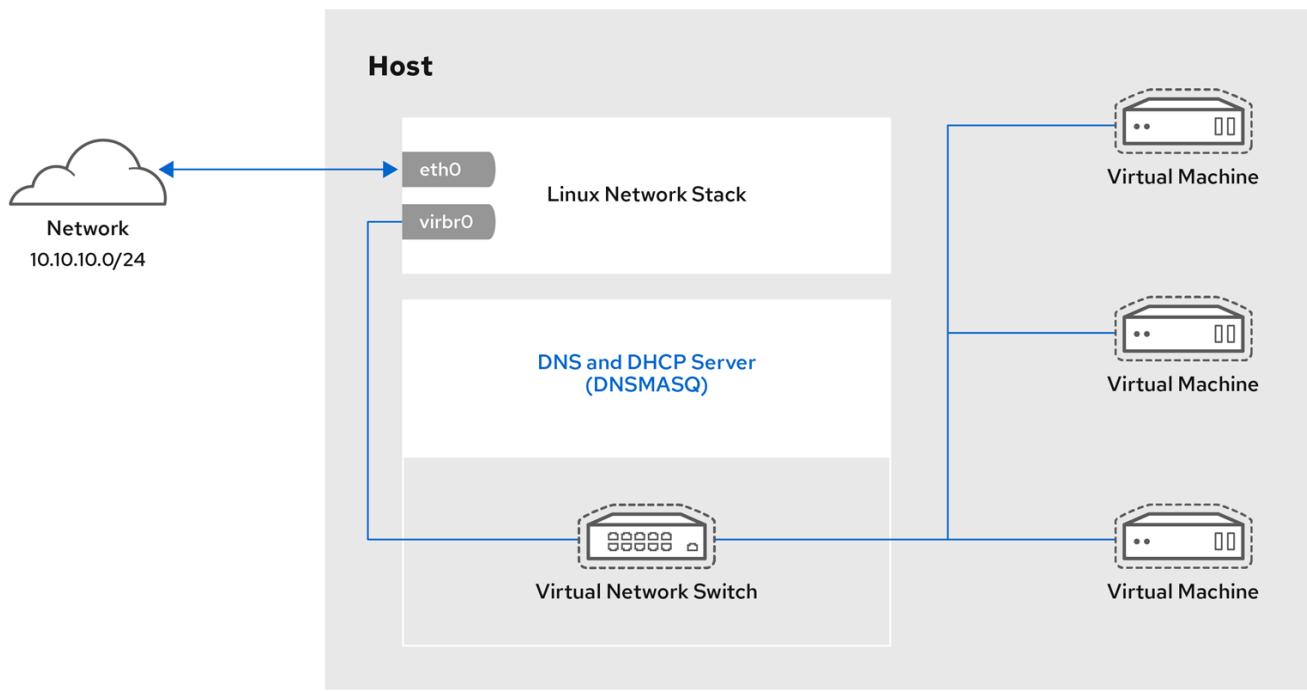
- [路由模式](#)
- [网桥模式](#)
- [隔离模式](#)
- [开放模式](#)

13.1.2. 虚拟网络默认配置

当 **libvirtd** 服务首次安装在虚拟化主机上时，在网络地址转换(NAT)模式下它包含一个初始虚拟网络配置。默认情况下，主机上的所有虚拟机都连接到名为 **default** 的同一 **libvirt** 虚拟网络。这个网络上的虚拟机可以连接到主机上和主机之外的网络位置，但有以下限制：

- 网络上的虚拟机对主机和主机上的其他虚拟机可见，但网络流量会受到客户机操作系统网络堆栈中的防火墙以及附加到客户机接口的 **libvirt** 网络过滤规则的影响。
- 网络上的虚拟机可以连接到主机之外但对它们不可见的位置。出站流量受 NAT 规则以及主机系统的防火墙影响。

下图演示了默认虚拟机网络配置：



13.2. 使用 WEB 控制台管理虚拟机网络接口

使用 RHEL 8 web 控制台，您可以管理连接到 web 控制台的虚拟机的虚拟网络接口。您可以：

- 查看网络接口的相关信息并进行编辑。
- 向虚拟机添加网络接口，以及 断开或删除接口。

13.2.1. 在 web 控制台中查看和编辑虚拟网络接口信息

使用 RHEL 8 web 控制台，您可以在所选虚拟机(VM)上查看和修改虚拟网络接口：

先决条件

- Web 控制台 VM 插件 已安装在您的系统上。

流程

1. 在 **Virtual Machines** 界面中，点击您要查看其信息的虚拟机。
此时将打开一个新页面，其中包含关于所选虚拟机基本信息的概述部分，以及用于访问虚拟机的图形界面的控制台部分。
2. 滚动 到网络接口。
网络接口部分显示为虚拟机配置的虚拟网络接口的信息，以及 添加、删除、编辑 或 拔掉 网络接口的选项。

Network interfaces						Add network interface
Type	Model type	MAC address	IP address	Source	State	
network	virtio	52:54:00	inet 192.168.122.9/24	default	up	Delete Unplug Edit

该信息包括：

- **类型** - 虚拟机的网络接口类型。类型包括虚拟网络、到 LAN 的桥接和直接连接。



注意

RHEL 8 及更新版本不支持通用以太网连接。

- **型号类型** - 虚拟网络接口的型号。
- **MAC 地址** - 虚拟网络接口的 MAC 地址。
- **IP 地址** - 虚拟网络接口的 IP 地址。
- **Source** - 网络接口源。这取决于网络类型。
- **State** - 虚拟网络接口的状态。

3. 要编辑虚拟网络接口设置，请点 **Edit**。此时会打开「虚拟网络接口设置」对话框。

52:54:00:b4:2a:62 virtual network interface settings

Interface type	Virtual network
Source	default
Model	(Linux, perf)
MAC address	52:64:00:b4:2a:63
<input type="button" value="Save"/> <input type="button" value="Cancel"/>	

4. 更改接口类型、源、型号或 MAC 地址。
5. 点 **Save**。已修改网络接口。



注意

对虚拟网络接口设置的更改仅在重启虚拟机后生效。

此外，MAC 地址只能在虚拟机关闭时修改。

其它资源

- [使用 web 控制台查看虚拟机信息](#)

13.2.2. 在 web 控制台中添加和连接虚拟网络接口

使用 RHEL 8 web 控制台，您可以创建一个虚拟网络接口并将虚拟机(VM)连接到它。

先决条件

- Web 控制台 VM 插件 [已安装在您的系统上](#)。

流程

1. 在 **Virtual Machines** 界面中，点击您要查看其信息的虚拟机。
此时将打开一个新页面，其中包含关于所选虚拟机基本信息的概述部分，以及用于访问虚拟机的图形界面的控制台部分。
2. 滚动 **到网络接口**。
网络接口部分显示了为虚拟机配置的虚拟网络接口的信息，以及 **添加**、**删除**、**编辑** 或 **插入** 网络接口的选项。

Network interfaces						Add network interface
Type	Model type	MAC address	IP address	Source	State	
network	virtio	52:54:	Unknown	default	down	Delete Plug Edit

3. 点击您要连接的虚拟网络接口所在行中的 **Plug**。
所选虚拟网络接口连接至虚拟机。

13.2.3. 在 web 控制台中断开连接和删除虚拟网络接口

使用 RHEL 8 web 控制台，您可以断开连接到所选虚拟机(VM)的虚拟网络接口。

先决条件

- Web 控制台 VM 插件 [已安装在您的系统上](#)。

流程

- 在 **Virtual Machines** 界面中，点击您要查看其信息的虚拟机。

此时将打开一个新页面，其中包含关于所选虚拟机基本信息的概述部分，以及用于访问虚拟机的图形界面的控制台部分。

- 滚动 **到网络接口**。

网络接口部分显示为虚拟机配置的虚拟网络接口的信息，以及 **添加、删除、编辑或拔掉** 网络接口的选项。

Network interfaces						Add network interface
Type	Model type	MAC address	IP address	Source	State	
network	virtio	52:54:00	inet 192.168.122.9/24	default	up	Delete Unplug Edit

- 在您要断开连接的虚拟网络接口行中点 **Unplug**。

所选虚拟网络接口断开与虚拟机的连接。

13.3. 推荐的虚拟机网络配置

在很多情况下，默认的 VM 网络配置已经足够了。但是，如果需要调整配置，您可以使用命令行界面(CLI)或 RHEL 8 web 控制台进行此操作。以下章节描述了这些情况下所选择的虚拟机网络设置。

13.3.1. 使用命令行界面配置外部可见虚拟机

默认情况下，新创建的虚拟机连接到使用 **virbr0**（主机上默认的虚拟网桥）的 NAT 类型网络。这可确保虚拟机可以使用主机的网络接口控制器(NIC)来连接到外部网络，但无法从外部系统访问虚拟机。

如果需要虚拟机出现在与 hypervisor 相同的外部网络上，您必须使用 [桥接模式](#)。为此，请将虚拟机连接到连接到 hypervisor 物理网络设备的网桥设备。要使用命令行界面，请遵循以下步骤。

先决条件

- 关闭使用默认 NAT 设置的 [现有虚拟机](#)。
- 管理程序的 IP 配置。这随着主机的网络连接而变化。例如，这个流程使用一种场景，其中主机使用以太网线连接到网络，主机的物理 NIC MAC 地址被分配给 DHCP 服务器上的一个静态 IP。因此，以太网接口被视为 hypervisor IP。

要获取以太网接口的 IP 配置，请使用 **ip addr** 工具：

```
# ip addr
[...]
enp0s25: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc fq_codel state UP
group default qlen 1000
    link/ether 54:ee:75:49:dc:46 brd ff:ff:ff:ff:ff:ff
    inet 10.0.0.148/24 brd 10.0.0.255 scope global dynamic noprefixroute enp0s25
```

流程

- 为主机上的物理接口创建和设置网桥连接。具体步骤请参阅 [配置网络桥接](#)。
请注意，在使用静态 IP 分配的场景中，您必须将物理以太网接口的 IPv4 设置移到网桥接口。
- 修改虚拟机的网络，以使用创建的桥接接口。例如，以下设置 *testguest* 使用 *bridge0*。

```
# virt-xml testguest --edit --network bridge=bridge0
Domain 'testguest' defined successfully.
```

- 启动虚拟机。

```
# virsh start testguest
```

- 在客户机操作系统中，调整系统网络接口的 IP 和 DHCP 设置，就好像虚拟机是与 hypervisor 位于同一网络中的另一台物理系统一样。
具体步骤将因虚拟机使用的客户端操作系统而异。例如，如果客户机操作系统是 RHEL 8，请参阅[配置以太网连接](#)。

验证

- 确保新创建的网桥正在运行，并且包含主机的物理接口和虚拟机的接口。

```
# ip link show master bridge0
2: enp0s25: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc fq_codel master
bridge0 state UP mode DEFAULT group default qlen 1000
    link/ether 54:ee:75:49:dc:46 brd ff:ff:ff:ff:ff:ff
10: vnet0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc fq_codel master
bridge0 state UNKNOWN mode DEFAULT group default qlen 1000
    link/ether fe:54:00:89:15:40 brd ff:ff:ff:ff:ff:ff
```

- 确保虚拟机出现在与 hypervisor 相同的外部网络上：

- 在客户机操作系统中，获取系统的网络 ID。例如，如果它是 Linux 客户机：

```
# ip addr
[...]
enp0s0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc fq_codel state
UP group default qlen 1000
    link/ether 52:54:00:09:15:46 brd ff:ff:ff:ff:ff:ff
    inet 10.0.0.150/24 brd 10.0.0.255 scope global dynamic noprefixroute enp0s0
```

- 从外部系统连接到本地网络，使用获取的 ID 连接到虚拟机。

```
# ssh root@10.0.0.150
root@10.0.0.150's password:
Last login: Mon Sep 24 12:05:36 2019
root~#*
```

如果连接正常工作，则网络已配置成功。

故障排除

- 在某些情况下，比如当虚拟机托管在客户端上时使用客户端到站点的 VPN，使用桥接模式使您的虚拟机可用于外部位置是不可能的。
要临时解决这个问题，您可以为虚拟机[使用 nftables 设置目标 NAT](#)。

其它资源

- [使用 web 控制台配置外部可见虚拟机](#)
- [网桥模式中的虚拟网络](#)

13.3.2. 使用 web 控制台配置外部可见的虚拟机

默认情况下，新创建的虚拟机连接到使用 **virbr0**（主机上默认的虚拟网桥）的 NAT 类型网络。这可确保虚拟机可以使用主机的网络接口控制器(NIC)来连接到外部网络，但无法从外部系统访问虚拟机。

如果需要虚拟机出现在与 hypervisor 相同的外部网络上，您必须使用 [桥接模式](#)。为此，请将虚拟机连接到连接到 hypervisor 物理网络设备的网桥设备。要使用 RHEL 8 web 控制台，请遵循以下步骤。

先决条件

- Web 控制台 VM 插件[已安装在您的系统上](#)。
- 关闭使用默认 NAT 设置的[现有虚拟机](#)。
- 管理程序的 IP 配置。这随着主机的网络连接而变化。例如，这个流程使用一种场景，其中主机使用以太网线连接到网络，主机的物理 NIC MAC 地址被分配给 DHCP 服务器上的一个静态 IP。因此，以太网接口被视为 hypervisor IP。
要获取以太网接口的 IP 配置，请转到 web 控制台中的 **Networking** 选项卡，并参阅 **Interfaces** 部分。

Interfaces		
Name	IP Address	Sending
enp0s25	10.0.0.148/24, 2a00:1028:83a4:1eda:91c7:667:8845:fa2e/64	2.29 Mbps

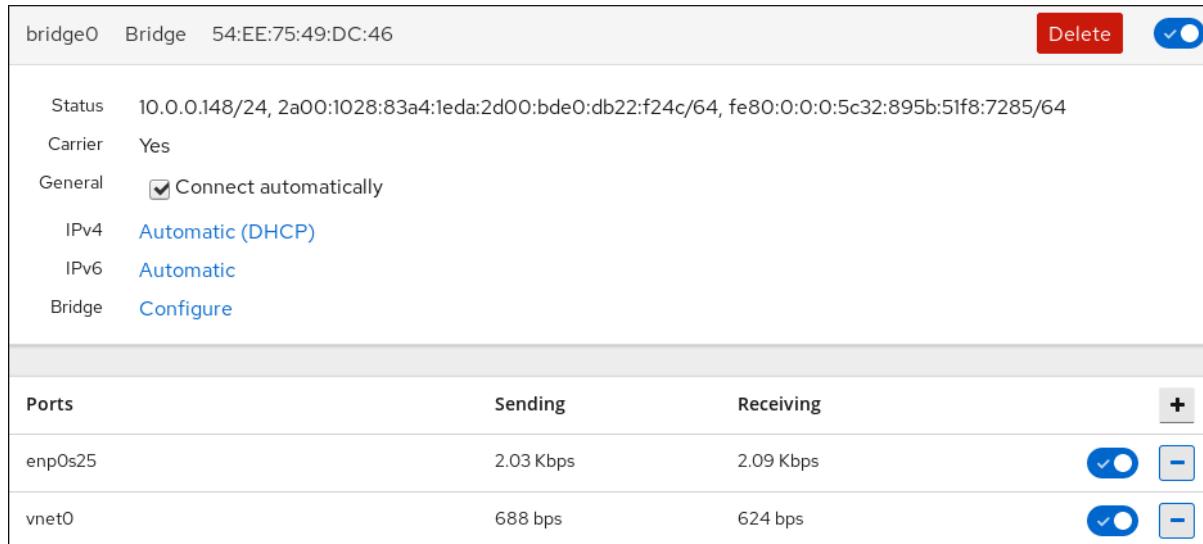
流程

- 为主机上的物理接口创建和设置网桥连接。具体步骤请参阅[在 web 控制台中配置网络桥接](#)。
请注意，在使用静态 IP 分配的场景中，您必须将物理以太网接口的 IPv4 设置移到网桥接口。
- 修改虚拟机的网络，以使用桥接接口。在虚拟机的 [网络接口](#) 选项卡中：
 - 点击 **Add Network Interface**
 - 在 **Add Virtual Network Interface** 对话框中：
 - 将 **Interface Type** 设成 **Bridge to LAN**
 - Source 到新创建的网桥，比如 **bridge0**
 - 点 **添加**
 - 可选：对连接到虚拟机的所有其他接口，点击 **Unplug**。
 - 点击 **Run** 来启动虚拟机。

4. 在客户机操作系统中，调整系统网络接口的 IP 和 DHCP 设置，就好像虚拟机是与 hypervisor 位于同一网络中的另一台物理系统一样。
具体步骤将因虚拟机使用的客户端操作系统而异。例如，如果客户机操作系统是 RHEL 8，请参阅[配置以太网连接](#)。

验证

- 在主机 Web 控制台的 **Networking** 选项卡中，单击新创建的网桥所在的行，以确保它正在运行，并且包含主机的物理接口和虚拟机的接口。



The screenshot shows the 'Networking' configuration screen. A table lists network interfaces. The first row is for a bridge named 'bridge0'. The 'Ports' section below shows two ports: 'enp0s25' and 'vnet0'. Both ports have their 'Sending' and 'Receiving' speeds displayed.

Ports	Sending	Receiving	
enp0s25	2.03 Kbps	2.09 Kbps	<input checked="" type="checkbox"/> <input type="button" value="-"/>
vnet0	688 bps	624 bps	<input checked="" type="checkbox"/> <input type="button" value="-"/>

- 确保虚拟机出现在与 hypervisor 相同的外部网络上。

- 在客户机操作系统中，获取系统的网络 ID。例如，如果它是 Linux 客户机：

```
# ip addr
[...]
enp0s0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc fq_codel state UP
group default qlen 1000
    link/ether 52:54:00:09:15:46 brd ff:ff:ff:ff:ff:ff
    inet 10.0.0.150/24 brd 10.0.0.255 scope global dynamic noprefixroute enp0s0
```

- 从外部系统连接到本地网络，使用获取的 ID 连接到虚拟机。

```
# ssh root@10.0.0.150
root@110.34.5.18's password:
Last login: Mon Sep 24 12:05:36 2019
root~#*
```

如果连接正常工作，则网络已配置成功。

故障排除

- 在某些情况下，比如当虚拟机托管在客户端上时使用客户端到站点的 VPN，使用桥接模式使您的虚拟机可用于外部位置是不可能的。

要临时解决这个问题，您可以[使用 nftables 为虚拟机设置目标 NAT](#)。

其它资源

- [使用命令行界面配置外部可见虚拟机](#)

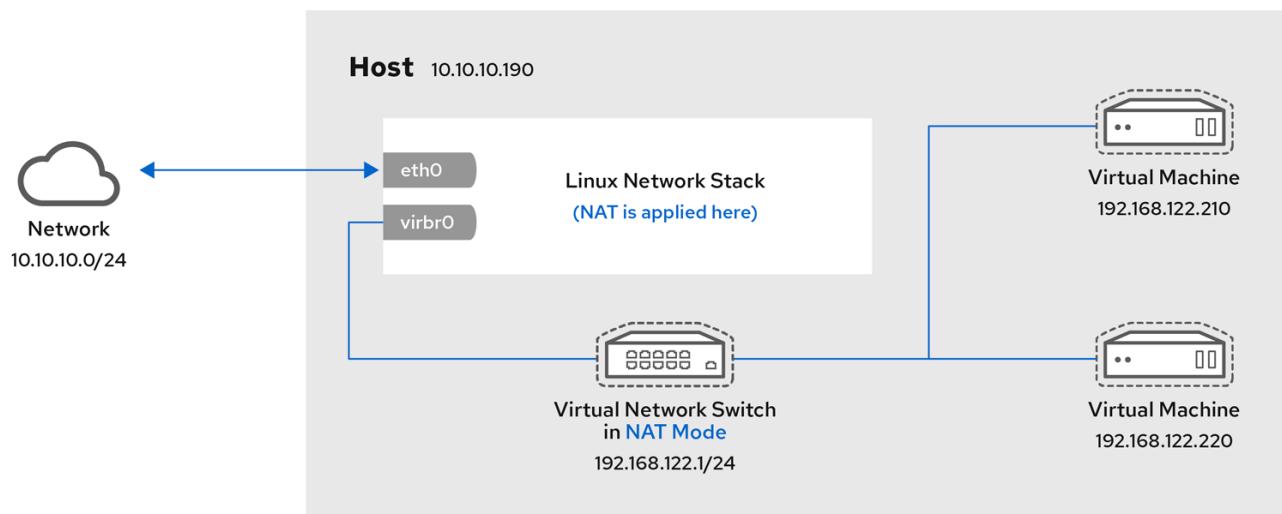
- 网桥模式中的虚拟网络

13.4. 虚拟机网络连接的类型

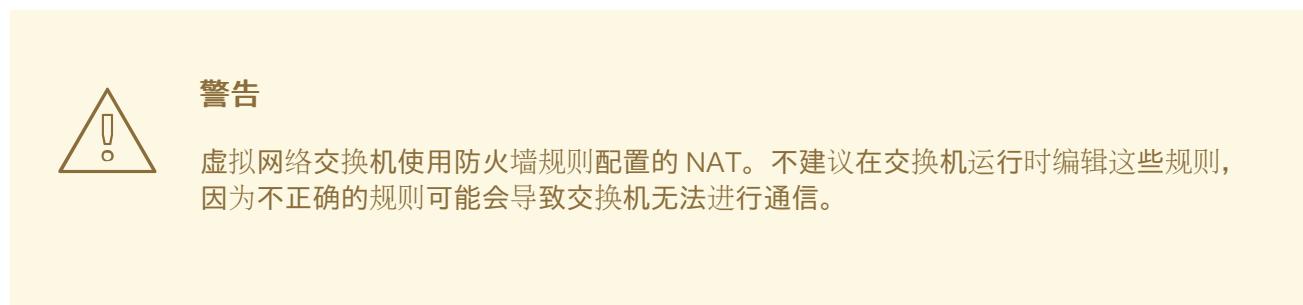
要修改虚拟机的网络属性和行为，更改虚拟机使用的虚拟网络或接口类型。以下小节描述了 RHEL 8 中虚拟机可用的连接类型。

13.4.1. 使用网络地址转换进行虚拟联网

默认情况下，虚拟网络交换机运行在网络地址转换(NAT)模式下。它们使用 IP 伪装而不是源-NAT(SNAT)或目的-NAT(DNAT)。IP 伪装可让连接的虚拟机使用主机机器的 IP 地址与任何外部网络通信。当虚拟网络交换机运行在 NAT 模式时，主机外的计算机无法与主机内的虚拟机进行通信。



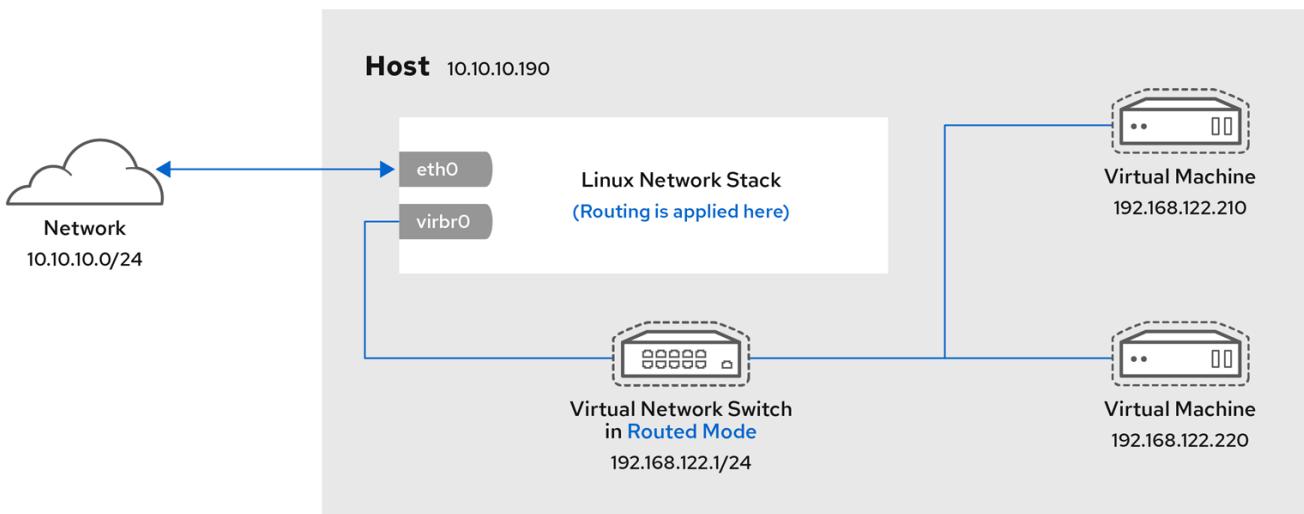
RHEL_52_1219



13.4.2. 路由模式中的虚拟网络

当使用 *Routed* 模式时，虚拟交换机会连接到连接到主机的物理 LAN 中，并在不使用 NAT 的情况下传输数据。虚拟交换机可以检查所有流量，并使用网络数据包中包含的信息来做出路由决策。使用此模式时，虚拟机(VM)都位于一个子网中，与主机是分开的。虚拟机子网通过虚拟交换机进行路由，该交换机位于主机上。这会启用传入连接，但对外部网络上的系统需要额外的路由表条目。

路由模式使用基于 IP 地址的路由：



RHEL_52_1219

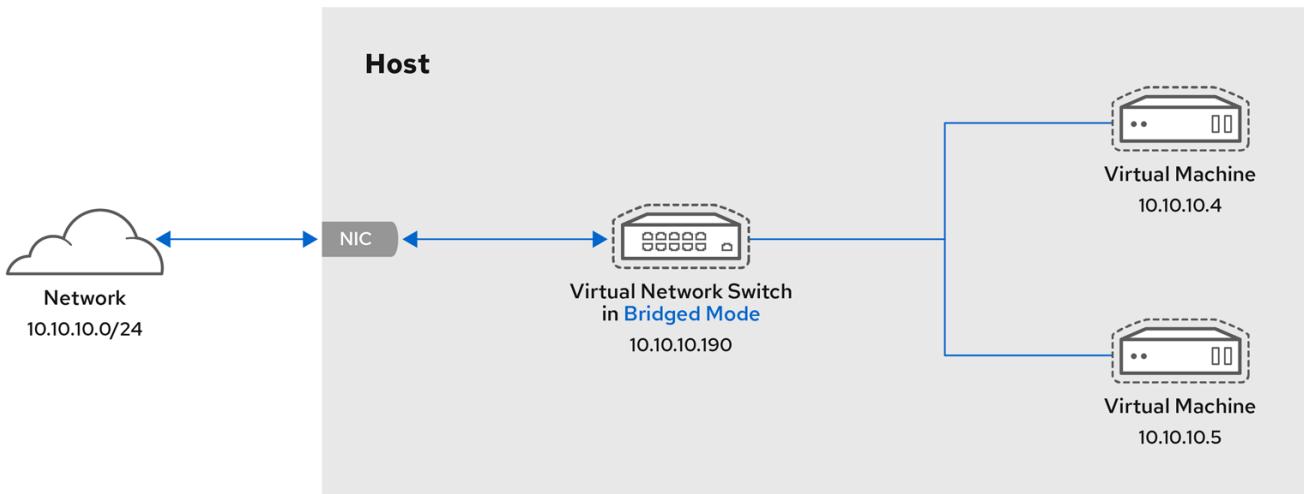
使用路由模式的通用拓扑是虚拟服务器托管(VSH)。VSH 提供程序可能有多个主机计算机，各自具有两个物理网络连接。一个接口用于管理和记帐，另一个用于虚拟机连接。每个虚拟机都有自己的公共 IP 地址，但主机使用私有的 IP 地址，以便只有内部管理员可以管理虚拟机。

+ image::vn-10-routed-mode-datacenter.png[]

13.4.3. 网桥模式中的虚拟网络

在大多数虚拟机网络模式中，虚拟机会自动创建并连接到 **virbr0** 虚拟网桥。相反，在 **桥接模式** 中，VM 会连接到主机上的一个已存在的 Linux 网桥。因此，虚拟机可以在物理网络中直接看到。这就可以允许进入的连接，但不需要任何额外的路由表条目。

网桥模式使用基于 MAC 地址的连接切换：



RHEL_52_1219

在桥接模式中，虚拟机出现在与主机相同的子网中。同一物理网络中的所有其他物理机器都可以检测虚拟机并访问它。

网桥网络绑定

通过使用绑定将多个物理网桥接口连在一起，可以在 hypervisor 上使用多个物理网桥接口。然后，可以将绑定添加到网桥，之后也可以将虚拟机添加到网桥。但是，绑定驱动程序有多种操作模式，而且并非所有这些模式都能与虚拟机正在使用的桥接一起工作。

以下 [绑定模式](#) 很有用：

- 模式 1
- 模式 2
- 模式 4

相反，使用模式 0、3、5 或 6 可能会导致连接失败。另请注意，应使用独立于介质的接口(MII)监控来监控绑定模式，因为地址解析协议(ARP)监控无法正常工作。

有关绑定模式的详情，请参考[红帽知识库](#)。

常见情况

使用桥接模式的最常见用例包括：

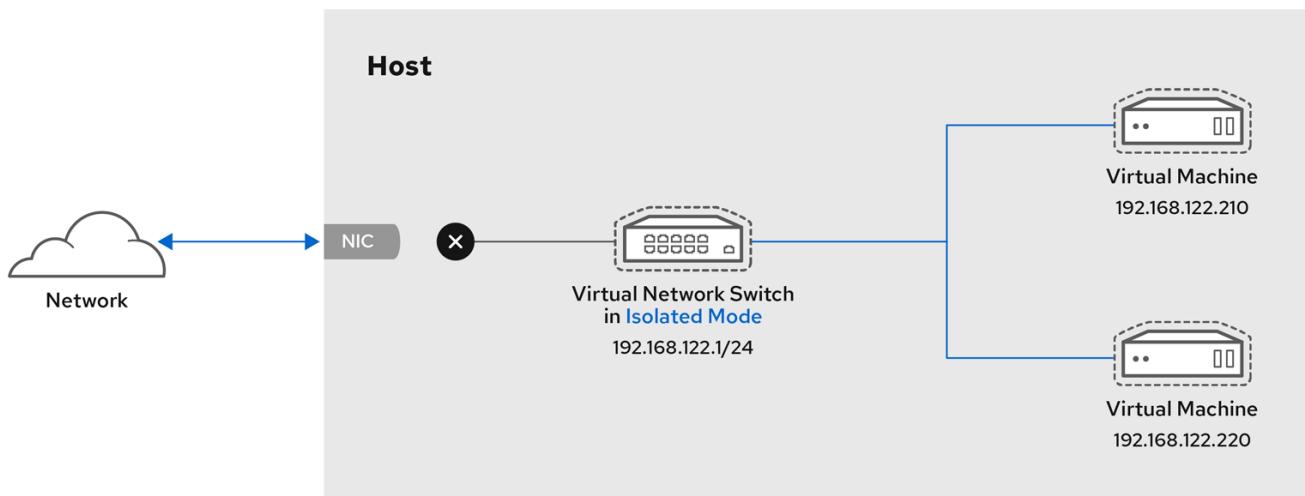
- 主机机器和虚拟机一起出现在现有网络中，最终用户看不到虚拟机和物理机器之间的不同。
- 在不更改现有物理网络配置设置的情况下部署虚拟机。
- 部署需要被现有物理网络轻松访问的虚拟机。将虚拟机放置到必须访问 DHCP 服务的物理网络中。
- 将虚拟机连接到使用虚拟 LAN(VLAN)的现有网络。
- 一个 demilitarized zone (DMZ) 网络。对于带有虚拟机的 DMZ 部署，红帽建议在物理网络路由器和交换机上设置 DMZ，并使用桥接模式将虚拟机连接到物理网络。

其它资源

- [使用命令行界面配置外部可见虚拟机](#)
- [使用 web 控制台配置外部可见虚拟机](#)
- [bridge_opts 参数的说明](#)

13.4.4. 隔离模式中的虚拟网络

使用 [隔离](#) 模式时，连接到虚拟交换机的虚拟机可以相互通信，并可以与主机进行通信，但它们的流量不会经过主机外部，而且它们无法接收来自主机外部的流量。对于基本功能，需要在这个模式中使用 [dnsmasq](#)，如 DHCP。



RHEL_52_1219

13.4.5. 开放模式中的虚拟网络

当对网络使用 **开放模式**时，**libvirt**不会为网络生成任何防火墙规则。因此，**libvirt** 不会覆盖主机提供的防火墙规则，因此用户可以手动管理虚拟机的防火墙规则。

13.4.6. 虚拟机连接类型的比较

下表提供了有关所选的虚拟机(VM)网络配置可以连接到的位置的信息，以及它们对哪些位置可见。

表 13.1. 虚拟机连接类型

	连接到主机	连接到主机上的其他虚拟机	连接到外部位置	可查看外部位置
网桥模式	是	是	是	是
NAT	是	是	是	否
路由模式	是	是	是	是
隔离模式	是	是	否	否
开放模式		取决于主机的防火墙规则		

13.5. 从 PXE 服务器启动虚拟机

使用预引导执行环境(PXE)的虚拟机可以从网络引导并加载它们的配置。本章描述了如何使用 **libvirt** 在虚拟或桥接网络上从 PXE 服务器引导虚拟机。



警告

这些流程仅作为示例提供。在操作之前，请确保您有足够的备份。

13.5.1. 在虚拟网络中设置 PXE 引导服务器

这个流程描述了如何配置 libvirt 虚拟网络以提供预启动执行环境(PXE)。这可让主机上的虚拟机被配置为从虚拟网络上提供的引导镜像引导。

先决条件

- 本地 PXE 服务器 (DHCP 和 TFTP)，例如：
 - libvirt 内部服务器
 - 手动配置 dhcpcd 和 tftpd
 - dnsmasq
 - Cobbler 服务器
- PXE 引导映像，如 Cobbler 配置的 **PXELINUX** 或手工进行配置。

流程

1. 将 PXE 引导镜像和配置放在 **/var/lib/tftpboot** 文件夹中。

2. 设置文件夹权限：

```
# chmod -R a+r /var/lib/tftpboot
```

3. 设置文件夹所有权：

```
# chown -R nobody: /var/lib/tftpboot
```

4. 更新 SELinux 上下文：

```
# chcon -R --reference /usr/sbin/dnsmasq /var/lib/tftpboot
# chcon -R --reference /usr/libexec/libvirt_leaseshelper /var/lib/tftpboot
```

5. 关闭虚拟网络：

```
# virsh net-destroy default
```

6. 在默认编辑器中打开虚拟网络配置文件：

```
# virsh net-edit default
```

7. 编辑 `<ip>` 元素以包含合适的地址、网络掩码、DHCP 地址范围和引导文件，其中 `BOOT_FILENAME` 是引导镜像文件的名称。

```
<ip address='192.168.122.1' netmask='255.255.255.0'>
  <tftp root='/var/lib/tftpboot' />
  <dhcp>
    <range start='192.168.122.2' end='192.168.122.254' />
    <bootp file='BOOT_FILENAME' />
  </dhcp>
</ip>
```

8. 启动虚拟网络：

```
# virsh net-start default
```

验证

- 验证 `default` 虚拟网络是否处于活跃状态：

```
# virsh net-list
Name      State  Autostart Persistent
-----
default   active  no      no
```

其它资源

- [准备使用 PXE 从网络安装](#)

13.5.2. 使用 PXE 和虚拟网络引导虚拟机

要从虚拟网络上的预引导执行环境(PXE)服务器引导虚拟机(VM)，您必须启用 PXE 引导。

先决条件

- 在虚拟网络上设置 PXE 引导服务器，如在[虚拟网络 中设置 PXE 引导服务器](#) 中所述。

流程

- 创建启用了 PXE 引导的新虚拟机。例如，若要从 `default` 虚拟网络上提供的 PXE 安装一个新的 10 GB qcow2 镜像文件：

```
# virt-install --pxe --network network=default --memory 2048 --vcpus 2 --disk size=10
```

- 另外，您可以手动编辑现有虚拟机的 XML 配置文件：

- 确保 `<os>` 元素中有一个 `<boot dev='network'>` 元素：

```
<os>
  <type arch='x86_64' machine='pc-i440fx-rhel7.0.0'>hvm</type>
  <boot dev='network' />
  <boot dev='hd' />
</os>
```

- ii. 确定客户端网络已配置为使用您的虚拟网络：

```
<interface type='network'>
  <mac address='52:54:00:66:79:14'/'>
  <source network='default'/'>
  <target dev='vnet0'/'>
  <alias name='net0'/'>
  <address type='pci' domain='0x0000' bus='0x00' slot='0x03' function='0x0'/'>
</interface>
```

验证

- 使用 **virsh start** 命令启动虚拟机。如果正确配置了 PXE，则虚拟机可以从 PXE 服务器上提供的引导镜像启动。

13.5.3. 使用 PXE 和桥接网络启动虚拟机

要从桥接网络上提供的预引导执行环境(PXE)服务器启动虚拟机(VM)，您必须启用 PXE 引导。

先决条件

- 启用了网络桥接。
- 网桥网络上提供了 PXE 引导服务器。

流程

- 创建启用了 PXE 引导的新虚拟机。例如，若要在 **breth0** 桥接网络上从 PXE 安装一个新的 10 GB qcow2 镜像文件：

```
# virt-install --pxe --network bridge=breth0 --memory 2048 --vcpus 2 --disk size=10
```

- 另外，您可以手动编辑现有虚拟机的 XML 配置文件：

- 确保 **<os>** 元素中有一个 **<boot dev='network'>** 元素：

```
<os>
  <type arch='x86_64' machine='pc-i440fx-rhel7.0.0'>hvm</type>
  <boot dev='network'/'>
  <boot dev='hd'/'>
</os>
```

- 确保虚拟机配置为使用桥接网络：

```
<interface type='bridge'>
  <mac address='52:54:00:5a:ad:cb'/'>
  <source bridge='breth0'/'>
  <target dev='vnet0'/'>
  <alias name='net0'/'>
  <address type='pci' domain='0x0000' bus='0x00' slot='0x03' function='0x0'/'>
</interface>
```

验证

- 使用 **virsh start** 命令启动虚拟机。如果正确配置了 PXE，则虚拟机可以从 PXE 服务器上提供的引导镜像启动。

其它资源

- [配置网络桥接](#)

13.6. 其它资源

- [配置和管理网络](#)
- [将特定的网络接口卡作为 SR-IOV 设备](#) 来提高虚拟机性能。

第 14 章 在主机及其虚拟机间共享文件

您可能需要在主机系统和其运行的虚拟机(VM)之间共享数据。要快速高效地完成此操作，您可以在系统上设置 NFS 文件共享。

14.1. 使用 NFS 在主机及其虚拟机间共享文件

为了在 RHEL 8 主机系统和连接到的虚拟机(VM)间有效文件共享，您可以导出虚拟机可以挂载和访问的 NFS 共享。

先决条件

- **nfs-utils** 软件包已安装在主机上。

```
# yum install nfs-utils -y
```

- 要与虚拟机共享的目录。如果您不想共享任何现有目录，请创建一个新目录，例如：*shared-files*。

```
# mkdir shared-files
```

- 当连接到虚拟机时，主机可以通过网络可见并可访问。通常，如果虚拟机使用 NAT 和 网桥类型虚拟网络，则会出现这种情况。
- 可选：为了提高安全性，请确保您的虚拟机与 NFS 版本 4 或更高版本兼容。

流程

1. 在主机上，将您要共享的文件的目录导出为网络文件系统(NFS)。

- a. 获取您要共享文件的每个虚拟机的 IP 地址。以下示例获取 *testguest1* 和 *testguest2* 的 IP 地址。

```
# virsh domifaddr testguest1
Name      MAC address      Protocol      Address
-----
vnet0    52:53:00:84:57:90  ipv4          192.168.124.220/24

# virsh domifaddr testguest2
Name      MAC address      Protocol      Address
-----
vnet1    52:53:00:65:29:21  ipv4          192.168.124.17/24
```

- b. 编辑主机上的 **/etc/exports** 文件，并添加一行，其中包含您要共享的目录、您要共享它的虚拟机 IP，以及共享选项。

```
<shared_directory> <VM1-IP(options)> <VM2-IP(options)> [...]
```

例如，以下命令在 *testguest1* 和 *testguest2* 主机上共享了 **/usr/local/shared-files** 目录，使虚拟机能够编辑该目录的内容：

```
/usr/local/shared-files/ 192.168.124.220(rw,sync) 192.168.124.17(rw,sync)
```



注意

如果要与 Windows 虚拟机共享目录，您必须确保 Windows NFS 客户端在共享目录中具有写入权限。执行此操作的一种简单方法是，使用 **/etc/exports** 文件中的 **all_squash**、**sonuid** 和 **anongid** 选项。

例如：

```
/usr/local/shared-files/
192.168.124.220(rw,sync,all_squash,anonuid=<directory-owner-
UID>,anongid=<directory-owner-GID>)
```

& *directory-owner-UID >* 和 *<directory-owner-GID >* 是拥有主机上共享目录的本地用户的 UID 和 GID。

要探索管理 NFS 客户端权限的其他选项，请遵循 [保护 NFS](#) 指南。

- 导出更新的文件系统。

```
# exportfs -a
```

- 确保 **nfs-server** 服务正在运行。

```
# systemctl start nfs-server
```

- 获取主机系统的 IP 地址。这可用于以后在虚拟机上挂载共享目录。

```
# ip addr
[...]
5: virbr0: [BROADCAST,MULTICAST,UP,LOWER_UP] mtu 1500 qdisc noqueue state UP
group default qlen 1000
    link/ether 52:54:00:32:ff:a5 brd ff:ff:ff:ff:ff:ff
    inet 192.168.124.1/24 brd 192.168.124.255 scope global virbr0
        valid_lft forever preferred_lft forever
[...]
```

请注意，相关的网络是您要与之共享文件的虚拟机用于连接主机的网络。通常，这是 **virbr0**。

- 将共享目录挂载到 **/etc/exports** 文件中指定的 Linux 虚拟机上。

```
# mount 192.168.124.1:/usr/local/shared-files /mnt/host-share
```

在本例中：

- 192.168.124.1** 是主机的 IP 地址。
- /usr/local/shared-files** 是主机上导出的目录的文件系统路径。
- /mnt/host-share** 是虚拟机的挂载点。挂载点必须是空目录。

- 要在 **/etc/exports** 文件中指定的 Windows 虚拟机上挂载共享目录：

- 以管理员身份打开 PowerShell shell 提示符。

- b. 安装 **NFS-Client** 软件包。对于 Windows 的服务器和桌面版本，安装命令会有所不同。
在 Windows 的服务器中：

```
# Install-WindowsFeature NFS-Client
```

在 Windows 的桌面版本中：

```
# Enable-WindowsOptionalFeature -FeatureName ServicesForNFS-ClientOnly,
ClientForNFS-Infrastructure -Online -NoRestart
```

- c. 将主机导出的目录挂载到 Windows 虚拟机上。

```
# C:\Windows\system32\mount.exe -o anon \\192.168.124.1\usr\local\shared-files Z:
```

在本例中：

- **192.168.124.1** 是主机的 IP 地址。
- **/usr/local/shared-files** 是主机上导出的目录的文件系统路径。
- **z**：是将用作挂载点的驱动器符。您必须选择一个没有在系统中使用的驱动器符。

验证

- 要验证您可以在主机和虚拟机间共享文件，请列出虚拟机上共享目录的内容。在以下示例中，将 `<mount_point>` 替换为挂载的共享目录的文件系统路径。

```
$ ls <mount_point>
shared-file1 shared-file2 shared-file3
```

其它资源

- [配置 NFS 服务器导出](#)
- [将 NFS 服务器配置为在防火墙后运行](#)
- [保护 NFS](#)

第 15 章 保护虚拟机

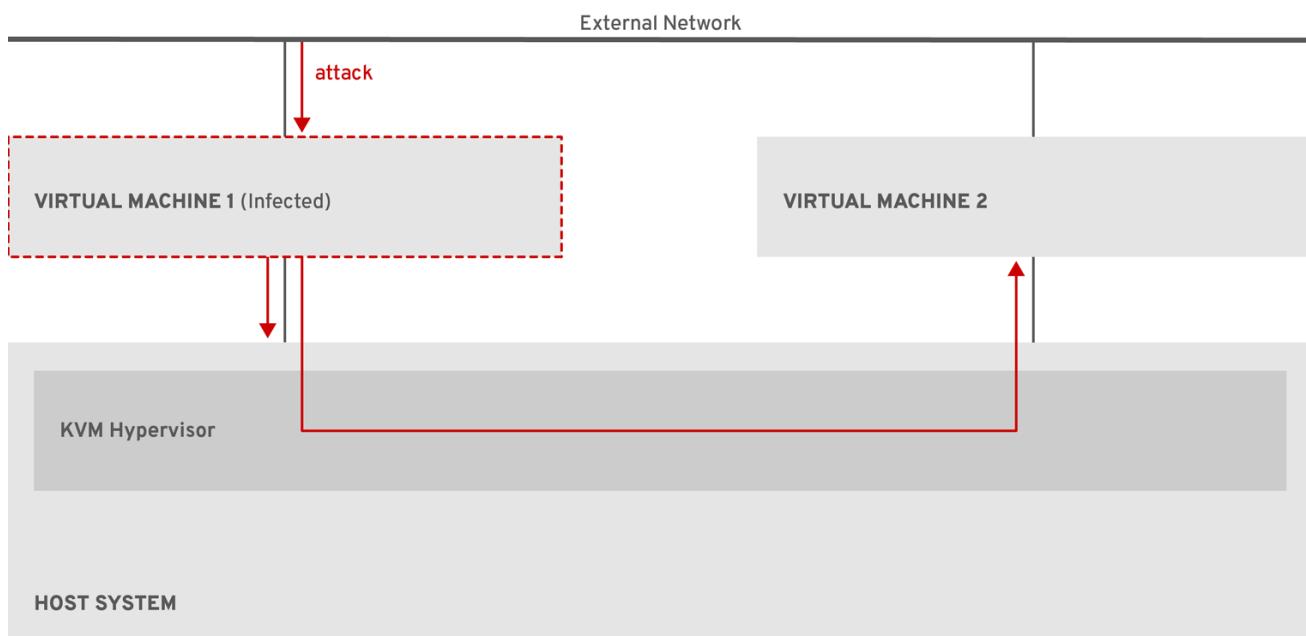
作为带有虚拟机(VM)的 RHEL 8 系统的管理员，确保您的虚拟机尽可能安全，降低虚拟机和主机操作系统受恶意软件破坏的风险。

本文档概述了 RHEL 8 主机上的 [保护虚拟机的机制](#)，并提供了 [方法列表](#) 来提高虚拟机的安全性。

15.1. 虚拟机中的安全性是如何工作的

通过使用虚拟机，可在单一主机机器中托管多个操作系统。这些系统通过 hypervisor 与主机连接，通常也通过虚拟网络连接。因此，每个虚拟机都可以用作使用恶意软件攻击主机的向量，主机可用作攻击任何虚拟机的向量。

图 15.1. 虚拟化主机上潜在的恶意攻击向量

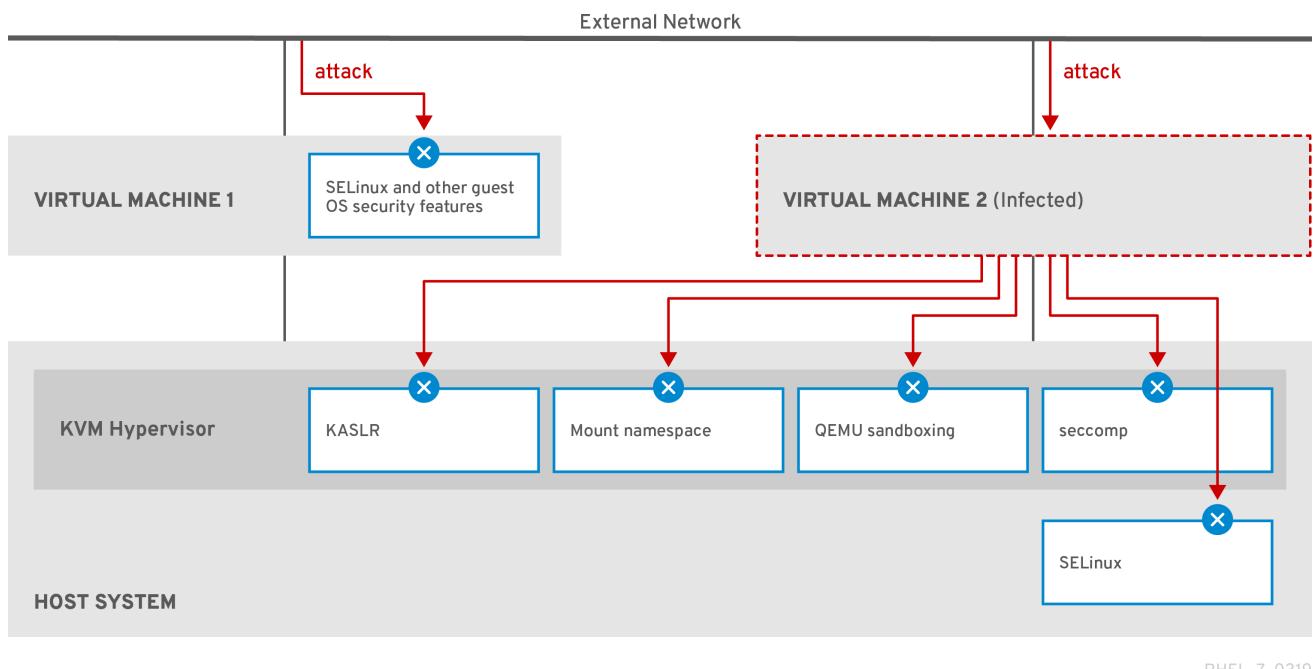


RHEL_7_0319

因为虚拟机监控程序使用主机内核来管理虚拟机，所以在虚拟机操作系统中运行的服务通常会被利用来将恶意代码注入主机系统。但是，您可以通过使用主机和您客户机系统上的 [一些安全特性](#) 来保护您的系统免受此类安全威胁。

这些特性（如 SELinux 或 QEMU 沙盒）提供了各种措施，使恶意代码难于攻击 hypervisor，并在您的主机和虚拟机之间进行传输。

图 15.2. 防止对虚拟化主机进行恶意软件攻击



RHEL_7_0319

RHEL 8 为虚拟机的安全性提供的许多功能始终处于活动状态，不需要启用或配置。[详情请查看虚拟机安全自动功能。](#)

此外，您可以遵循各种最佳实践来最大程度降低虚拟机和 hypervisor 的漏洞。如需更多信息，请参阅[保护虚拟机的最佳实践](#)。

15.2. 保护虚拟机的最佳实践

根据以下步骤，您的虚拟机被恶意代码利用，并用作攻击向量攻击您的主机系统的风险会大幅降低。

在客户端中：

- 象保护物理机器一样保护虚拟机的安全。增强安全性的具体方法取决于客户端操作系统。如果您的虚拟机正在运行 RHEL 8，请参阅[保护 Red Hat Enterprise Linux 8](#) 以了解有关提高客户机系统安全性的详细说明。

在主机端：

- 当远程管理虚拟机时，请使用加密的工具（如 **SSH**）和网络协议（如 **SSL**）连接到虚拟机。
- 确定 SELinux 处于 Enforcing 模式：

```
# getenforce
Enforcing
```

如果 SELinux 被禁用或者处于 *Permissive* 模式，请参阅[使用 SELinux](#) 文档来激活 Enforcing 模式。



注意

SELinux Enforcing 模式也启用了 sVirt RHEL 8 特性。这是用于虚拟化的一组特殊的 SELinux 布尔值，可以[手动调整](#)以进行精细的虚拟机安全管理。

- 使用带有 *SecureBoot* 的虚拟机：

SecureBoot 是一种特性，可确保虚拟机运行加密签名的操作系统。这可防止因为恶意软件攻击而更改了操作系统的虚拟机引导。

SecureBoot 只能在安装使用 OVMF 固件的 Linux 虚拟机时使用。具体步骤请参阅 [创建 SecureBoot 虚拟机](#)。

- 不要使用 **qemu-*** 命令，如 **qemu-kvm**。

QEMU 是 RHEL 8 中虚拟化架构的基本组件，但难以手动管理，而不正确的 QEMU 配置可能会导致安全漏洞。因此，红帽不支持使用 **qemu-*** 命令。反之，使用 *libvirt* 工具，如 **virsh**、**virt-install** 和 **virt-xml**，根据最佳实践来编排 QEMU。

其它资源

- [RHEL 中虚拟化的 SELinux 布尔值](#)

15.3. 创建 SECUREBOOT 虚拟机

您可以创建一个使用 *SecureBoot* 特性的 Linux 虚拟机(VM)，这将确保您的虚拟机运行加密签名的操作系统。如果虚拟机的客户机操作系统已被恶意软件更改，这将非常有用。在这种场景下，*SecureBoot* 会阻止虚拟机启动，从而停止可能将恶意软件传播给您的主机。

先决条件

- 虚拟机使用 Q35 机器类型。
- **edk2-OVMF** 软件包已安装：

```
# yum install edk2-ovmf
```

- 操作系统 (OS) 安装源可存在于本地或者网络中。可以是以下格式之一：
 - 安装介质的 ISO 镜像
 - 现有虚拟机安装的磁盘镜像



警告

在 RHEL 8 中无法从主机 CD-ROM 或者 DVD-ROM 设备安装。当使用 RHEL 8 中的任何虚拟机安装方法时，如果选择了 CD-ROM 或者 DVD-ROM 作为安装源，则安装将失败。如需更多信息，请参阅[红帽知识库](#)。

- 可选：对于快速、简单的配置安装，可以使用 Kickstart 文件。

流程

1. 使用 **virt-install** 命令来创建虚拟机，如 [使用命令行界面“创建虚拟机”](#)。对于 **--boot** 选项，请使用 **uefi,nvram_template=/usr/share/OVMF/OVMF_VARS.secboot.fd** 值。这会使用 **OVMF_VARS.secboot.fd** 和 **OVMF_CODE.secboot.fd** 文件作为虚拟机非易失性 RAM(NVRAM)设置的模板，该设置启用了 *SecureBoot* 特性。

例如：

```
# virt-install --name rhel8sb --memory 4096 --vcpus 4 --os-variant rhel8.0 --boot uefi,nvram_template=/usr/share/OVMF/OVMF_VARS.secboot.fd --disk boot_order=2,size=10 --disk boot_order=1,device=cdrom,bus=scsi,path=/images/RHEL-8.0-installation.iso
```

- 根据屏幕上的说明，按照操作系统安装过程进行操作。

验证

- 客户机操作系统安装后，通过打开 [图形客户机控制台](#) 中的终端或 [使用 SSH](#) 连接到客户机 OS，以访问虚拟机的命令行。
- 要确认是否在虚拟机上启用了 SecureBoot，请使用 **mokutil --sb-state** 命令：

```
# mokutil --sb-state
SecureBoot enabled
```

其它资源

- [在 AMD64、Intel 64 和 64 位 ARM 上安装 RHEL 8](#)

15.4. 限制虚拟机用户可以使用哪些操作

在某些情况下，在 RHEL 8 上托管的虚拟机(VM)用户可以执行的默认操作可能会导致安全隐患。如果是这种情况，您可以通过配置 **libvirt** 守护进程配置来在主机上使用 **polkit** 策略工具包来限制虚拟机用户可用的操作。

流程

- 可选：确保与 **libvirt** 有关的系统 **polkit** 控制策略根据您的偏好而设置。
 - 在 **/usr/share/polkit-1/actions/** 和 **/usr/share/polkit-1/rules.d/** 目录中找到所有与 libvirt 相关的文件。


```
# ls /usr/share/polkit-1/actions | grep libvirt
# ls /usr/share/polkit-1/rules.d | grep libvirt
```
 - 打开文件并检查规则设置。
有关读取 **polkit** 控制策略语法的信息，请使用 **man polkit**。
 - 修改 **libvirt** 控制策略。要做到这一点：
 - 在 **/etc/polkit-1/rules.d/** 目录中创建一个新的 **.rules** 文件。
 - 将自定义策略添加到此文件中，并保存。
有关 **libvirt** 控制策略的更多信息和示例，请参阅 [libvirt 上游文档](#)。
- 配置您的虚拟机，以使用由 **polkit** 确定的访问策略。
为此，请取消对 **/etc/libvirt/libvirtd.conf** 文件中的 **access_drivers = ["polkit"]** 行的注释。

```
# sed -i 's/#access_drivers = \[ "polkit" \]/access_drivers = \[ "polkit" \]/' /etc/libvirt/libvirtd.conf
```

3. 重新启动 **libvirdt** 服务。

```
# systemctl restart libvirtd
```

验证

- 作为您要限制其虚拟机操作的用户，请执行一个受限操作。
例如，如果非特权用户被限制查看在系统会话中创建的虚拟机：

```
$ virsh -c qemu:///system list --all
Id  Name      State
-----
```

如果这个命令没有列出任何虚拟机，即使系统上有一个或多个虚拟机，则 **polkit** 成功地限制了非特权用户的操作。

故障排除

- 目前，配置 **libvirt** 以使用 **polkit**，使它不可能使用 [RHEL 8 web 控制台](#) 连接到虚拟机，因为与 **libvirt-dbus** 服务不兼容。
如果您需要在 web 控制台中对虚拟机进行精细访问控制，请创建一个自定义 D-Bus 策略。具体说明请参阅红帽知识库中的 [如何在 Cockpit 中配置虚拟机的精细控制](#)。

其它资源

- **man polkit** 命令
- 有关 [polkit 访问控制策略的 libvirt 上游信息](#)

15.5. 虚拟机安全性的自动功能

除了手动方式提高了您在 [保护虚拟机的最佳实践](#) 中列出的虚拟机的安全性外，很多安全功能是由 **libvirt** 软件套件提供的，并在 RHEL 8 中使用虚拟化时自动启用。它们是：

系统和会话连接

要访问 RHEL 8 中所有可用的虚拟机管理工具，您需要使用 **libvirt** 的 **系统连接**(**qemu:///system**)。为此，您必须在系统上拥有 root 权限，或者是 **libvirt** 用户组的成员。

不属于 **libvirt** 组中的非 root 用户只能访问 **libvirt** 的 **会话连接**(**qemu:///session**)，后者必须在访问资源时遵守本地用户的访问权限。例如，使用会话连接，您无法检测或访问系统连接中或由其他用户创建的虚拟机。另外，可用的 VM 网络配置选项也有很大限制。



注意

RHEL 8 文档假设您有系统连接特权。

虚拟机分离

单个虚拟机作为隔离进程在主机上运行，并依赖于主机内核强制的安全性。因此，虚拟机无法读取或访问同一主机上其他虚拟机的内存或存储。

QEMU 沙盒

此特性可防止 QEMU 代码执行可能会破坏主机安全性的系统调用。

内核地址空间随机化(KASLR)

启用对内核镜像解压缩的物理和虚拟地址进行随机化。因此，KASLR 会根据内核对象的位置防止客户机安全漏洞。

15.6. 用于虚拟化的 SELinux 布尔值

要在 RHEL 8 系统上精细配置虚拟机的安全性，您可以在主机上配置 SELinux 布尔值，以确保 hypervisor 以特定的方式工作。

要列出所有与虚拟化相关的布尔值及其状态，请使用 **getsebool -a | grep virt** 命令：

```
$ getsebool -a | grep virt
[...]
virt_sandbox_use_netlink --> off
virt_sandbox_use_sys_admin --> off
virt_transition_userdomain --> off
virt_use_comm --> off
virt_use_execmem --> off
virt_use_fusefs --> off
[...]
```

要启用特定的布尔值，请以 root 用户身份使用 **setsebool -P boolean_name on** 命令。要禁用布尔值，请使用 **setsebool -P boolean_name off**。

下表列出了 RHEL 8 中提供的与虚拟化相关的布尔值，以及它们在启用时的作用：

表 15.1. SELinux 虚拟化布尔值

SELinux 布尔值	描述
staff_use_svirt	启用非 root 用户创建并转换虚拟机至 sVirt。
unprivuser_use_svirt	启用非特权用户创建虚拟机并将其转换至 sVirt。
virt_sandbox_use_audit	启用沙盒容器来发送审核信息。
virt_sandbox_use_netlink	启用沙盒容器使用 netlink 系统调用。
virt_sandbox_use_sys_admin	启用沙盒容器使用 sys_admin 系统调用，如 mount。
virt_transition_userdomain	启用虚拟进程作为用户域运行。
virt_use_comm	启用 virt 使用串行/并行通信端口。
virt_use_execmem	使受限的虚拟客户机可以使用可执行的内存和可执行的堆栈。
virt_use_fusefs	启用 virt 读取 FUSE 挂载的文件。
virt_use_nfs	启用 virt 管理 NFS 挂载的文件。
virt_use_rawip	启用 virt 与 rawip 套接字交互。

SELinux 布尔值	描述
virt_use_samba	启用 virt 管理 CIFS 挂载的文件。
virt_use_sanlock	启用受限制的虚拟客户机与 sanlock 交互。
virt_use_usb	启用 virt 使用 USB 设备。
virt_use_xserver	启用虚拟机与 X 窗口系统交互。

15.7. 在 IBM Z 上设置 IBM SECURE EXECUTION

当使用 IBM Z 硬件来运行 RHEL 8 主机时，您可以通过为虚拟机配置 IBM 安全执行来提高虚拟机(VM)的安全性。

IBM Secure Execution（也称 Protected Virtualization）可防止主机系统访问虚拟机的状态和内存内容。因此，即使主机被攻击，也无法用作攻击客户端操作系统的向量。另外，安全执行也可以用来防止不可信主机从虚拟机获取敏感信息。

以下流程描述了如何将 IBM Z 主机上的现有虚拟机转换为安全虚拟机。

先决条件

- 系统硬件是以下之一：
 - IBM z15 或更高版本
 - IBM LinuxONE III 或更高版本
- 为您的系统启用安全执行功能。要验证，请使用：

```
# grep facilities /proc/cpuinfo | grep 158
```

如果这个命令显示任何输出，代表您的 CPU 与安全执行兼容。

- 内核包含对安全执行的支持。要确认，请使用：

```
# ls /sys/firmware | grep uv
```

如果命令生成任何输出，那么内核支持安全执行。

- 主机 CPU 型号包含 **解包** 功能。要确认，请使用：

```
# virsh domcapabilities | grep unpack
<feature policy='require' name='unpack'/>
```

如果命令生成上述输出，则您的 CPU 主机型号与安全执行兼容。

- 虚拟机的 CPU 模式设置为 **host-model**。若要对此进行确认，请使用以下命令，并将 **vm-name** 替换为您的虚拟机的名称：

```
# virsh dumpxml vm-name | grep "<cpu mode='host-model'>"
```

如果命令生成任何输出，则虚拟机的 CPU 模式设置正确。

- 您已获取并验证了 IBM Z 主机密钥文档。有关此操作的说明，请参阅 IBM 文档中的 [验证主机密钥文档](#)。

流程

在主机上 执行以下步骤：

1. 将 `prot_virt=1` 内核参数添加到主机的 [引导配置](#) 中。

```
# grubby --update-kernel=ALL --args="prot_virt=1"
```

2. 更新引导菜单：

```
# zipl
```

3. 使用 `virsh edit` 修改您要保护的虚拟机的 XML 配置。

4. 将 `<launchSecurity type="s390-pv"/>` 添加到 `</ devices>` 行下的。例如：

```
[...]
</memballoon>
</devices>
<launchSecurity type="s390-pv"/>
</domain>
```

5. 如果配置的 `<devices>` 部分包含 `virtio-rng` device (`<rng model="virtio">`)，请删除 `< rng>` `</rng>` 块的所有行。

在您要保护的虚拟机的 [客户机操作系统中](#) 执行以下步骤。

1. 创建一个参数文件。例如：

```
# touch ~/secure-parameters
```

2. 在 `/boot/loader/entries` 目录中，识别最新版本的引导加载程序条目：

```
# ls /boot/loader/entries -l
[...]
-rw-r--r--. 1 root root 281 Oct 9 15:51 3ab27a195c2849429927b00679db15c1-4.18.0-
240.el8.s390x.conf
```

3. 检索引导装载程序条目的内核选项行：

```
# cat /boot/loader/entries/3ab27a195c2849429927b00679db15c1-4.18.0-
240.el8.s390x.conf | grep options
options root=/dev/mapper/rhel-root
crashkernel=auto
rd.lvm.lv=rhel/root rd.lvm.lv=rhel/swap
```

4. 将 `options` 行的内容和 `swiotlb=262144` 添加到创建的参数文件中。

```
# echo "root=/dev/mapper/rhel-root crashkernel=auto rd.lvm.lv=rhel/root
rd.lvm.lv=rhel/swap swiotlb=262144" > ~/secure-parameters
```

5. 生成一个 IBM 安全执行镜像。

例如，以下命令会根据 `/boot/vmlinuz-4.18.0-240.el8.s390x` 映像，使用 `secure-parameters` 文件创建 `/boot/secure-image` 安全镜像、`/boot/initramfs-4.18.0-240.el8.s390x.img` 初始 RAM 磁盘文件和 `HKD-8651-000201C048.crt` 主机密钥文档。

```
# genprotimg -i /boot/vmlinuz-4.18.0-240.el8.s390x -r /boot/initramfs-4.18.0-
240.el8.s390x.img -p ~/secure-parameters -k HKD-8651-00020089A8.crt -o
/boot/secure-image
```

使用 `genprotimg` 工具创建安全镜像，其中包含内核参数、初始 RAM 磁盘和引导镜像。

6. 更新虚拟机的引导菜单，以从安全镜像引导。此外，删除以 `initrd` 和 `options` 开头的行，因为不需要它们。

例如，在 RHEL 8.3 虚拟机中，可以在 `/boot/loader/entries/` 目录中编辑引导菜单：

```
# cat /boot/loader/entries/3ab27a195c2849429927b00679db15c1-4.18.0-
240.el8.s390x.conf
title Red Hat Enterprise Linux 8.3
version 4.18.0-240.el8.s390x
linux /boot/secure-image
[...]
```

7. 创建可引导磁盘镜像：

```
# zipl -V
```

8. 安全地删除原始的未保护的文件。例如：

```
# shred /boot/vmlinuz-4.18.0-240.el8.s390x
# shred /boot/initramfs-4.18.0-240.el8.s390x.img
# shred secure-parameters
```

原始引导镜像、初始 RAM 镜像和内核参数文件未受保护，如果未删除它们，则启用了安全执行的虚拟机仍然容易受到黑客攻击或敏感数据挖掘的攻击。

验证

- 在主机上，使用 `virsh dumpxml` 工具确认受保护虚拟机的 XML 配置。配置必须包含 `<launchSecurity type="s390-pv"/>` 元素，且没有 `<rng model="virtio"/>` 行。

```
# virsh dumpxml vm-name
[...]
<cpu mode='host-model'/>
<devices>
  <disk type='file' device='disk'>
    <driver name='qemu' type='qcow2' cache='none' io='native'>
      <source file='/var/lib/libvirt/images/secure-guest.qcow2' />
      <target dev='vda' bus='virtio' />
    </disk>
  <interface type='network'>
```

```

<source network='default'/>
<model type='virtio' />
</interface>
<console type='pty' />
<memballoon model='none' />
</devices>
<launchSecurity type="s390-pv"/>
</domain>

```

其它资源

- [有关启动安全虚拟服务器的 IBM 文档](#)
- [IBM 关于 genprotimg 的文档](#)
- [配置内核命令行参数](#)

15.8. 将加密 COPROCESSORS 附加到 IBM Z 上的虚拟机

要在 IBM Z 主机上的虚拟机中使用硬件加密，请从加密的 coprocessor 设备创建介质设备并将其分配给预期的虚拟机。具体步骤请查看以下说明。

先决条件

- 您的主机运行在 IBM Z 硬件上。
- 加密 coprocessor 与设备分配兼容。要进行确认，请确保协处理器的类型列为 **CEX4** 或更高版本。

```
# lszcrypt -V

CARD.DOMAIN TYPE MODE      STATUS REQUESTS PENDING HWTYPE QDEPTH
FUNCTIONS DRIVER

-----
05      CEX5C CCA-Coproc online    1      0     11    08 S--D--N-- cex4card
05.0004  CEX5C CCA-Coproc online    1      0     11    08 S--D--N-- cex4queue
05.00ab  CEX5C CCA-Coproc online    1      0     11    08 S--D--N-- cex4queue
```

- *mdevctl* 软件包已安装。
- **vfio_ap** 内核模块已加载。要验证，请使用：

```
# lsmod | grep vfio_ap
vfio_ap      24576  0
[...]
```

要载入模块，请使用：

```
# modprobe vfio_ap
```

- **s390utils** 版本支持 **ap** 处理：

```
# lszdev --list-types
...
```

```
ap      Cryptographic Adjunct Processor (AP) device
...
...
```

流程

1. 获取您要分配给虚拟机的设备的十进制值。例如，对于设备 **05.0004** 和 **05.00ab**：

```
# echo "obase=10; ibase=16; 04" | bc
4
# echo "obase=10; ibase=16; AB" | bc
171
```

2. 在主机上，将设备重新分配给 **vfio-ap** 驱动程序：

```
# chzdev -t ap apmask=-5 aqmask=-4,-171
```



注意

要永久分配设备，请使用 **-p** 标志。

3. 验证是否正确分配了加密设备。

```
# lszcrypt -V
CARD.DOMAIN TYPE MODE      STATUS REQUESTS PENDING HWTYPE QDEPTH
FUNCTIONS DRIVER
-----
05      CEX5C CCA-Coproc -      1    0   11   08 S--D--N-- cex4card
05.0004  CEX5C CCA-Coproc -      1    0   11   08 S--D--N-- vfio_ap
05.00ab  CEX5C CCA-Coproc -      1    0   11   08 S--D--N-- vfio_ap
```

如果域队列的 DRIVER 值变为 **vfio_ap**，则重新分配成功。

4. 生成设备 UUID。

```
# uuidgen
669d9b23-fe1b-4ecb-be08-a2fabca99b71
```

在此流程的下列步骤中，将 **669d9b23-fe1b-4ecb-be08-a2fabca99b71** 替换为您生成的 UUID。

5. 使用 UUID，创建一个新的 **vfio_ap** 设备。

以下示例演示了创建永久介质设备并为它分配队列。例如，以下命令可将域适配器 **0x05** 和域队列 **0x0004** 和 **0x00ab** 分配给设备 **669d9b23-fe1b-4ecb-be08-a2fabca99b71**。

```
# mdevctl define --uuid 669d9b23-fe1b-4ecb-be08-a2fabca99b71 --parent matrix --type
vfio_ap-passthrough
# mdevctl modify --uuid 669d9b23-fe1b-4ecb-be08-a2fabca99b71 --
addattr=assign_adapter --value=0x05
# mdevctl modify --uuid 669d9b23-fe1b-4ecb-be08-a2fabca99b71 --
addattr=assign_domain --value=0x0004
# mdevctl modify --uuid 669d9b23-fe1b-4ecb-be08-a2fabca99b71 --
addattr=assign_domain --value=0x00ab
```

6. 启动中介设备。

```
# mdevctl start --uuid 669d9b23-fe1b-4ecb-be08-a2fabca99b71
```

7. 检查配置是否已正确应用

```
# cat /sys/devices/vfio_ap/matrix/mdev_supported_types/vfio_ap-
passthrough/devices/669d9b23-fe1b-4ecb-be08-a2fabca99b71/matrix
05.0004
05.00ab
```

如果输出中包含您之前分配给 **vfio-ap** 的队列的数字值，则该过程成功。

8. 将介质设备附加到虚拟机。

- 为加密策略中介设备创建并打开 XML 文件。例如：

```
# vim crypto-dev.xml
```

- 在文件中添加以下行并保存：

```
<hostdev mode='subsystem' type='mdev' managed='no' model='vfio-ap'>
<source>
<address uuid='669d9b23-fe1b-4ecb-be08-a2fabca99b71'/>
</source>
</hostdev>
```

- 使用 XML 文件将介质设备附加到虚拟机。例如，要将 **crypto-dev.xml** 文件中定义的设备永久附加到正在运行的 **testguest1** 虚拟机：

```
# virsh attach-device testguest1 crypto-dev.xml --live --config
```

--live 选项仅将设备附加到正在运行的虚拟机中，且引导之间没有持久性。**--config** 选项使配置更改持久。您可以只使用 **--config** 选项将设备附加到关闭的虚拟机。

请注意，每个 UUID 每次只能分配给一个虚拟机。

9. 将所需的控制域分配给介质设备。

```
# mdevctl modify --uuid 669d9b23-fe1b-4ecb-be08-a2fabca99b71 --
addattr=assign_control_domain --value=0x00ab
# mdevctl modify --uuid 669d9b23-fe1b-4ecb-be08-a2fabca99b71 --
addattr=assign_control_domain --value=0x0004
```

验证

1. 确定客户端操作系统检测到分配的加密设备。

```
# lszcrypt -V
```

CARD.DOMAIN	TYPE	MODE	STATUS	REQUESTS	PENDING	HWTTYPE	QDEPTH

```
05      CEX5C CCA-Coproc online    1      0    11   08 S--D--N-- cex4card
05.0004  CEX5C CCA-Coproc online    1      0    11   08 S--D--N-- cex4queue
05.00ab  CEX5C CCA-Coproc online    1      0    11   08 S--D--N-- cex4queue
```

客户端操作系统中的这个命令的输出将与在有相同加密 coprocessor 设备的主机逻辑分区中相同。

2. 在客户端操作系统中，确认已成功为加密设备分配了控制域。

```
# lszcrypt -d C
```

```
DOMAIN 00 01 02 03 04 05 06 07 08 09 0a 0b 0c 0d 0e 0f
```

```
00 . . . . .
10 . . . . .
20 . . . . . B . .
30 . . B B . . . .
40 . . . . .
50 . . . . .
60 . . . . .
70 . . . . .
80 . . . . .
90 . . . . .
a0 . . . . .
b0 . . . . .
c0 . . . . .
d0 . . . . .
e0 . . . . .
f0 . . . . .
```

如果 **lszcrypt -d C** 在加密设备列表中显示 **B intersections**，则控制域分配成功。

15.9. 在 WINDOWS 虚拟机中启用标准硬件安全性

要保护 Windows 虚拟机，您可以使用 Windows 设备的标准硬件功能启用基本级别的安全性。

先决条件

- 请确定您安装了最新的 WHQL 认证的 VirtIO 驱动程序。
- 确保虚拟机固件支持 UEFI 引导。
- 在您的主机上安装 **edk2-OVMF** 软件包。

```
# yum install edk2-ovmf
```

- 在您的主机上安装 **vTPM** 软件包。

```
# yum install swtpm libtpms
```

- 确保虚拟机使用 Q35 机器架构。
- 请确定您有 Windows 安装介质。

流程

- 通过在虚拟机 XML 配置中的 `<devices>` 部分中添加以下参数来启用 TPM 2.0。

```
<devices>
[...]
<tpm model='tpm-crb'>
    <backend type='emulator' version='2.0' />
</tpm>
[...]
</devices>
```

- 在 UEFI 模式中安装 Windows。有关如何操作的更多信息，请参阅[创建 SecureBoot 虚拟机](#)。
- 在 Windows 虚拟机上安装 VirtIO 驱动程序。有关如何操作的更多信息，请参阅[在 Windows 客户端上安装 virtio 驱动程序](#)。
- 在 UEFI 中，启用安全引导。有关如何进行此操作的更多信息，请参阅[安全引导](#)。

验证

- 确定 Windows 机器中的设备安全性页面显示以下信息：
`settings > Update & Security > Windows Security > device Security`

Your device meets the requirements for standard hardware security.

15.10. 在 WINDOWS 虚拟机上启用增强的硬件安全

为进一步保护 Windows 虚拟机(VM)，您可以启用基于虚拟化的代码完整性保护，也称为 hypervisor 保护的代码完整性(HVCI)。

先决条件

- 确保启用了标准硬件安全。如需更多信息，请参阅[Windows 虚拟机上启用标准硬件安全性](#)。
- 确保启用了 Hyper-V enlightenments。如需更多信息，请参阅[启用 Hyper-Vlightenments](#)。

流程

- 打开 Windows 虚拟机的 XML 配置。以下示例打开 *Example-L1* 虚拟机的配置：

```
# virsh edit Example-L1
```

- 在 `<<cpu>>` 部分下，指定 CPU 模式并添加策略标记。



重要

- 对于 Intel CPU，启用 **vmx** 策略标记。
- 对于 AMD CPU，启用 **svm** 策略标记。
- 如果您不想指定自定义 CPU，您可以将 `<cpu mode=>` 设置为 **host-passthrough**。

```
<cpu mode='custom' match='exact' check='partial'>
  <model fallback='allow'>Skylake-Client-IBRS</model>
  <topology sockets='1' dies='1' cores='4' threads='1' />
  <feature policy='require' name='vmx' />
</cpu>
```

3. 保存 XML 配置并重启虚拟机。
4. 在虚拟机操作系统中，导航到 **Core 隔离详情页面**：
`settings > Update & Security > Windows Security > Device Security > Core isolated details`
5. 切换开关以启用 **内存完整性**。
6. 重启虚拟机。



注意

有关启用 HVCI 的其他方法，请查看相关的 Microsoft 文档。

验证

- 确保 Windows 虚拟机上的 **Device Security** 页面显示以下信息：
`settings > Update & Security > Windows Security > device Security`

Your device meets the requirements for enhanced hardware security.
- 或者，检查 Windows 虚拟机上的系统信息：
 - a. 在命令提示符下运行 **msinfo32.exe**。
 - b. 检查是否在 **基于虚拟化的安全服务运行** 下列出了 **凭据保护**，**Hypervisor 强制的代码完整性**。

第 16 章 优化虚拟机性能

与主机相比，虚拟机的性能总会有所降低。以下章节解释了导致这种恶化的原因，并提供了有关如何在 RHEL 8 中将虚拟化的性能影响降到最低的说明，以便您的硬件基础架构资源能尽可能被高效地使用。

16.1. 影响虚拟机性能的因素

虚拟机作为用户空间进程在主机上运行。因此管理程序需要转换主机的系统资源，以便虚拟机可使用它们。因此，部分资源会被转换消耗，因此虚拟机无法获得与主机相同的性能效率。

虚拟化对系统性能的影响

虚拟机性能损失的原因包括：

- 虚拟 CPU (vCPU) 是主机上的线程，由 Linux 调度程序处理。
- VM 不会自动继承主机内核的优化功能，比如 NUMA 或巨页。
- 主机的磁盘和网络 I/O 设置可能会对虚拟机有显著的性能影响。
- 网络流量通常通过基于软件的网桥到达虚拟机。
- 根据主机设备及其型号，特定硬件的模拟可能会产生大量的开销。

虚拟化对虚拟机性能影响的严重程度受到各种因素的影响，具体包括：

- 并行运行的虚拟机数量。
- 每个虚拟机使用的虚拟设备数量。
- 虚拟机使用的设备类型。

降低虚拟机性能损失

RHEL 8 提供很多功能，可用于降低虚拟化的负面影响。值得注意的是：

- [Tuned 服务](#) 可自动优化虚拟机的资源分布和性能。
- [块 I/O 调优](#) 可以提高虚拟机块设备（如磁盘）的性能。
- [NUMA 调优](#) 可以提高 vCPU 的性能。
- 可以通过多种方式优化 [虚拟网络](#)。



重要

调整虚拟机性能会对其他虚拟化功能造成负面影响。例如，它可以使迁移修改过的虚拟机更为困难。

16.2. 使用 Tuned 优化虚拟机性能

Tuned 程序是一种调优配置集交付机制，它可针对某些工作负载特性调整 RHEL，比如对 CPU 密集型任务或存储网络吞吐量响应的要求。它提供很多预先配置的调优配置文件，以便在多个特定用例中增强性能并降低功耗。您可以编辑这些配置集，或创建新配置集来创建适合您的环境的性能解决方案，包括虚拟环境。

要为虚拟化优化 RHEL 8，请使用以下配置集：

- 对于 RHEL 8 虚拟机，请使用 **virtual-guest** 配置集。它基于普遍适用的 **吞吐量性能** 配置文件，但也减少了虚拟内存的交换。
- 对于 RHEL 8 虚拟化主机，请使用 **virtual-host** 配置集。这可提高脏内存页面的主动回写，这有助于主机性能。

先决条件

- TuneD 服务** 已安装并启用。

流程

启用特定的 **TuneD** 配置集：

- 列出可用的 **TuneD** 配置集。

```
# tuned-adm list

Available profiles:
- balanced      - General non-specialized TuneD profile
- desktop       - Optimize for the desktop use-case
[...]
- virtual-guest - Optimize for running inside a virtual guest
- virtual-host  - Optimize for running KVM guests
Current active profile: balanced
```

- 可选：创建新 **TuneD** 配置集或编辑现有的 **TuneD** 配置集。

如需更多信息，请参阅[自定义 TuneD 配置集](#)。

- 激活 **TuneD** 配置集。

```
# tuned-adm profile selected-profile
```

- 要优化虚拟化主机，请使用 *virtual-host* 配置集。

```
# tuned-adm profile virtual-host
```

- 在 RHEL 虚拟机操作系统中，使用 *virtual-guest* 配置集。

```
# tuned-adm profile virtual-guest
```

其它资源

- [监控和管理系统状态和性能](#)

16.3. 配置虚拟机内存

要提高虚拟机(VM)的性能，您可以将额外的主机 RAM 分配给虚拟机。类似地，您可以减少分配给虚拟机的内存量，从而使主机内存可以分配给其他虚拟机或任务。

要执行这些操作，您可以使用 [Web 控制台](#) 或 [命令行界面](#)。

16.3.1. 使用 web 控制台添加和删除虚拟机内存

要提高虚拟机(VM)的性能或释放它使用的主机资源，您可以使用 Web 控制台来调整分配给虚拟机的内存量。

先决条件

- 客户端操作系统正在运行内存 balloon 驱动程序。请执行以下命令校验：

- 确保虚拟机的配置包含 **memballoon** 设备：

```
# virsh dumpxml testguest | grep membaloond
<memballoon model='virtio'>
</memballoon>
```

如果此命令显示任何输出，并且型号未设置为 **none**，则存在 **memballoon** 设备。

- 确保 balloon 驱动程序在客户机操作系统中运行。

- 在 Windows 客户机中，驱动程序作为 **virtio-win** 驱动程序软件包的一部分安装。具体步骤请查看 [为 Windows 虚拟机安装 KVM 半虚拟驱动程序](#)。
- 在 Linux 客户机中，通常默认包含驱动程序，并在存在 **memballoon** 设备时激活。

- Web 控制台 VM 插件 [已安装在您的系统上](#)。

流程

- 可选：包含有关虚拟机最大内存和当前使用的内存的信息。这将作为您更改的基准，并进行验证。

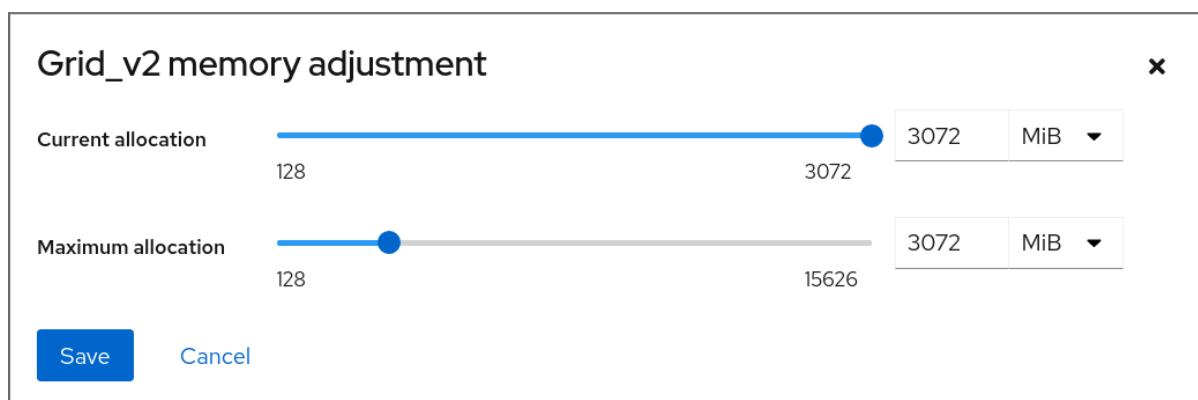
```
# virsh dominfo testguest
Max memory: 2097152 KiB
Used memory: 2097152 KiB
```

- 在 **Virtual Machines** 界面中，点击您要查看其信息的虚拟机。

此时将打开一个新页面，其中包含关于所选虚拟机基本信息的概述部分，以及用于访问虚拟机的图形界面的控制台部分。

- 单击概述窗格中 **Memory** 行旁边的 **edit**。

此时将显示 **Memory Adjustment** 对话框。



- 为所选虚拟机配置虚拟 CPU。

- 最大分配** - 设置虚拟机可用于其进程的最大主机内存量。您可以在创建虚拟机时指定最大内存，或可以在以后增大。您可以将内存指定为 MiB 或 GiB 的倍数。

只有在关闭虚拟机上才能调整最大内存分配。

- **当前分配** - 设置分配给虚拟机的实际内存量。这个值可以小于最大分配量，但不能超过它。您可以调整值，来控制虚拟机的进程可使用的内存。您可以将内存指定为 MiB 或 GiB 的倍数。
如果没有指定这个值，则默认分配是 **Maximum allocation** 值。

5. 点击 **Save**。

调整了虚拟机的内存分配。

其它资源

- [使用命令行界面添加和删除虚拟机内存](#)
- [优化虚拟机 CPU 性能](#)

16.3.2. 使用命令行界面添加和删除虚拟机内存

若要提高虚拟机(VM)的性能或释放其使用的主机资源，您可以使用 CLI 来调整分配给虚拟机的内存量。

先决条件

- 客户端操作系统正在运行内存 balloon 驱动程序。请执行以下命令校验：

1. 确保虚拟机的配置包含 **memballoon** 设备：

```
# virsh dumpxml testguest | grep membaloone
<memballoon model='virtio'>
</memballoon>
```

如果此命令显示任何输出，并且型号未设置为 **none**，则存在 **memballoon** 设备。

2. 确定 balloon 驱动程序正在客户端操作系统中运行。

- 在 Windows 客户机中，驱动程序作为 **virtio-win** 驱动程序软件包的一部分安装。具体步骤请查看 [为 Windows 虚拟机安装 KVM 半虚拟驱动程序](#)。
- 在 Linux 客户机中，通常默认包含驱动程序，并在存在 **memballoon** 设备时激活。

流程

1. 可选：包含有关虚拟机最大内存和当前使用的内存的信息。这将作为您更改的基准，并进行验证。

```
# virsh dominfo testguest
Max memory: 2097152 KiB
Used memory: 2097152 KiB
```

2. 调整分配给虚拟机的最大内存。增加这个值可以提高虚拟机的性能风险，降低这个值会降低虚拟机在主机上的性能占用空间。请注意，此更改只能在关闭的虚拟机上执行，因此调整正在运行的虚拟机需要重新启动才能生效。

例如，将 *testguest* 虚拟机可以使用的最大内存更改为 4096 MiB：

```
# virt-xml testguest --edit --memory memory=4096,currentMemory=4096
Domain 'testguest' defined successfully.
Changes will take effect after the domain is fully powered off.
```

要增加正在运行的虚拟机的最大内存，您可以将内存设备附加到虚拟机。这也被称为**内存热插拔**。详情请参阅[将设备附加到虚拟机](#)。



警告

不支持从正在运行的虚拟机中删除内存设备（也称为内存热拔），红帽不建议这样做。

- 可选：您还可以调整虚拟机当前使用的内存，最多不超过最大分配数。这调整了虚拟机在主机上的内存负载，直到下一次重启为止，而不需要更改最大的虚拟机分配。

```
# virsh setmem testguest --current 2048
```

验证

- 确认虚拟机使用的内存已更新：

```
# virsh dominfo testguest
Max memory: 4194304 KiB
Used memory: 2097152 KiB
```

- 可选：如果您调整了当前虚拟机内存，您可以获取虚拟机的内存 balloon 统计，以评估它如何有效地控制其内存使用量。

```
# virsh domstats --balloon testguest
Domain: 'testguest'
balloon.current=365624
balloon.maximum=4194304
balloon.swap_in=0
balloon.swap_out=0
balloon.major_fault=306
balloon.minor_fault=156117
balloon.unused=3834448
balloon.available=4035008
balloon.usable=3746340
balloon.last-update=1587971682
balloon.disk_caches=75444
balloon.hugetlb_pgalloc=0
balloon.hugetlb_pgfail=0
balloon.rss=1005456
```

其它资源

- [使用 web 控制台添加和删除虚拟机内存](#)

- 优化虚拟机 CPU 性能

16.3.3. 其它资源

- 将设备附加到虚拟机.

16.4. 优化虚拟机 I/O 性能

虚拟机(VM)的输入和输出(I/O)能力可能会显著限制虚拟机的整体效率。要解决这个问题，您可以通过配置块 I/O 参数来优化虚拟机的 I/O。

16.4.1. 在虚拟机中调整块 I/O

当一个或多个虚拟机正在使用多个块设备时,可能需要通过修改虚拟设备的 I/O 优先级来调整虚拟设备的 I/O 权重。

增加设备的 I/O 权重会增加设备的 I/O 带宽的优先级，从而为它提供更多主机资源。同样的，降低设备的权重可使其消耗较少的主机资源。



注意

每个设备的 **weight** 值必须在 **100** 到 **1000** 之间。或者，该值可以是 **0**，它会从每个设备列表中删除该设备。

流程

显示和设置虚拟机的块 I/O 参数：

1. 显示虚拟机当前的 **<blkio>** 参数：

```
# virsh dumpxml VM-name
```

```
<domain>
[...]
<blkiotune>
  <weight>800</weight>
  <device>
    <path>/dev/sda</path>
    <weight>1000</weight>
  </device>
  <device>
    <path>/dev/sdb</path>
    <weight>500</weight>
  </device>
</blkiotune>
[...]
</domain>
```

2. 编辑指定设备的 I/O 加权：

```
# virsh blkiotune VM-name --device-weights device, I/O-weight
```

例如：以下将 *liftrul* 虚拟机中的 */dev/sda* 设备的权重改为 500。

```
# virsh blkiotune liftbrul --device-weights /dev/sda, 500
```

16.4.2. 虚拟机中的磁盘 I/O 节流

当多个虚拟机同时运行时，它们可能会使用过量的磁盘 I/O 而干扰系统性能。KVM 虚拟化中的磁盘 I/O 节流使得能够对从虚拟机发送到主机的磁盘 I/O 请求设定限制。这可以防止虚拟机过度使用共享资源并影响其他虚拟机的性能。

要启用磁盘 I/O 节流，请对从附加到虚拟机的每个块设备发送给主机的磁盘 I/O 请求设置限制。

流程

1. 使用 **virsh domblklist** 命令列出指定虚拟机上所有磁盘设备的名称。

```
# virsh domblklist rollin-coal
Target   Source
-----
vda      /var/lib/libvirt/images/rollin-coal.qcow2
sda      -
sdb      /home/horridly-demanding-processes.iso
```

2. 找到您要节流的虚拟磁盘挂载的主机块设备。

例如，如果您想要从上一步中节流 **sdb** 虚拟磁盘，以下输出显示该磁盘挂载在 **/dev/nvme0n1p3** 分区上。

```
$ lsblk
NAME           MAJ:MIN RM  SIZE RO TYPE MOUNTPOINT
zram0          252:0   0    4G  0 disk [SWAP]
nvme0n1        259:0   0 238.5G 0 disk
└─nvme0n1p1    259:1   0  600M 0 part /boot/efi
└─nvme0n1p2    259:2   0   1G  0 part /boot
└─nvme0n1p3    259:3   0 236.9G 0 part
  └─luks-a1123911-6f37-463c-b4eb-fxzy1ac12fea 253:0   0 236.9G 0 crypt /home
```

3. 使用 **virsh blkiotune** 命令为块设备设置 I/O 限制。

```
# virsh blkiotune VM-name --parameter device,limit
```

以下示例将 **rollin-coal** 上的 **sdb** 磁盘节流为每秒 1000 个读写 I/O 操作，每秒的读写 I/O 操作吞吐量 50 MB。

```
# virsh blkiotune rollin-coal --device-read-iops-sec /dev/nvme0n1p3,1000 --device-write-iops-sec /dev/nvme0n1p3,1000 --device-write-bytes-sec /dev/nvme0n1p3,52428800 --device-read-bytes-sec /dev/nvme0n1p3,52428800
```

附加信息

- 磁盘 I/O 节流可用于各种情况，例如，当属于不同客户的虚拟机在同一台主机上运行时，或者为不同的虚拟机提供服务质量保障时。磁盘 I/O 节流还可用来模拟较慢的磁盘。
- I/O 节流可以独立应用于附加到虚拟机的每个块设备，并支持对吞吐量和 I/O 操作的限制。

- 红帽不支持使用 **virsh blkdeviotune** 命令在虚拟机上配置 I/O 节流。有关使用 RHEL 8 作为虚拟机主机时不支持的功能的更多信息，请参阅 [RHEL 8 虚拟化中的不支持的功能](#)。

16.4.3. 启用多队列 virtio-scsi

在虚拟机(VM)中使用 **virtio-scsi** 存储设备时，*multi-queue virtio-scsi* 特性可提高存储性能和可扩展性。它允许每个虚拟 CPU(vCPU)使用单独的队列和中断，而不影响其他 vCPU。

流程

- 要为特定虚拟机启用 multi-queue virtio-scsi 支持，请在虚拟机的 XML 配置中添加以下内容，其中 *N* 是 vCPU 队列的总数：

```
<controller type='scsi' index='0' model='virtio-scsi'>
  <driver queues='N' />
</controller>
```

16.5. 优化虚拟机 CPU 性能

与主机中的物理 CPU 非常相似，vCPU 对虚拟机(VM)性能至关重要。因此，优化 vCPU 会对虚拟机的资源效率产生重大影响。优化 vCPU：

1. 调整分配给虚拟机的主机 CPU 数。您可以使用 [CLI](#) 或 [Web 控制台](#) 进行此操作。
 2. 确保 vCPU 模型与主机的 CPU 型号一致。例如，将 *testguest1* 虚拟机设置为使用主机的 CPU 型号：
- ```
virt-xml testguest1 --edit --cpu host-model
```
3. [停止内核同页合并\(KSM\)](#)。
  4. 如果您的主机使用非统一内存访问(NUMA)，您也可以为其虚拟机 [配置 NUMA](#)。这会尽可能将主机的 CPU 和内存进程映射到虚拟机的 CPU 和内存进程上。实际上，NUMA 调优为 vCPU 提供了对分配给虚拟机的系统内存更精简的访问，这可以提高 vCPU 的处理效率。

详情请参阅 [在虚拟机中配置 NUMA 和 vCPU 性能调整场景示例](#)。

### 16.5.1. 使用命令行界面添加和删除虚拟 CPU

要提高或优化虚拟机(VM)的 CPU 性能，您可以添加或删除分配给虚拟机的虚拟 CPU(vCPU)。

当在运行的虚拟机上执行时，这也被称为 vCPU 热插和热拔。但请注意，RHEL 8 不支持 vCPU 热拔，红帽不建议使用它。

#### 先决条件

- 可选：查看目标虚拟机中的 vCPU 的当前状态。例如，显示 *testguest* 虚拟机上的 vCPU 数量：

```
virsh vcpucount testguest
maximum config 4
maximum live 2
current config 2
current live 1
```

此输出显示 *testguest* 目前使用 1 个 vCPU，另外 1 个 vCPU 可以热插入以提高虚拟机性能。但是，重新引导后，vCPU *testguest* 使用的数量会改为 2，而且能够热插 2 个 vCPU。

## 流程

- 调整可以附加到虚拟机的最大 vCPU 数量，其在虚拟机下次启动时生效。  
例如，要将 *testguest* 虚拟机的最大 vCPU 数量增加到 8：

```
virsh setvcpus testguest 8 --maximum --config
```

请注意，最大值可能会受 CPU 拓扑、主机硬件、hypervisor 和其他因素的限制。

- 将当前附加到虚拟机的 vCPU 数量调整到上一步中配置的最大值。例如：

- 将附加到正在运行的 *testguest* 虚拟机的 vCPU 数量增加到 4：

```
virsh setvcpus testguest 4 --live
```

这会增加虚拟机的性能和主机的 *testguest* 负载占用，直到虚拟机下次引导为止。

- 将附加到 *testguest* 虚拟机的 vCPU 数量永久减少至 1：

```
virsh setvcpus testguest 1 --config
```

这会降低虚拟机的性能和 *testguest* 的主机负载占用。但是，如果需要可热插入虚拟机以暂时提高性能。

## 验证

- 确认虚拟机的 vCPU 的当前状态反映了您的更改。

```
virsh vcpucount testguest
maximum config 8
maximum live 4
current config 1
current live 4
```

## 其它资源

- [使用 Web 控制台管理虚拟 CPU](#)

### 16.5.2. 使用 Web 控制台管理虚拟 CPU

使用 RHEL 8 web 控制台，您可以查看并配置 web 控制台连接的虚拟机使用的虚拟 CPU。

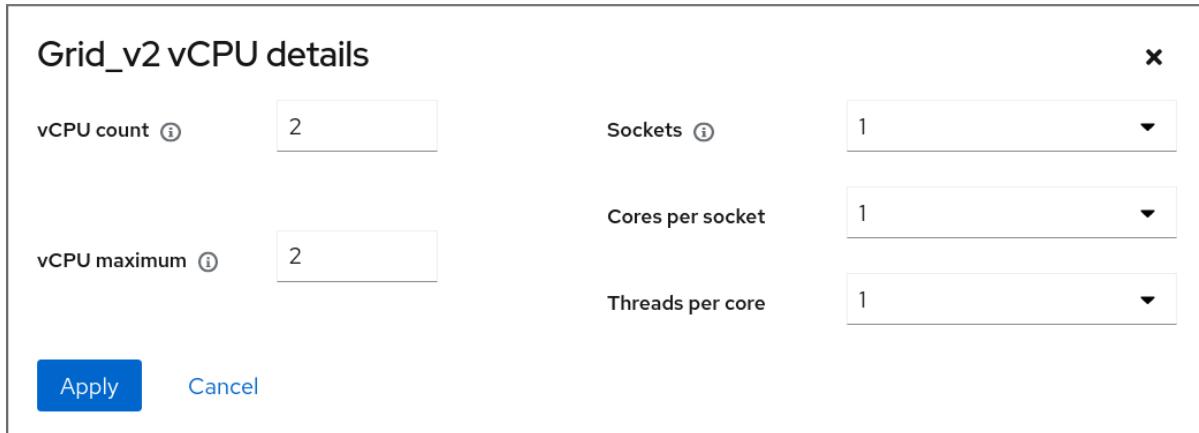
## 先决条件

- Web 控制台 VM 插件 [已安装在您的系统上](#)。

## 流程

- 在 **Virtual Machines** 界面中，点击您要查看其信息的虚拟机。  
此时将打开一个新页面，其中包含关于所选虚拟机基本信息的概述部分，以及用于访问虚拟机的图形界面的控制台部分。

- 单击概述窗格中 vCPU 数旁边的 **edit**。
- 此时会出现 vCPU 详情对话框。



- 为所选虚拟机配置虚拟 CPU。
  - vCPU 数量** - 当前正在使用的 vCPU 数量。



### 注意

vCPU 数量不能超过 vCPU 的最大值。

- vCPU 最大** - 可为虚拟机配置的最大虚拟 CPU 数。如果这个值大于 **vCPU Count**, 可以为虚拟机附加额外的 vCPU。
- 插槽** - 向虚拟机公开的插槽数量。
- 每个插槽的内核数** - 向虚拟机公开的每个插槽的内核数。
- 每个内核的线程数** - 向虚拟机公开的每个内核的线程数。

请注意，**插槽**、**每个插槽的内核数**和**每个内核的线程数**选项调整了虚拟机的 CPU 拓扑。这可能对 vCPU 性能有好处，但可能会影响客户机操作系统中某些软件的功能。如果您的部署不需要不同的设置，请保留默认值。

- 点应用。
- 配置了虚拟机的虚拟 CPU。



### 注意

对虚拟 CPU 设置的更改仅在重启虚拟机后生效。

## 其它资源

- [使用命令行界面添加和删除虚拟 CPU](#)

### 16.5.3. 在虚拟机中配置 NUMA

以下方法可用于配置 RHEL 8 主机上虚拟机(VM)的非统一内存访问(NUMA)设置。

## 先决条件

- 主机是一个与 NUMA 兼容的机器。要检测是否是这种情况，请使用 **virsh nodeinfo** 命令，并查看 **NUMA cell(s)** 行：

```
virsh nodeinfo
CPU model: x86_64
CPU(s): 48
CPU frequency: 1200 MHz
CPU socket(s): 1
Core(s) per socket: 12
Thread(s) per core: 2
NUMA cell(s): 2
Memory size: 67012964 KiB
```

如果行的值为 2 或更高，则主机与 NUMA 兼容。

## 流程

为便于使用，您可以使用自动化工具和服务设置虚拟机的 NUMA 配置。但是，手动 NUMA 设置可能会显著提高性能。

### 自动方法

- 将虚拟机的 NUMA 策略设为 **Preferred**。例如，对于 *testguest5* 虚拟机要这样做：

```
virt-xml testguest5 --edit --vcpus placement=auto
virt-xml testguest5 --edit --numatune mode=preferred
```

- 在主机上启用自动 NUMA 均衡：

```
echo 1 > /proc/sys/kernel/numa_balancing
```

- 使用 **numad** 命令，自动将 VM CPU 与内存资源匹配。

```
numad
```

### 手动方法

1. 将特定 vCPU 线程固定到特定主机 CPU 或者 CPU 范围。在非 NUMA 主机和虚拟机上也可以这样做，我们推荐您使用一种安全的方法来提高 vCPU 性能。

例如，以下命令将 *testguest6* 虚拟机的 vCPU 线程 0 到 5 分别固定到主机 CPU 1、3、5、7、9 和 11：

```
virsh vcpupin testguest6 0 1
virsh vcpupin testguest6 1 3
virsh vcpupin testguest6 2 5
virsh vcpupin testguest6 3 7
virsh vcpupin testguest6 4 9
virsh vcpupin testguest6 5 11
```

之后，您可以验证操作是否成功：

```
virsh vcpupin testguest6
VCPU CPU Affinity

0 1
```

```

1 3
2 5
3 7
4 9
5 11

```

2. 固定 vCPU 线程后，您还可以将与指定虚拟机关联的 QEMU 进程线程固定到特定的主机 CPU 或 CPU 范围。例如：以下命令将 *testguest6* 的 QEMU 进程线程 固定到 CPU 13 和 15，确认成功：

```

virsh emulatorpin testguest6 13,15
virsh emulatorpin testguest6
emulator: CPU Affinity

*: 13,15

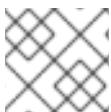
```

3. 最后，您也可以指定将哪些主机 NUMA 节点专门分配给某个虚拟机。这可提高虚拟机 vCPU 的主机内存用量。例如，以下命令将 *testguest6* 设置为使用主机 NUMA 节点 3 到 5，确认成功：

```

virsh numatune testguest6 --nodeset 3-5
virsh numatune testguest6

```



### 注意

为了获得最佳性能，建议使用以上列出的所有手动调优方法

## 已知问题

- 目前无法在 IBM Z 主机上执行 NUMA 调整

## 其它资源

- [vCPU 性能调整场景示例](#)
- 使用 **numastat** 工具 [查看系统的当前 NUMA 配置](#)

### 16.5.4. vCPU 性能调整场景示例

为了获得最佳 vCPU 性能，红帽建议一起使用手动 **vcpupin**、**emulatorpin** 和 **numatune** 设置，例如在以下场景中。

## 起始场景

- 您的主机有以下与硬件相关的信息：
  - 2 个 NUMA 节点
  - 每个节点上的 3 个 CPU 内核
  - 每个内核有 2 个线程

此类机器的 **virsh nodeinfo** 输出类似于：

```

virsh nodeinfo
CPU model: x86_64

```

```
CPU(s): 12
CPU frequency: 3661 MHz
CPU socket(s): 2
Core(s) per socket: 3
Thread(s) per core: 2
NUMA cell(s): 2
Memory size: 31248692 KiB
```

- 您打算将现有的虚拟机修改为有 8 个 vCPU，这意味着单个 NUMA 节点无法容纳它。因此，您应该在每个 NUMA 节点上分发 4 个 vCPU，并使 vCPU 拓扑尽可能接近主机拓扑。这意味着，作为给定物理 CPU 的同级线程运行的 vCPU 应该固定到同一核上的主机线程。详情请查看以下解决方案。

## 解决方案

- 获取主机拓扑的信息：

```
virsh capabilities
```

输出应包含类似如下的部分：

```
<topology>
<cells num="2">
 <cell id="0">
 <memory unit="KiB">15624346</memory>
 <pages unit="KiB" size="4">3906086</pages>
 <pages unit="KiB" size="2048">0</pages>
 <pages unit="KiB" size="1048576">0</pages>
 <distances>
 <sibling id="0" value="10" />
 <sibling id="1" value="21" />
 </distances>
 <cpus num="6">
 <cpu id="0" socket_id="0" core_id="0" siblings="0,3" />
 <cpu id="1" socket_id="0" core_id="1" siblings="1,4" />
 <cpu id="2" socket_id="0" core_id="2" siblings="2,5" />
 <cpu id="3" socket_id="0" core_id="0" siblings="0,3" />
 <cpu id="4" socket_id="0" core_id="1" siblings="1,4" />
 <cpu id="5" socket_id="0" core_id="2" siblings="2,5" />
 </cpus>
 </cell>
 <cell id="1">
 <memory unit="KiB">15624346</memory>
 <pages unit="KiB" size="4">3906086</pages>
 <pages unit="KiB" size="2048">0</pages>
 <pages unit="KiB" size="1048576">0</pages>
 <distances>
 <sibling id="0" value="21" />
 <sibling id="1" value="10" />
 </distances>
 <cpus num="6">
 <cpu id="6" socket_id="1" core_id="3" siblings="6,9" />
 <cpu id="7" socket_id="1" core_id="4" siblings="7,10" />
 <cpu id="8" socket_id="1" core_id="5" siblings="8,11" />
 <cpu id="9" socket_id="1" core_id="3" siblings="6,9" />
```

```

<cpu id="10" socket_id="1" core_id="4" siblings="7,10" />
<cpu id="11" socket_id="1" core_id="5" siblings="8,11" />
</cpus>
</cell>
</cells>
</topology>
```

2. 可选：使用 [适用的工具和实用程序](#) 测试虚拟机的性能。

3. 在主机上设置并挂载 1GiB 巨页：

a. 在主机的内核命令行中添加以下行：

```
default_hugepagesz=1G hugepagesz=1G
```

b. 使用以下内容创建 **/etc/systemd/system/hugetlb-gigantic-pages.service** 文件：

```

[Unit]
Description=HugeTLB Gigantic Pages Reservation
DefaultDependencies=no
Before=dev-hugepages.mount
ConditionPathExists=/sys/devices/system/node
ConditionKernelCommandLine=hugepagesz=1G

[Service]
Type=oneshot
RemainAfterExit=yes
ExecStart=/etc/systemd/hugetlb-reserve-pages.sh

[Install]
WantedBy=sysinit.target
```

c. 使用以下内容创建 **/etc/systemd/hugetlb-reserve-pages.sh** 文件：

```

#!/bin/sh

nodes_path=/sys/devices/system/node/
if [! -d $nodes_path]; then
 echo "ERROR: $nodes_path does not exist"
 exit 1
fi

reserve_pages()
{
 echo $1 > $nodes_path/$2/hugepages/hugepages-1048576kB/nr_hugepages
}

reserve_pages 4 node1
reserve_pages 4 node2
```

这会从 *node1* 保留 4 个 1GiB 巨页，并在 *node2* 中保留 4 个 1GiB 巨页。

d. 使在上一步中创建的脚本可执行：

```
chmod +x /etc/systemd/hugetlb-reserve-pages.sh
```

- e. 在引导时启用巨页保留：

```
systemctl enable hugetlb-gigantic-pages
```

4. 使用 **virsh edit** 命令编辑您要优化的虚拟机的 XML 配置，在本例中为 *super-VM*：

```
virsh edit super-vm
```

5. 用以下方法调整虚拟机的 XML 配置：

- 将虚拟机设置为使用 8 个静态 vCPU。使用 **<vcpu/>** 元素来执行此操作。
- 将每个 vCPU 线程固定到拓扑中镜像的对应主机 CPU 线程。为此，请在 **<cputune>** 部分中使用 **<vcputpin/>** 元素。

请注意，如上面 **virsh capabilities** 工具所示，主机 CPU 线程在各自的内核中不是按顺序排序的。此外，vCPU 线程应固定到同一 NUMA 节点上最多可用的主机核集合。有关图表，请查看以下 **示例拓扑** 部分。

步骤 a. 和 b. 的 XML 配置类似：

```
<cputune>
<vcputpin vcpu='0' cpuset='1' />
<vcputpin vcpu='1' cpuset='4' />
<vcputpin vcpu='2' cpuset='2' />
<vcputpin vcpu='3' cpuset='5' />
<vcputpin vcpu='4' cpuset='7' />
<vcputpin vcpu='5' cpuset='10' />
<vcputpin vcpu='6' cpuset='8' />
<vcputpin vcpu='7' cpuset='11' />
<emulatorpin cpuset='6,9' />
</cputune>
```

- c. 将虚拟机设置为使用 1 GiB 巨页：

```
<memoryBacking>
<hugepages>
<page size='1' unit='GiB' />
</hugepages>
</memoryBacking>
```

- d. 配置虚拟机的 NUMA 节点，使其使用主机上对应的 NUMA 节点的内存。要做到这一点，请在 **<numatune>** 部分中使用 **<memnode/>** 元素：

```
<numatune>
<memory mode="preferred" nodeset="1" />
<memnode cellid="0" mode="strict" nodeset="0" />
<memnode cellid="1" mode="strict" nodeset="1" />
</numatune>
```

- e. 确保 CPU 模式设为 **host-passthrough**，且 CPU 在 **passthrough** 模式下使用缓存：

```
<cpu mode="host-passthrough">
<topology sockets="2" cores="2" threads="2" />
<cache mode="passthrough" />
```

## 验证

- 确认生成的虚拟机 XML 配置包含类似如下内容：

```
[...]
<memoryBacking>
 <hugepages>
 <page size='1' unit='GiB'/>
 </hugepages>
</memoryBacking>
<vcpu placement='static'>8</vcpu>
<cputune>
 <vcpuin vcpu='0' cpuset='1' />
 <vcpuin vcpu='1' cpuset='4' />
 <vcpuin vcpu='2' cpuset='2' />
 <vcpuin vcpu='3' cpuset='5' />
 <vcpuin vcpu='4' cpuset='7' />
 <vcpuin vcpu='5' cpuset='10' />
 <vcpuin vcpu='6' cpuset='8' />
 <vcpuin vcpu='7' cpuset='11' />
 <emulatorpin cpuset='6,9' />
</cputune>
<numatune>
 <memory mode="preferred" nodeset="1" />
 <memnode cellid="0" mode="strict" nodeset="0" />
 <memnode cellid="1" mode="strict" nodeset="1" />
</numatune>
<cpu mode="host-passthrough">
 <topology sockets="2" cores="2" threads="2" />
 <cache mode="passthrough" />
<numa>
 <cell id="0" cpus="0-3" memory="2" unit="GiB">
 <distances>
 <sibling id="0" value="10" />
 <sibling id="1" value="21" />
 </distances>
 </cell>
 <cell id="1" cpus="4-7" memory="2" unit="GiB">
 <distances>
 <sibling id="0" value="21" />
 <sibling id="1" value="10" />
 </distances>
 </cell>
</numa>
</cpu>
</domain>
```

- 可选：使用 [适用的工具和实用程序](#) 测试虚拟机的性能来评估虚拟机优化的影响。

## 拓扑示例

- 下表演示了 vCPU 和主机 CPU 之间的连接：

表 16.1. 主机拓扑

CPU 线程	0	3	1	4	2	5	6	9	7	10	8	11
内核	0		1		2		3		4		5	
插槽			0						1			
NUMA 节点			0						1			

表 16.2. VM 拓扑

vCPU 线程	0	1	2	3	4	5	6	7
内核	0		1		2		3	
插槽			0			1		
NUMA 节点			0			1		

表 16.3. 合并主机和虚拟机拓扑

vCPU 线程	0	1	2	3	4	5	6	7				
主机 CPU 线程	0	3	1	4	2	5	6	9	7	10	8	11
内核	0		1		2		3		4		5	
插槽			0						1			
NUMA 节点			0						1			

在这种情况下，有 2 个 NUMA 节点和 8 个 vCPU。因此，应该为每个节点固定 4 个 vCPU 线程。

另外，红帽建议在每个节点上至少保留一个 CPU 线程用于主机系统操作。

因为在这个示例中，每个 NUMA 节点都有 3 个核，每个核都有 2 个主机 CPU 线程，节点 0 的设置转换如下：

```
<vcpuin vcpu='0' cpuset='1'/>
<vcpuin vcpu='1' cpuset='4'/>
<vcpuin vcpu='2' cpuset='2'/>
<vcpuin vcpu='3' cpuset='5'/>
```

## 16.5.5. 停止内核相同页面合并

虽然内核相同页面合并(KSM)提高了内存密度，但它会增加 CPU 的使用率，并且可能会对总体性能产生不利影响，具体取决于工作负载。在这种情况下，您可以通过停用 KSM 来提高虚拟机(VM)的性能。

根据您的要求，您可以对单个会话停用 KSM，或永久停用。

## 流程

- 要对单个会话停用 KSM， 请使用 **systemctl** 工具停止 **ksm** 和 **ksmtuned** 服务。

```
systemctl stop ksm
systemctl stop ksmtuned
```

- 要永久停用 KSM， 请使用 **systemctl** 工具来禁用 **ksm** 和 **ksmtuned** 服务。

```
systemctl disable ksm
Removed /etc/systemd/system/multi-user.target.wants/ksm.service.
systemctl disable ksmtuned
Removed /etc/systemd/system/multi-user.target.wants/ksmtuned.service.
```



### 注意

取消激活 KSM 前在虚拟机间共享的内存页将保持共享。要停止共享，请使用以下命令删除系统中的所有 **PageKSM** 页面：

```
echo 2 > /sys/kernel/mm/ksm/run
```

匿名页面替换了 KSM 页面后，**khugepaged** 内核服务将在虚拟机物理内存中重建透明的大页面。

## 16.6. 优化虚拟机网络性能

由于虚拟机的网络接口卡(NIC)的虚拟性质，虚拟机会丢失其分配的主机网络带宽的一部分，这可以降低虚拟机的整体工作负载效率。以下提示可最大程度降低虚拟化对虚拟 NIC(vNIC)吞吐量的负面影响。

## 流程

使用以下任一方法并观察它是否对虚拟机网络性能有帮助：

### 启用 **vhost\_net** 模块

在主机上，确保启用了 **vhost\_net** 内核特性：

```
lsmod | grep vhost
vhost_net 32768 1
vhost 53248 1 vhost_net
tap 24576 1 vhost_net
tun 57344 6 vhost_net
```

如果这个命令的输出为空，请启用 **vhost\_net** 内核模块：

```
modprobe vhost_net
```

### 设置 **multi-queue virtio-net**

要为虚拟机设置 *multi-queue virtio-net* 特性，请使用 **virsh edit** 命令来编辑虚拟机的 XML 配置。在 XML 中，将以下内容添加到 **<devices>** 部分，并将 **N** 替换为虚拟机中的 vCPU 个数，最多 16 个：

```
<interface type='network'>
 <source network='default' />
```

```
<model type='virtio'/>
<driver name='vhost' queues='N' />
</interface>
```

如果虚拟机正在运行，重启它以使更改生效。

## 批量网络数据包

在具有长传输路径的 Linux VM 配置中，在将数据包提交给内核之前对其进行批处理可以提高缓存利用率。要设置数据包批处理，请在主机上使用以下命令，并将 *tap0* 替换为虚拟机使用的网络接口的名称：

```
ethtool -C tap0 rx-frames 64
```

## SR-IOV

如果您的主机 NIC 支持 SR-IOV，请为您的 vNIC 使用 SR-IOV 设备分配。如需更多信息，请参阅[管理 SR-IOV 设备](#)。

## 其它资源

- [了解虚拟网络](#)

## 16.7. 虚拟机性能监控工具

要确定什么消耗了最多的 VM 资源，以及虚拟机性能的哪方面需要优化，可以使用性能诊断工具，包括通用的和特定于虚拟机的。

### 默认操作系统性能监控工具

要进行标准性能评估，您可以使用主机和客户机操作系统默认提供的工具：

- 在 RHEL 8 主机上，以 root 用户身份使用 **top** 工具或 **系统监控** 应用程序，并在输出中查找 **qemu** 和 **virt**。这显示了您的虚拟机消耗的主机系统资源量。
  - 如果监控工具显示任何 **qemu** 或 **virt** 进程消耗了大量的主机 CPU 或内存量，请使用 **perf** 工具进行调查。详情请查看以下信息。
  - 另外，如果 **vhost\_net** 线程进程（如 *vhost\_net-1234*）被显示为消耗大量主机 CPU 容量，请考虑使用[虚拟网络优化功能](#)，如 **multi-queue virtio-net**。
- 在客户机操作系统上，使用系统上可用的性能工具和应用程序来评估哪些进程消耗了最多的系统资源。
  - 在 Linux 系统上，您可以使用 **top** 工具。
  - 在 Windows 系统中，您可以使用 **Task Manager** 应用程序。

### perf kvm

您可以使用 **perf** 实用程序收集有关 RHEL 8 主机性能的特定虚拟化统计。要做到这一点：

1. 在主机上安装 **perf** 软件包：

```
yum install perf
```

2. 使用 **perf kvm stat** 命令之一来显示您的虚拟化主机的 perf 统计信息：

- 若要实时监控 hypervisor，请使用 **perf kvm stat live** 命令。
- 若要记录一段时间内 hypervisor 的 perf 数据，请使用 **perf kvm stat record** 命令激活日志记录。命令被取消或中断后，数据保存在 **perf.data.guest** 文件中，可以使用 **perf kvm stat report** 命令进行分析。

3. 分析 **VM-EXIT** 事件类型及其分发的 **perf** 输出。例如，**PAUSE\_INSTRUCTION** 事件应当不常发生，但在以下输出中，此事件的频繁发生表明主机 CPU 没有很好地处理正在运行的 vCPU。在这种场景下，请考虑关闭某些处于活动状态的虚拟机，从这些虚拟机中删除 vCPU，或者 [调优 vCPU 的性能](#)。

```
perf kvm stat report
```

Analyze events for all VMs, all VCPUs:

VM-EXIT	Samples	Samples%	Time%	Min Time	Max Time	Avg time
EXTERNAL_INTERRUPT	365634	31.59%	18.04%	0.42us	58780.59us 204.08us ( +/- 0.99% )	
MSR_WRITE	293428	25.35%	0.13%	0.59us	17873.02us	1.80us ( +/- 4.63% )
PREEMPTION_TIMER	276162	23.86%	0.23%	0.51us	21396.03us	3.38us ( +/- 5.19% )
PAUSE_INSTRUCTION	189375	16.36%	11.75%	0.72us	29655.25us	256.77us ( +/- 0.70% )
HLT	20440	1.77%	69.83%	0.62us	79319.41us	14134.56us ( +/- 0.79% )
VMCALL	12426	1.07%	0.03%	1.02us	5416.25us	8.77us ( +/- 7.36% )
EXCEPTION_NMI	27	0.00%	0.00%	0.69us	1.34us	0.98us ( +/- 3.50% )
EPT_MISCONFIG	5	0.00%	0.00%	5.15us	10.85us	7.88us ( +/- 11.67% )

Total Samples:1157497, Total events handled time:413728274.66us.

**perf kvm stat** 输出中表明有问题的其它事件类型包括：

- INSN\_EMULATION** - 建议次优化的 [虚拟机 I/O 配置](#)。

有关使用 **perf** 来监控虚拟化性能的更多信息，请参阅 **perf-kvm** 手册页。

## numastat

要查看系统当前的 NUMA 配置，您可以使用 **numastat** 工具，该工具通过安装 **numactl** 软件包来提供。

以下显示了一个有 4 个运行虚拟机的主机，各自从多个 NUMA 节点获取内存。这不是 vCPU 性能的最佳方案，并[保证调整](#)：

```
numastat -c qemu-kvm
```

Per-node process memory usage (in MBs)

PID	Node 0	Node 1	Node 2	Node 3	Node 4	Node 5	Node 6	Node 7	Total
-----	--------	--------	--------	--------	--------	--------	--------	--------	-------

```
51722 (qemu-kvm) 68 16 357 6936 2 3 147 598 8128
51747 (qemu-kvm) 245 11 5 18 5172 2532 1 92 8076
53736 (qemu-kvm) 62 432 1661 506 4851 136 22 445 8116
53773 (qemu-kvm) 1393 3 1 2 12 0 0 6702 8114

Total 1769 463 2024 7462 10037 2672 169 7837 32434
```

相反，以下显示单个节点为每个虚拟机提供内存，这效率显著提高。

```
numastat -c qemu-kvm

Per-node process memory usage (in MBs)
PID Node 0 Node 1 Node 2 Node 3 Node 4 Node 5 Node 6 Node 7 Total

51747 (qemu-kvm) 0 0 7 0 8072 0 1 0 8080
53736 (qemu-kvm) 0 0 7 0 0 0 8113 0 8120
53773 (qemu-kvm) 0 0 7 0 0 0 1 8110 8118
59065 (qemu-kvm) 0 0 8050 0 0 0 0 0 8051

Total 0 0 8072 0 8072 0 8114 8110 32368
```

## 16.8. 其它资源

- [优化 Windows 虚拟机](#)

# 第 17 章 安装和管理 WINDOWS 虚拟机

要将 Microsoft Windows 用作 RHEL 8 主机上虚拟机(VM)的客户机操作系统，红帽建议采取额外步骤以确保这些虚拟机正常运行。

为此，以下章节提供了有关安装和优化主机上 Windows 虚拟机的信息，以及在这些虚拟机中安装和配置驱动程序的信息。

## 17.1. 安装 WINDOWS 虚拟机

您可以在 RHEL 8 主机上创建完全虚拟化的 Windows 机器，在虚拟机(VM)中启动图形 Windows 安装程序，并优化安装的 Windows 客户机操作系统(OS)。

要创建虚拟机并安装 Windows 客户机操作系统，请使用 **virt-install** 命令或 RHEL 8 web 控制台。

### 先决条件

- Windows OS 安装源可以是以下之一，并在本地或网络上可用：
  - 安装介质的 ISO 镜像
  - 现有虚拟机安装的磁盘镜像
- 带有 KVM **virtio** 驱动程序的存储介质。  
要创建此介质，请参阅 [在主机中准备 virtio 驱动程序安装介质](#)。
- 如果要安装 Windows 11，必须在主机上安装 **edk2-ovmf**、**swtpm** 和 **libtpms** 软件包。

### 流程

1. 创建虚拟机。具体步骤请参阅 [创建虚拟机](#)，但请注意以下具体信息。

- 如果使用 **virt-install** 工具来创建虚拟机，请在命令中添加以下选项：

- 带有 KVM **virtio** 驱动程序的存储介质.例如：

```
--disk path=/usr/share/virtio-win/virtio-win.iso,device=cdrom
```

- 要安装的 Windows 版本。例如，对于 Windows 10 和 11：

```
--os-variant win10
```

要获得可用 Windows 版本列表以及相应的选项，请使用以下命令：

```
osinfo-query os
```

- 如果您要安装 Windows 11，启用 **统一可扩展固件接口 (UEFI)**和 **虚拟信任平台模块 (vTPM)**：

```
--boot uefi --tpm model=tpm-crb,backend.type=emulator,backend.version=2.0
```

- 如果使用 web 控制台创建虚拟机，在 **Create new virtual machine** 窗口的 **Operating system** 字段中指定您的 Windows 版本。

- 如果要在 Windows 11 和 Windows Server 2022 之前安装 Windows 版本，点 **Create 开始安装并运行**。
- 如果您要安装 Windows 11，或者您想要使用额外的 Windows Server 2022 功能，点 **Create and edit and enable UEFI 和 vTPM 使用 CLI 确认**：
  - A. 打开虚拟机的 XML 配置：

```
virsh edit windows-vm
```

- B. 在 **os** 元素中添加 **firmware='efi'** 选项：

```
<os firmware='efi'>
<type arch='x86_64' machine='pc-q35-6.2'>hvm</type>
<boot dev='hd'/>
</os>
```

- C. 在 **devices** 元素中添加 **tpm** 设备：

```
<devices>
<tpm model='tpm-crb'>
 <backend type='emulator' version='2.0' />
</tpm>
</devices>
```

- D. 点击虚拟机表中的 **Install** 启动 Windows 安装。

## 2. 在虚拟机中安装 Windows OS。

有关如何安装 Windows 操作系统的详情，请参考相关微软安装文档。

3. 如果使用 Web 控制台创建虚拟机，请使用 Disks 接口将带有 virtio 驱动程序的存储介质附加到虚拟机。具体步骤请查看 [使用 web 控制台将现有磁盘附加到虚拟机](#)。
4. 在 Windows 客户机操作系统中配置 KVM virtio 驱动程序。详情请查看 [为 Windows 虚拟机安装 KVM 半虚拟驱动程序](#)。

## 其它资源

- [优化 Windows 虚拟机](#)
- [在 Windows 虚拟机中启用标准硬件安全性](#)
- [在 Windows 虚拟机上启用增强的硬件安全](#)
- [虚拟机 XML 配置示例](#)

## 17.2. 优化 WINDOWS 虚拟机

当在 RHEL 8 上托管的虚拟机(VM)中使用 Microsoft Windows 作为客户机操作系统时，客户端的性能可能会受到负面影响。

因此，红帽建议您使用以下组合来优化 Windows 虚拟机：

- 使用半虚拟驱动程序。如需更多信息，请参阅[为 Windows 虚拟机安装 KVM 半虚拟驱动程序](#)。

- 启用 Hyper-V enlightenments。如需更多信息, [请参阅启用 Hyper-Vlightenments。](#)
- 配置 NetKVM 驱动程序参数。如需更多信息, [请参阅配置 NetKVM 驱动程序参数。](#)
- 优化或禁用 Windows 后台进程。如需更多信息, [请参阅 优化 Windows 虚拟机上的后台进程。](#)

### 17.2.1. 为 Windows 虚拟机安装 KVM 半虚拟驱动程序

提高 Windows 虚拟机(VM)性能的主要方法是在客户机操作系统上为 Windows 安装 KVM 半虚拟化(virtio)驱动程序。

要在 Windows 虚拟机上安装驱动程序, 请执行以下操作 :

- 准备主机机器上的安装介质。如需更多信息, [请参阅 在主机 中准备 virtio 驱动程序安装介质。](#)
- 将安装介质附加到现有 Windows 虚拟机, 或者在创建新 Windows 虚拟机时附加该介质。如需更多信息, [请参阅在 RHEL 上安装 Windows 虚拟机。](#)
- 在 Windows 客户机操作系统中安装 virtio 驱动程序。如需更多信息, [请参阅在 Windows 客户端 中安装 virtio 驱动程序。](#)
- 在 Windows 客户机操作系统上安装 QEMU 客户机代理。如需更多信息, [请参阅在 Windows 客户机上安装 QEMU 客户机代理。](#)

#### 17.2.1.1. Windows virtio 驱动程序如何工作

半虚拟化驱动程序通过降低 I/O 延迟并将吞吐量提高到几乎裸机水平来提高虚拟机(VM)的性能。红帽建议您为运行 I/O 密集型任务和应用程序的虚拟机使用半虚拟化驱动程序。

**virtio** 驱动程序是 KVM 的半虚拟化设备驱动程序, 可用于在 KVM 主机上运行的 Windows 虚拟机。这些驱动程序由 **virtio-win** 软件包提供, 其中包括以下设备的驱动程序 :

- 块 (存储) 设备
- 网络接口控制器
- 视频控制器
- 内存 ballooning 设备
- 半虚拟串口设备
- 熵源设备
- 半虚拟 panic 设备
- 输入设备, 如鼠标、键盘或平板电脑
- 一组小型模拟设备



#### 注意

有关模拟、virtio 和分配的设备的附加信息, [请参阅管理虚拟设备。](#)

使用 KVM virtio 驱动程序, 以下 Microsoft Windows 版本应该与物理系统运行类似 :

- Windows Server 版本：请参阅红帽知识库中的带有 KVM 的 Red Hat Enterprise Linux 认证的客户机操作系统。
- Windows Desktop（非服务器）版本：
  - Windows 10 (32 位和 64 位版本)

### 17.2.1.2. 在主机中准备 virtio 驱动程序安装介质

要在 Windows 虚拟机(VM)上安装或更新 KVM virtio 驱动程序，您必须首先在主机上准备 virtio 驱动程序安装介质。为此，请将 virtio-win 软件包提供的 .iso 文件作为存储设备附加到 Windows 虚拟机。

#### 先决条件

- 确定在 RHEL 8 主机系统中启用了虚拟化。如需更多信息，请参阅[启用虚拟化](#)。
- 确保具有虚拟机的 root 访问权限。

#### 流程

1. 刷新您的订阅数据：

```
subscription-manager refresh
All local data refreshed
```

2. 获取 virtio-win 软件包的最新版本。

- 如果没有安装 virtio-win：

```
yum install -y virtio-win
```

- 如果安装了 virtio-win：

```
yum upgrade -y virtio-win
```

如果安装成功，则 virtio-win 驱动程序文件位于 /usr/share/virtio-win/ 目录中。其中包括 ISO 文件和一个 drivers 目录，目录中包含驱动程序文件，每个架构及受支持的 Windows 版本各对应一个文件。

```
ls /usr/share/virtio-win/
drivers/ guest-agent/ virtio-win-1.9.9.iso virtio-win.iso
```

3. 将 virtio-win.iso 文件作为存储设备附加到 Windows 虚拟机。

- [创建新 Windows 虚拟机时](#)，请使用 virt-install 命令选项附加该文件。
- 在现有 Windows 虚拟机上安装驱动程序时，请使用 virt-xml 工具将文件作为 CD-ROM 附加：

```
virt-xml WindowsVM --add-device --disk virtio-win.iso,device=cdrom
Domain 'WindowsVM' defined successfully.
```

#### 其它资源

- 在 Windows 客户机操作系统中安装 virtio 驱动程序。

### 17.2.1.3. 在 Windows 客户端中安装 virtio 驱动程序

要在 Windows 客户机操作系统上安装 KVM virtio 驱动程序，您必须添加一个包含驱动程序（在创建虚拟机(VM)或之后）的存储设备，并在 Windows 客户机操作系统中安装驱动程序。

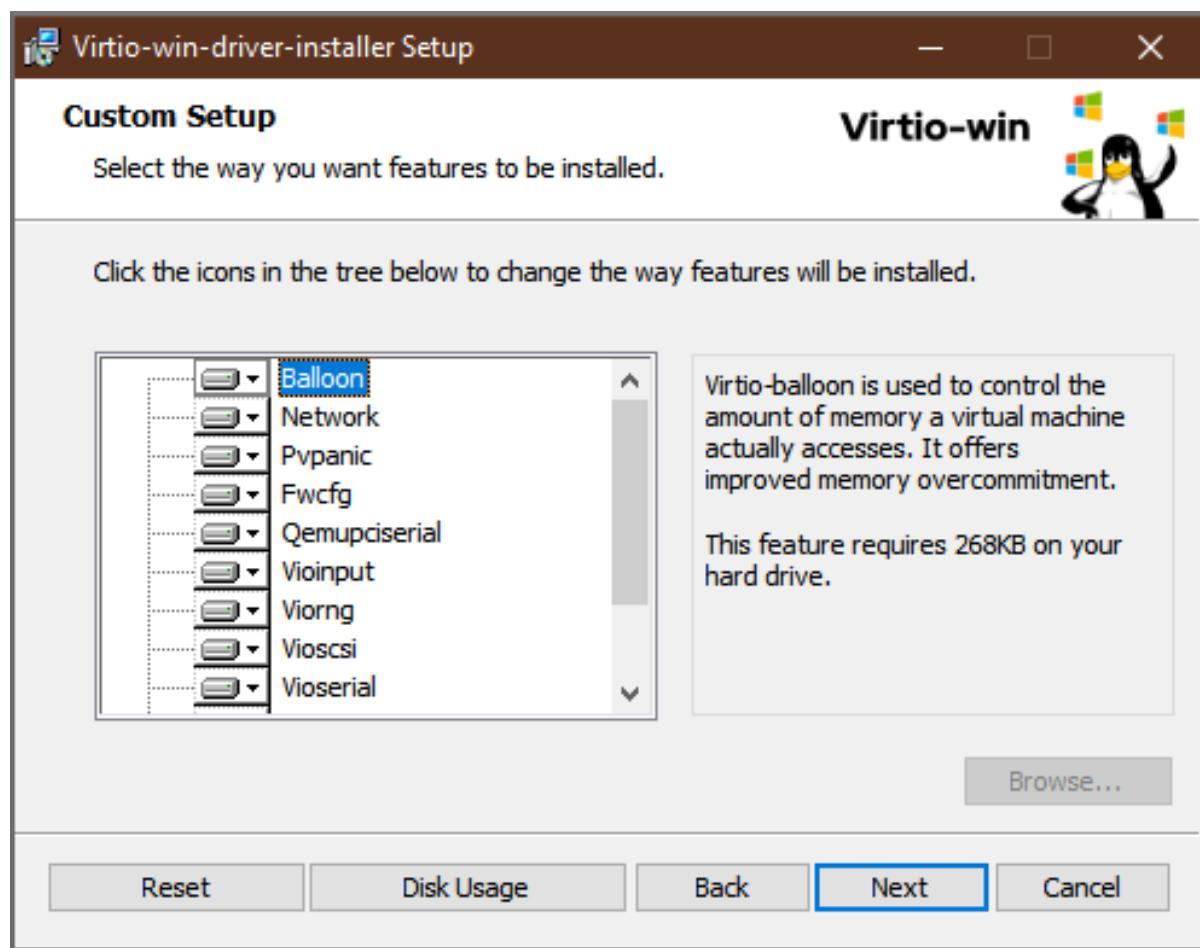
此流程提供了使用图形界面安装驱动程序的说明。您也可以使用 [Microsoft Windows Installer\(MSI\)](#) 命令行界面。

#### 先决条件

- 带有 KVM virtio 驱动程序的安装介质必须附加到虚拟机。有关准备该介质的步骤，请参考 [在主机中准备 virtio 驱动程序安装介质](#)。

#### 流程

1. 在 Windows 客户机操作系统中，打开 **File Explorer** 应用程序。
2. 单击 **This PC**。
3. 在 **Devices and drives** 窗格中，打开 **virtio-win** 介质。
4. 根据在虚拟机上安装的操作系统，运行安装程序之一：
  - 如果使用 32 位操作系统，请运行 **virtio-win-gt-x86.msi** 安装程序。
  - 如果使用 64 位操作系统，请运行 **virtio-win-gt-x64.msi** 安装程序。
5. 在打开的 **Virtio-win-driver-installer** 设置向导中，请按照显示的说明进行操作，直到您到达 **Custom Setup** 步骤。



6. 在 Custom Setup 窗口中，选择您要安装的设备驱动程序。会自动选择建议的驱动程序集，驱动程序的描述显示在列表右侧。
7. 单击 next，然后单击 Install。
8. 安装完成后，点完成。
9. 重启虚拟机以完成驱动程序安装。

### 验证

1. 在 Windows 虚拟机上，进入设备管理器：
  - a. 点 Start
  - b. 搜索设备管理器
2. 确保设备使用正确的驱动程序：
  - a. 点设备打开 Driver Properties 窗口。
  - b. 进入 Driver 选项卡。
  - c. 单击 Driver Details。

### 后续步骤

- 如果您安装了 NetKVM 驱动程序，可能还需要配置 Windows 客户机的网络参数。如需更多信息，请参阅[配置 NetKVM 驱动程序参数](#)。

### 17.2.1.4. 在 Windows 客户端中更新 virtio 驱动程序

要在 Windows 客户机操作系统(OS)上更新 KVM virtio 驱动程序，如果 Windows OS 版本支持它，您可以使用 Windows Update 服务。如果没有，请从附加到 Windows 虚拟机(VM)的virtio 驱动程序安装介质中重新安装驱动程序。

#### 先决条件

- 安装了 virtio 驱动程序的 Windows 客户机操作系统。
- 如果不使用 Windows Update，则必须将带有最新 KVM virtio 驱动程序的安装介质附加到 Windows 虚拟机。有关准备该介质的步骤，请参考[在主机中准备 virtio 驱动程序安装介质](#)。

#### 步骤 1：使用 Windows Update 更新驱动程序

在 Windows 10 中，Windows Server 2016 及更新版本的操作系统，检查 Windows 更新图形界面是否有可用的 驱动程序更新：

1. 启动 Windows 虚拟机并登录到其客户机操作系统。
2. 导航到 Optional updates 页面：  
Settings → Windows Update → Advanced options → Optional updates
3. 安装 Red Hat, Inc的所有更新。

#### 步骤 2：通过重新安装驱动程序来更新驱动程序

在 Windows 10 和 Windows Server 2016 之前的操作系统上，或者操作系统无法访问 Windows Update，请重新安装驱动程序。这会将 Windows 客户机操作系统网络配置恢复为默认值(DHCP)。如果要保留自定义网络配置，则需要使用 netsh 程序创建备份和恢复它：

1. 启动 Windows 虚拟机并登录到其客户机操作系统。
2. 打开 Windows 命令提示：
  - a. 使用 Super+R 键盘快捷键。
  - b. 在显示的窗口中，键入 cmd 并按 Ctrl+Shift+Enter 以以管理员身份运行。
3. 使用 Windows 命令提示备份 OS 网络配置：
 

```
C:\WINDOWS\system32\netsh dump > backup.txt
```
4. 从附加的安装介质重新安装 KVM virtio 驱动程序。执行以下操作之一：
  - 使用 Windows 命令提示重新安装驱动程序，其中 X 是安装介质驱动器符。以下命令安装所有 virtio 驱动程序。
    - 如果使用 64 位 vCPU：

```
C:\WINDOWS\system32\msiexec.exe /i X:\virtio-win-gt-x64.msi /passive /norestart
```
  - 如果使用 32 位 vCPU：

```
C:\WINDOWS\system32\msiexec.exe /i X:\virtio-win-gt-x86.msi /passive /norestart
```

- 在不重启虚拟机的情况下，[使用图形界面](#)重新安装驱动程序。
5. 使用 Windows 命令提示恢复 OS 网络配置：

```
C:\WINDOWS\system32\netsh -f backup.txt
```

6. 重启虚拟机以完成驱动程序安装。

## 其它资源

- [Microsoft 文档中有关 Windows 更新的文档](#)

### 17.2.1.5. 在 Windows 客户端中安装 QEMU 客户机代理

要允许主机在 Windows 虚拟机(VM)上执行各种操作，您可以安装 QEMU 客户机代理。为此，请将包含 QEMU 客户机代理安装程序的存储设备添加到现有虚拟机或者创建新虚拟机时，并在 Windows 客户机操作系统上安装驱动程序。

此流程提供了使用图形界面安装客户机代理的说明。您也可以使用 [Microsoft Windows Installer\(MSI\)](#)命令行界面。

## 先决条件

- 带有客户机代理的安装介质附加到虚拟机。有关准备该介质的步骤，请参考[在主机中准备 virtio 驱动程序安装介质](#)。

## 流程

1. 在 Windows 客户机操作系统中，打开 File Explorer 应用程序。
2. 单击 This PC。
3. 在 Devices and drives 窗格中，打开 virtio-win 介质。
4. 打开 guest-agent 文件夹。
5. 根据在虚拟机上安装的操作系统，运行以下命令之一：
  - 如果使用 32 位操作系统，请运行 `qemu-ga-i386.msi` 安装程序。
  - 如果使用 64 位操作系统，请运行 `qemu-ga-x86_64.msi` 安装程序。

## 验证

1. 在 Windows 虚拟机上，进入 Services 窗口。  
Computer Management > Services
2. 确保 QEMU 客户机代理服务的 Status 为 Running。

### 17.2.2. 启用 Hyper-V enlightenments

Hyper-V enlightenments 为 KVM 提供了一个模拟 Microsoft Hyper-V hypervisor 的方法。这提高了 Windows 虚拟机的性能。

以下小节提供了有关支持的 Hyper-V enlightenments 以及如何启用它们的信息。

### 17.2.2.1. 在 Windows 虚拟机上启用 Hyper-V enlightenments

Hyper-V enlightenments 在运行在 RHEL 8 主机上的 Windows 虚拟机(VM)中提供了更好的性能。有关如何启用它们的步骤，请查看以下操作。

#### 流程

1. 使用 `virsh edit` 命令打开虚拟机的 XML 配置。例如：

```
virsh edit windows-vm
```

2. 将以下 `<hyperv>` 子部分添加到 XML 的`<features>` 部分：

```
<features>
[...]
<hyperv>
<relaxed state='on' />
<vapic state='on' />
<spinlocks state='on' retries='8191' />
<vpindex state='on' />
<runtime state='on' />
<syncic state='on' />
<stimer state='on' >
 <direct state='on' />
</stimer>
<frequencies state='on' />
</hyperv>
[...]
</features>
```

如果 XML 已包含 `<hyperv>` 子部分，请按上所示进行修改。

3. 按如下方式更改配置的 `clock` 部分：

```
<clock offset='localtime'>
...
<timer name='hypervclock' present='yes' />
</clock>
```

4. 保存并退出 XML 配置。

5. 如果虚拟机正在运行，请重新启动它。

#### 验证

- 使用 `virsh dumpxml` 命令显示正在运行的虚拟机的 XML 配置。如果包括以下段，则虚拟机上启用了 Hyper-V enlightenments。

```
<hyperv>
<relaxed state='on' />
<vapic state='on' />
<spinlocks state='on' retries='8191' />
<vpindex state='on' />
<runtime state='on' />
<syncic state='on' />
```

```

<stimer state='on'>
 <direct state='on' />
</stimer>
<frequencies state='on' />
</hyperv>

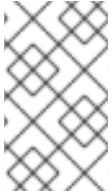
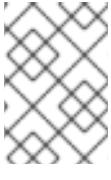
<clock offset='localtime'>
...
<timer name='hypervclock' present='yes' />
</clock>

```

### 17.2.2.2. 可配置 Hyper-V enlightenments

您可以配置特定的 Hyper-V 功能来优化 Windows 虚拟机。下表提供了有关这些可配置 Hyper-V 功能及其值的信息。

表 17.1. 可配置 Hyper-V 功能

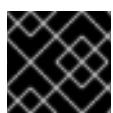
Enlightenment	描述	值
crash	<p>向虚拟机提供 MSR，可用于在虚拟机崩溃时存储信息和日志。 QEMU 日志中提供的信息。</p>  <p><b>注意</b></p> <p>如果启用了 hv_crash，则不会创建 Windows 崩溃转储。</p>	on, off
evmcs	<p>在 L0(KVM) 和 L1(Hyper-V) hypervisor 之间实施半虚拟化协议，从而使 L2 更快地退出到 hypervisor。</p>  <p><b>注意</b></p> <p>这个功能只供 Intel 处理器使用。</p>	on, off
frequencies	启用 Hyper-V frequency Machine Specific Registers (MSR)。	on, off
ipi	启用半虚拟化处理器间中断(IPI)支持。	on, off

Enlightenment	描述	值
no-nonarch-coresharing	通知客户端操作系统：虚拟处理器永远不会共享物理核，除非它们被报告为同级 SMT 线程。Windows 和 Hyper-V 客户机需要这些信息才能适当地减轻与并发多线程 (SMT) 相关的 CPU 漏洞。	on, off, auto
reenlightenment	仅在迁移期间发生时间戳计数器 (TSC) 频率更改时通知。它还允许 guest 继续使用旧频率，直到准备好切换至新频率。	on, off
relaxed	禁用 Windows 完整性检查，当虚拟机在负载较重的主机上运行时，该检查通常会导致 BSOD。这和 Linux 内核选项 no_timer_check 类似，它会在 Linux 在 KVM 中运行时自动启用。	on, off
runtime	设定运行客户机代码以及代表客户端代码的处理器时间。	on, off
spinlocks	<ul style="list-style-type: none"> <li>● 虚拟机的操作系统用来通知 Hyper-V，调用虚拟处理器正在尝试获取一个可能由同一分区中其他虚拟处理器占有的资源。</li> <li>● Hyper-V 用来向虚拟机的操作系统指明在向 Hyper-V 指示过度旋转情形之前应该尝试获取 spinlock 的次数。</li> </ul>	on, off
stimer	为虚拟处理器启用合成计时器。请注意，某些 Windows 版本在未提供这种启示时将恢复使用 HPET（或在 HPET 不可用时使用 RTC），这可能导致大量 CPU 消耗，即使虚拟 CPU 处于空闲状态。	on, off
stimer-direct	当通过正常的中断发送过期事件时，启用合成计时器。	on, off.
sync	与 stimer 一起激活合成计时器。Windows 8 以周期性模式使用此功能。	on, off

Enlightenment	描述	值
time	<p>启用以下虚拟机可用的特定于 Hyper-V 的时钟源,</p> <ul style="list-style-type: none"> <li>● 基于 MSR 的 82 个 Hyper-V 时钟源 (HV_X64_MSR_TIME_REF_COUNT, 0x40000020)</li> <li>● 参考通过 MSR 启用的 TSC 83 页面 (HV_X64_MSR_REFERENCED_TSC, 0x40000021)</li> </ul>	on, off
tlbflush	清除虚拟处理器的 TLB。	on, off
vapic	启用虚拟 APIC，其提供 MSR 对高使用率、内存映射高级可编程中断控制器(APIC)寄存器的加速访问。	on, off
vendor_id	设置 Hyper-V 厂商 id。	<ul style="list-style-type: none"> <li>● on, off</li> <li>● id 值 - 最多 12 个字符的字符串</li> </ul>
vpindex	启用虚拟处理器索引。	on, off

### 17.2.3. 配置 NetKVM 驱动程序参数

安装 NetKVM 驱动程序后，您可以对其进行配置以更好地适合您的环境。以下流程中列出的参数可以使用 Windows 设备管理器(**devmgmt.msc**)进行配置。



#### 重要

修改驱动程序的参数会导致 Windows 重新加载该驱动程序。这会中断现有的网络活动。

#### 先决条件

- NetKVM 驱动程序安装在虚拟机上。  
如需更多信息，请参阅[Windows 虚拟机安装 KVM 半虚拟驱动程序](#)。

#### 流程

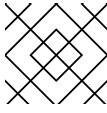
1. 打开 Windows Device Manager。  
有关打开设备管理器的详情，请参考 Windows 文档。
2. 找到 Red Hat VirtIO Ethernet Adapter。
  - a. 在 Device Manager 窗口中，点 Network adapters 旁边的 +。

- b. 在网络适配器列表中，双击 Red Hat VirtIO Ethernet Adapter。该设备的 Properties 窗口将打开。
3. 查看设备参数。  
在 Properties 窗口中，点击 Advanced 选项卡。
4. 修改设备参数。
  - a. 点击您要修改的参数。  
此时会显示那个参数的选项。
  - b. 根据需要修改选项。  
有关 NetKVM 参数选项的信息，请参考 [NetKVM 驱动程序参数](#)。
  - c. 点 OK 保存更改。

#### 17.2.4. NetKVM 驱动程序参数

下表提供了可配置的 NetKVM 驱动程序日志记录参数。

表 17.2. 日志参数

参数	描述
Logging.Enable	确定是否启用日志记录的布尔值。默认值为 Enabled。
Logging.Level	<p>定义日志级别的整数。当整数增加时，日志的详细程度也会增加。</p> <ul style="list-style-type: none"> <li>● 默认值为 0（仅错误）。</li> <li>● 1-2 添加配置信息。</li> <li>● 3-4 添加数据包流信息。</li> <li>● 5-6 添加中断以及 DPC 级别追踪信息。</li> </ul> <p> <b>注意</b> 高日志级别会减慢您的虚拟机速度。</p>

下表提供了可配置的 NetKVM 驱动程序初始参数信息。

表 17.3. 初始参数

参数	描述
分配 MAC	为半虚拟 NIC 定义本地管理的 MAC 地址的字符串。默认不设置。

参数	描述
Init.ConnectionRate(Mb)	代表每秒（以 MB 为单位）的连接率的整数。 Windows 2008 及更新版本的默认值为 10G（每秒 10,000MB）。
Init.Do802.1PQ	启用 Priority/VLAN 标签填充和删除支持的布尔值。 默认值为 Enabled。
Init.MTUSize	定义最大传输单元（MTU）的整数。默认值为 1500。 从 500 到 65500 的值都可以接受。
Init.MaxTxBuffers	代表将被分配的 TX 环描述符数的整数。  默认值为 1024。  有效值有：16、32、64、128、256、512 和 1024。
init.MaxRxBuffers	代表将要分配的 RX 环描述符数的整数。  默认值为 256。  有效值有：16、32、64、128、256、512 和 1024。
Offload.Tx.Checksum	指定 TX checksum 卸载模式。  在 Red Hat Enterprise Linux 8 中，这个参数的有效值为：  * All（默认），为 IPv4 和 IPv6 都启用 IP、TCP 和 UDP checksum offloading  * TCP/UDP(v4,v6)，为 IPv4 和 IPv6 都启用 TCP 和 UDP checksum offloading  * TCP/UDP(v4)，只为 IPv4 启用 TCP 和 UDP checksum offloading  * TCP(v4)，只为 IPv4 启用 TCP checksum

### 17.2.5. 在 Windows 虚拟机中优化后台进程

要优化运行 Windows OS 的虚拟机(VM)的性能，您可以配置或禁用各种 Windows 进程。



#### 警告

如果您更改配置，某些进程可能无法按预期工作。

### 流程

您可以通过执行以下任一组合来优化 Windows 虚拟机：

- 删除未使用的设备，如 USB 或 CD-ROM，并禁用端口。
- 禁用后台服务，如 SuperFetch 和 Windows Search。有关停止服务的详情，请参考[禁用系统服务](#)或[Stop-Service](#)。
- 禁用 `useplatformclock`。为此，请运行以下命令：

```
bcdedit /set useplatformclock No
```

- 检查和禁用不必要的调度任务，如调度的磁盘清除。有关如何操作的更多信息，请参阅[禁用调度任务](#)。
- 确定磁盘没有加密。
- 减少服务器应用程序的周期性活动。您可以编辑对应的计时器。如需更多信息，请参阅[多媒体计时器](#)。
- 关闭虚拟机上的 Server Manager 应用程序。
- 禁用 antivirus 软件。请注意，禁用 antivirus 可能会破坏虚拟机的安全。
- 禁用屏保。
- 在没有使用时，仍然将 Windows OS 保持在登录屏幕中。

### 17.3. 在 WINDOWS 虚拟机中启用标准硬件安全性

要保护 Windows 虚拟机，您可以使用 Windows 设备的标准硬件功能启用基本级别的安全性。

#### 先决条件

- 请确定您安装了最新的 WHQL 认证的 VirtIO 驱动程序。
- 确保虚拟机固件支持 UEFI 引导。
- 在您的主机上安装 edk2-OVMF 软件包。

```
yum install edk2-ovmf
```

- 在您的主机上安装 vTPM 软件包。

```
yum install swtpm libtpms
```

- 确保虚拟机使用 Q35 机器架构。
- 请确定您有 Windows 安装介质。

#### 流程

1. 通过在虚拟机 XML 配置中的 `<devices>` 部分中添加以下参数来启用 TPM 2.0。

```
<devices>
[...]
```

```

<tpm model='tpm-crb'>
 <backend type='emulator' version='2.0' />
</tpm>
[...]
</devices>

```

2. 在 UEFI 模式中安装 Windows。有关如何操作的更多信息，请参阅[创建 SecureBoot 虚拟机](#)。
3. 在 Windows 虚拟机上安装 VirtIO 驱动程序。有关如何操作的更多信息，请参阅[在 Windows 客户端上安装 virtio 驱动程序](#)。
4. 在 UEFI 中，启用安全引导。有关如何进行此操作的更多信息，请参阅[安全引导](#)。

## 验证

- 确定 Windows 机器中的设备安全性页面显示以下信息：  
settings > Update & Security > Windows Security > device Security

Your device meets the requirements for standard hardware security.

## 17.4. 在 WINDOWS 虚拟机上启用增强的硬件安全

为进一步保护 Windows 虚拟机(VM)，您可以启用基于虚拟化的代码完整性保护，也称为 hypervisor 保护的代码完整性(HVCI)。

### 先决条件

- 确保启用了标准硬件安全。如需更多信息，请参阅[在 Windows 虚拟机上启用标准硬件安全性](#)。
- 确保启用了 Hyper-V enlightenments。如需更多信息，请参阅[启用 Hyper-Vlightenments](#)。

### 流程

1. 打开 Windows 虚拟机的 XML 配置。以下示例打开 Example-L1 虚拟机的配置：

# virsh edit Example-L1

2. 在 <cpu> 部分下，指定 CPU 模式并添加策略标记。

### 重要

- 对于 Intel CPU，启用 vmx 策略标记。
- 对于 AMD CPU，启用 svm 策略标记。
- 如果您不想指定自定义 CPU，您可以将 <cpu mode='custom'> 设置为 host-passthrough。

```

<cpu mode='custom' match='exact' check='partial'>
 <model fallback='allow'>Skylake-Client-IBRS</model>
 <topology sockets='1' dies='1' cores='4' threads='1' />
 <feature policy='require' name='vmx' />
</cpu>

```

3. 保存 XML 配置并重启虚拟机。
4. 在虚拟机操作系统中，导航到 Core 隔离详情页面：  
settings > Update & Security > Windows Security > Device Security > Core isolated details
5. 切换开关以启用 内存完整性。
6. 重启虚拟机。



#### 注意

有关启用 HVCI 的其他方法，请查看相关的 Microsoft 文档。

#### 验证

- 确保 Windows 虚拟机上的 Device Security 页面显示以下信息：  
settings > Update & Security > Windows Security > device Security
  - Your device meets the requirements for enhanced hardware security.
- 或者，检查 Windows 虚拟机上的系统信息：
  - a. 在命令提示符下运行 msinfo32.exe。
  - b. 检查是否在基于虚拟化的安全服务运行 下列出了凭据保护，Hypervisor 强制的代码完整性。

## 17.5. 后续步骤

- 要在 RHEL 8 主机及其 Windows 虚拟机间共享文件，您可以使用 [NFS](#)。

## 第 18 章 创建嵌套虚拟机

在 RHEL 8 系统及更新版本上，可以创建嵌套的虚拟机(VM)。这意味着，运行在 RHEL 8 物理主机 (level 0 或 L0) 上的 RHEL 8 虚拟机（也称为 level 1 或 L1），可以充当 hypervisor，并创建自己的虚拟机 (level 2 或 L2)。

换句话说，L0 主机可以运行 L1 虚拟机(VM)，每个 L1 虚拟机都可以托管自己的 L2 虚拟机。请注意，在这种情况下，L0 和 L1 主机必须是 RHEL 8 系统，而 L2 客户机可以是任何受支持的 RHEL 或 Windows 系统。



### 警告

红帽目前仅作为 [技术预览](#) 提供嵌套虚拟化，因此不受支持。

此外，红帽不建议在用户的生产环境中使用嵌套虚拟化，因为 [功能上存在各种限制](#)。相反，嵌套虚拟化主要用于开发和测试场景，例如：

- 在受限环境中调试虚拟机监控程序
- 在有限的物理资源中测试较大的虚拟部署

还可以在多个架构上创建嵌套虚拟机，如 [Intel](#)、[AMD](#)、[IBM POWER9](#) 和 [IBM Z](#)。请注意，无论使用什么架构，嵌套都是一个 [技术预览](#)，因此不受红帽支持。

### 18.1. 在 INTEL 上创建嵌套虚拟机

按照以下步骤在 Intel 主机中启用和配置嵌套虚拟化。



### 警告

目前嵌套虚拟化仅作为 Intel 架构上的一个 [技术预览](#) 提供，因此不受支持。

#### 先决条件

- 一个 L0 RHEL 8 主机运行一个 L1 虚拟机(VM)。
- 管理程序 CPU 必须支持嵌套虚拟化。若要验证，可在 L0 hypervisor 上使用 `cat /proc/cpuinfo` 命令。如果命令的输出中包含 `vmx` 和 `ept` 标志，则可以创建 L2 虚拟机。这通常是 Intel Xeon v3 内核及更新版本上的情况。
- 确定在 L0 主机上启用了嵌套虚拟化：

```
cat /sys/module/kvm_intel/parameters/nested
```

- 如果命令返回 1 或 Y，则表示启用了该特性。跳过剩余的先决条件步骤，并继续执行流程部分。
- 如果该命令返回 0 或者 N，而您的系统支持嵌套虚拟化，则使用以下步骤启用该功能。
  - i. 卸载 `kvm_intel` 模块：

```
modprobe -r kvm_intel
```

- ii. 激活嵌套功能：

```
modprobe kvm_intel nested=1
```

- iii. 现在启用嵌套功能，但只有在下次重启 LO 主机时才启用。要永久启用它，请在 `/etc/modprobe.d/kvm.conf` 文件中添加以下行：

```
options kvm_intel nested=1
```

## 流程

1. 为嵌套虚拟化配置 L1 虚拟机。
  - a. 打开虚拟机的 XML 配置。以下示例打开 `Intel-L1` 虚拟机的配置：

```
virsh edit Intel-L1
```

- b. 通过编辑 `<cpu>` 元素将 VM 配置为使用 `host-passthrough` CPU 模式：

```
<cpu mode='host-passthrough'>
```

如果您需要虚拟机使用特定的 CPU 模型，请将虚拟机配置为使用自定义 CPU 模式。在 `<cpu>` 元素内，添加 `<feature policy='require' name='vmx'>` 元素，以及一个带有内部指定的 CPU 模型的 `<model>` 元素。例如：

```
<cpu mode ='custom' match ='exact' check='partial'>
 <model fallback='allow'>Haswell-noTSX</model>
 <feature policy='require' name='vmx'>
 ...
</cpu>
```

2. 在 L1 虚拟机中创建 L2 虚拟机。要做到这一点，请遵循 [创建 L1 虚拟机](#) 的相同步骤。

## 18.2. 在 AMD 上创建嵌套虚拟机

按照以下步骤在 AMD 主机中启用和配置嵌套虚拟化。



### 警告

目前嵌套虚拟化仅作为 AMD64 架构上的一个 [技术预览](#) 提供，因此不受支持。

## 先决条件

- 一个 L0 RHEL 8 主机运行一个 L1 虚拟机(VM)。
- 管理程序 CPU 必须支持嵌套虚拟化。若要验证，可在 L0 hypervisor 上使用 cat /proc/cpuinfo 命令。如果命令的输出中包含 svm 和 npt 标志，则可以创建 L2 虚拟机。这通常是 AMD EPYC 内核及更新版本上的情况。
- 确定在 L0 主机上启用了嵌套虚拟化：

```
cat /sys/module/kvm_amd/parameters/nested
```

- 如果命令返回 1 或 Y，则表示启用了该特性。跳过剩余的先决条件步骤，并继续执行流程部分。
- 如果命令返回 0 或 N，请使用以下步骤启用该功能。
  - i. 在 L0 主机上停止所有运行的虚拟机。
  - ii. 卸载 kvm\_amd 模块：

```
modprobe -r kvm_amd
```

- iii. 激活嵌套功能：

```
modprobe kvm_amd nested=1
```

- iv. 现在启用嵌套功能，但只有在下次重启 L0 主机时才启用。要永久启用它，请在 /etc/modprobe.d/kvm.conf 文件中添加以下内容：

```
options kvm_amd nested=1
```

## 流程

1. 为嵌套虚拟化配置 L1 虚拟机。

- a. 打开虚拟机的 XML 配置。以下示例将打开 AMD-L1 虚拟机的配置：

```
virsh edit AMD-L1
```

- b. 通过编辑 <cpu> 元素将 VM 配置为使用 host-passthrough CPU 模式：

```
<cpu mode='host-passthrough'>
```

如果您需要虚拟机使用特定的 CPU 模型，请将虚拟机配置为 使用自定义 CPU 模式。在 <cpu> 元素中，添加一个 <feature policy='require' name='svm'> 元素，以及一个 <model> 元素，使用内部指定的 CPU 模型。例如：

```
<cpu mode="custom" match="exact" check="none">
<model fallback="allow">EPYC-IBPB</model>
<feature policy="require" name="svm"/>
...
</cpu>
```

- 在 L1 虚拟机中创建 L2 虚拟机。要做到这一点，请遵循[创建 L1 虚拟机](#)的相同步骤。

### 18.3. 在 IBM Z 上创建嵌套的虚拟机

按照以下步骤在 IBM Z 主机中启用和配置嵌套虚拟化。



#### 警告

目前嵌套虚拟化仅作为 IBM Z 架构上的一个 [技术预览](#) 提供，因此不受支持。

#### 先决条件

- 一个 LO RHEL 8 主机运行一个 L1 虚拟机(VM)。
- 管理程序 CPU 必须支持嵌套虚拟化。若要验证确实是这种情况，请在 LO hypervisor 上使用 `cat /proc/cpuinfo` 命令。如果命令的输出中包含 `sie` 标志，则可以创建 L2 虚拟机。
- 确定在 LO 主机上启用了嵌套虚拟化：

```
cat /sys/module/kvm/parameters/nested
```

- 如果命令返回 1 或 Y，则表示启用了该特性。跳过剩余的先决条件步骤，并继续执行流程部分。
- 如果命令返回 0 或 N，请使用以下步骤启用该功能。

i. 在 LO 主机上停止所有运行的虚拟机。

ii. 卸载 kvm 模块：

```
modprobe -r kvm
```

iii. 激活嵌套功能：

```
modprobe kvm nested=1
```

iv. 现在启用嵌套功能，但只有在下次重启 LO 主机时才启用。要永久启用它，请在 `/etc/modprobe.d/kvm.conf` 文件中添加以下行：

```
options kvm nested=1
```

#### 流程

- 在 L1 虚拟机中创建 L2 虚拟机。要做到这一点，请遵循[创建 L1 虚拟机](#)的相同步骤。

### 18.4. 在 IBM POWER9 上创建嵌套虚拟机

按照以下步骤在 IBM POWER9 主机中启用和配置嵌套虚拟化。



## 警告

目前嵌套虚拟化仅作为 IBM POWER9 架构上的一个 [技术预览](#) 提供，因此不受支持。此外，在之前版本的 IBM POWER 系统（如 IBM POWER8）上不能创建嵌套虚拟机（VM）。

## 先决条件

- 一个 L0 RHEL 8 主机正在运行一个 L1 虚拟机。L1 虚拟机使用 RHEL 8 作为客户机操作系统。
- 在 L0 主机上启用嵌套虚拟化：

```
cat /sys/module/kvm_hv/parameters/nested
```

- 如果命令返回 1 或 Y，则表示启用了该特性。跳过剩余的先决条件步骤，并继续执行流程部分。
- 如果命令返回 0 或 N，请使用以下步骤启用该功能：

- i. 在 L0 主机上停止所有运行的虚拟机。

- ii. 卸载 kvm 模块：

```
modprobe -r kvm_hv
```

- iii. 激活嵌套功能：

```
modprobe kvm_hv nested=1
```

- iv. 现在启用嵌套功能，但只有在下次重启 L0 主机时才启用。要永久启用它，请在 /etc/modprobe.d/kvm.conf 文件中添加以下行：

```
options kvm_hv nested=1
```

## 流程

1. 为确保 L1 虚拟机能够创建 L2 虚拟机，请将 cap-nested-hv 参数添加到 L1 虚拟机的机器类型中。要做到这一点，使用 virsh edit 命令修改 L1 虚拟机的 XML 配置，并将以下行添加到 <features> 部分：

```
<nested-hv state='on'>
```

2. 在 L1 虚拟机中创建 L2 虚拟机。要做到这一点，请遵循 [创建 L1 虚拟机](#) 的相同步骤。为了显著提高 L2 虚拟机的性能，红帽建议在 L2 虚拟机的 XML 配置中也添加“cap-nested-hv”参数。具体步骤请查看上一步。

## 附加信息

- 请注意，使用 IBM POWER8 作为 L2 虚拟机的架构目前无法正常工作。

## 18.5. 嵌套虚拟化的限制

在使用嵌套虚拟化时请记住以下限制。



### 警告

红帽目前不支持嵌套虚拟化，仅作为 [技术预览](#) 提供嵌套。

#### 支持的构架

- L0 主机必须是 Intel、AMD、IBM POWER9 或 IBM Z 系统。嵌套虚拟化目前无法在其它构架中使用。

#### 支持的客户端操作系统

- 要创建嵌套虚拟机，您必须使用以下客户机操作系统(OS)：
  - 在 L0 主机上 - RHEL 8.2 及更新版本
  - 在 L1 虚拟机上 - RHEL 7.8 及更新版本, RHEL 8.2 及更新版本



### 注意

此支持不适用于在 L1 虚拟机中使用基于 RHEL 7 和 RHEL 8 的虚拟化产品。它们是：

- Red Hat Virtualization
- Red Hat OpenStack Platform
- OpenShift Virtualization

- 在 L2 虚拟机上 - 您必须使用以下操作系统之一：
  - RHEL 7.8 及更新版本
  - RHEL 8.2 及更新版本
  - Microsoft Windows Server 2016
  - Microsoft Windows Server 2019
- 另外，在 IBM POWER9 上，嵌套虚拟化当前仅在以下情况下可以工作：
  - L0 主机和 L1 虚拟机都使用 RHEL 8
  - L2 虚拟机使用 RHEL 8 或带有 rhel-alt 内核的 RHEL 7。
  - L1 VM 和 L2 虚拟机没有在 POWER8 兼容模式下运行。

#### Hypervisor 限制

- 目前，红帽仅在 RHEL-KVM 上支持嵌套。当 RHEL 用作 L0 hypervisor 时，您可以对 WSL 2 使用 RHEL 8 或 Windows 作为 L1 hypervisor。
- 在非 KVM L0 hypervisor（如 VMware ESXi 或 Amazon Web Services(AWS)）上使用 L1 RHEL 8 虚拟机时，在 RHEL 8 客户机操作系统中创建 L2 虚拟机可以正常工作，但不受支持。

## 功能限制

- 使用 L2 虚拟机作为 hypervisor，并创建 L3 客户机尚未经过合适的测试，预计无法正常工作。
- 如果在 L0 主机上启用了嵌套虚拟化，则迁移虚拟机目前无法在 AMD 系统上工作。
- 在 IBM Z 系统上，无法同时使用大页面支持存储和嵌套虚拟化。

```
modprobe kvm hpage=1 nested=1
modprobe: ERROR: could not insert 'kvm': Invalid argument
dmesg |tail -1
[90226.508366] kvm-s390: A KVM host that supports nesting cannot back its KVM
guests with huge pages
```

- L0 主机中的一些特性可能不适用于 L1 管理程序。  
例如，在 IBM POWER 9 硬件上，外部中断虚拟化引擎(XIVE)无法正常工作。但是，L1 虚拟机可以使用模拟的 XIVE 中断控制器来启动 L2 虚拟机。

## 第 19 章 诊断虚拟机问题

使用虚拟机(VM)时，您可能会遇到不同严重等级的问题。有些问题可能很快且容易修复，而另外一些问题，您可能需要捕获与虚拟机相关的数据和日志来报告或诊断问题。

以下章节提供了有关生成日志、诊断一些常见虚拟机问题以及报告这些问题的详细信息。

### 19.1. 生成 LIBVIRT 调试日志

要诊断虚拟机(VM)问题，生成和查看 libvirt 调试日志会很有帮助。当请求支持解决与虚拟机相关的问题时，附加调试日志也很有用。

以下章节解释了 [什么是调试日志](#)，如何 [将它们设为永久](#)，[在运行时启用它们](#)，以及在报告问题时 [附加它们](#)。

#### 19.1.1. 了解 libvirt 调试日志

调试日志是文本文件，其中包含关于虚拟机(VM)运行期间所发生的事件的数据。日志提供有关基本的服务器端功能方面的信息，如主机库和 libvirt 守护进程。日志文件还包含所有正在运行的虚拟机的标准错误输出(stderr)。

默认不启用 debug 日志记录，且必须在 libvirt 启动时启用。您可以对单个会话启用日志记录，也可以 [永久启用记录](#)。您还可以通过 [修改守护进程运行时设置](#)，在 libvirt 守护进程会话运行时启用日志。

在请求对虚拟机问题的支持时，[附加 libvirt 调试日志](#)也很有用。

#### 19.1.2. 为 libvirt 调试日志启用持久性设置

您可以将 libvirt 调试日志记录配置为在 libvirt 启动时自动启用。默认情况下，libvirtd 是 RHEL 8 中唯一的 libvirt 守护进程。要在 libvirt 配置中进行持久化更改，您必须编辑位于 /etc/libvirt 目录中的 libvirtd.conf 文件。

#### 流程

- 1 在编辑器中打开 libvirtd.conf 文件。
- 2 根据您的要求替换或设置过滤器。

表 19.1. 调试过滤器的值

1	记录 libvirt 生成的所有消息。
2	记录所有非调试信息。
3	记录所有警告和错误消息。这是默认值。
4	仅记录错误消息。

#### 例 19.1. 日志过滤器的守护进程设置示例

以下设置：

- 记录来自 `remote,util.json` 和 `rpc` 层的所有错误和警告消息
- 仅记录来自 `event` 层的错误消息。
- 将过滤的日志保存到 `/var/log/libvirt/libvirt.log`

```
log_filters="3:remote 4:event 3:util.json 3:rpc"
log_outputs="1:file:/var/log/libvirt/libvirt.log"
```

3. 保存并退出。
4. 重启 libvirt 守护进程。

```
$ systemctl restart libvirtd.service
```

### 19.1.3. 在运行时启用 libvirt 调试日志

您可以修改 libvirt 守护进程的运行时设置，以启用调试日志并将其保存到一个输出文件中。

这在无法重启 libvirt 守护进程时非常有用，因为重启解决了问题，或者因为有另一个进程（如迁移或备份）同时在运行。如果您要在不编辑配置文件或重启守护进程的情况下尝试命令，修改运行时设置也很有用。

#### 先决条件

- 确保 `libvirt-admin` 软件包已安装。

#### 流程

1. 可选：备份活跃的日志过滤器集合。

```
virt-admin daemon-log-filters >> virt-filters-backup
```



#### 注意

建议您备份一组处于活动状态的过滤器，以便在生成日志后恢复它们。如果您没有恢复过滤器，则会继续记录消息，这可能会影响系统性能。

2. 使用 `virt-admin` 工具启用调试，并根据您的要求设置过滤器。

表 19.2. 调试过滤器的值

1	记录 libvirt 生成的所有消息。
2	记录所有非调试信息。
3	记录所有警告和错误消息。这是默认值。
4	仅记录错误消息。

**例 19.2. 日志过滤器的 virt-admin 设置示例**

以下命令：

- 记录来自 **远程**、**util.json** 和 **rpc** 层的所有错误和警告消息
- 仅记录来自 **event** 层的错误消息。

```
virt-admin daemon-log-filters "3:remote 4:event 3:util.json 3:rpc"
```

**3. 使用 virt-admin 工具将日志保存到特定的文件或目录中。**

例如，以下命令将日志输出保存到 /var/log/libvirt/ 目录中的 libvirt.log 文件中：

```
virt-admin daemon-log-outputs "1:file:/var/log/libvirt/libvirt.log"
```

**4. 可选：您还可以删除过滤器来生成包含所有与虚拟机相关的信息的日志文件。但不建议您这样做，因为这个文件可能包含由 libvirt 模块生成的大量冗余信息。**

- 使用 virt-admin 工具指定一组空白的过滤器。

```
virt-admin daemon-log-filters
Logging filters:
```

**5. 可选：使用备份文件将过滤器恢复到其原始状态。**

使用保存的值执行第二步，以恢复过滤器。

#### 19.1.4. 将 libvirt 调试日志附加到支持请求中

您可能需要请求额外的支持来诊断和解决虚拟机(VM)问题。强烈建议将调试日志附加到支持请求中，以确保支持团队可以访问所需的所有信息，从而快速解决与虚拟机相关的问题。

##### 流程

- 要报告问题并请求支持，[创建一个支持问题单](#)。
- 根据遇到的问题，将以下日志与您的报告一起附加：
  - 对于 libvirt 服务的问题，请附加主机上的 /var/log/libvirt/libvirt.log 文件。
  - 对于特定虚拟机的问题，请附加对应的日志文件。  
例如，对于 testguest1 虚拟机，请附加 testguest1.log 文件，该文件可在 /var/log/libvirt/qemu/testguest1.log 中找到。

##### 其它资源

- [如何给红帽支持提供日志文件？](#)

## 19.2. 转储虚拟机内核

要分析虚拟机(VM)崩溃或出现故障的原因，您可以将虚拟机内核转储到磁盘上的文件，以便稍后分析和诊断。

本节提供了简要的 [内核转储介绍](#)，并解释了如何将 [虚拟机内核转储](#) 到一个特定的文件。

### 19.2.1. 虚拟机内核转储的工作原理

虚拟机(VM)需要许多运行的进程来准确且高效地工作。在某些情况下，运行的虚拟机在使用时可能会意外终止或出现故障。重新启动虚拟机可能会导致数据被重置或丢失，从而难以诊断导致虚拟机崩溃的确切问题。

在这种情况下，您可以在重启虚拟机前使用 `virsh dump` 工具将虚拟机内核保存（或 [转储](#)）到一个文件中。内核转储文件包含虚拟机的原始物理内存镜像，其中包含有关虚拟机的详细信息。此信息可用于手动诊断，或使用 `crash` 等工具来诊断虚拟机问题。

#### 其它资源

- [crash 手册页](#)
- [crash Github 存储库](#)

### 19.2.2. 创建虚拟机内核转储文件

虚拟机(VM)内核转储包含在任何给定的时间有关虚拟机状态的详细信息。此信息与虚拟机的快照类似，可以帮助您在虚拟机出现故障或突然关闭时检测问题。

#### 先决条件

- 请确定您有足够的磁盘空间保存该文件。请注意，虚拟机消耗的空间取决于分配给虚拟机的 RAM 量。

#### 流程

- 使用 `virsh dump` 工具。

例如，以下命令将 `lander1` 虚拟机的内核、内存和 CPU 通用寄存器文件转储到`/core/file` 目录中的 `gargantua.file`：

```
virsh dump lander1 /core/file/gargantua.file --memory-only
Domain 'lander1' dumped to /core/file/gargantua.file
```



#### 重要

`crash` 工具不再支持 `virsh dump` 命令的默认文件格式。要使用 `crash` 分析内核转储文件，您必须使用 `--memory-only` 选项创建文件。

另外，在创建内核转储文件时，必须使用 `--memory-only` 选项将其附加到红帽支持问题单中。

#### 故障排除

如果 `virsh dump` 命令失败并显示 `System is deadlocked on memory` 错误，请确保为内核转储文件分配了足够的内存。要做到这一点，请使用以下 `crashkernel` 选项值。或者，请勿使用 `crashkernel`，它会自动分配内核转储内存。

```
crashkernel=1G-4G:192M,4G-64G:256M,64G-:512M
```

## 其它资源

- **virsh dump --help 命令**
- **virsh 手册页**
- [开一个支持问题单](#)

## 19.3. 回溯虚拟机进程

当与虚拟机(VM)相关的进程出现故障时，您可以使用 **gstack** 命令和进程标识符(PID)来生成一个故障进程的执行堆栈跟踪。如果该进程是线程组的一部分，那么也会跟踪所有线程。

### 先决条件

- 确保 GDB 软件包已安装。  
有关安装 GDB 和可用组件的详情，请参阅[安装 GNU 调试器](#)。
- 请确定您知道要追踪进程的 PID。  
您可以使用 **pgrep** 命令，后跟进程的名称来查找 PID。例如：

```
pgrep libvirt
22014
22025
```

### 流程

- 使用 **gstack** 工具，后跟您要回溯的进程的 PID。  
例如：以下命令追踪 PID 为 22014 的 libvirt 进程。

```
gstack 22014
Thread 3 (Thread 0x7f33edaf7700 (LWP 22017)):
#0 0x00007f33f81aef21 in poll () from /lib64/libc.so.6
#1 0x00007f33f89059b6 in g_main_context_iterate.isra () from /lib64/libglib-2.0.so.0
#2 0x00007f33f8905d72 in g_main_loop_run () from /lib64/libglib-2.0.so.0
...
...
```

## 其它资源

- **gstack 手册页**
- [GNU 调试器\(GDB\)](#)

## 用于报告虚拟机问题并提供日志的其他资源

要请求额外的帮助和支持，您可以：

- 使用 **redhat-support-tool** 命令行选项、Red Hat Portal UI 或使用 FTP 的几个不同的方法来提一个服务请求。
  - 要报告问题并请求支持，请参阅创建[支持问题单](#)。
- 提交服务请求时上传 SOS 报告以及日志文件。  
这样可保证红帽支持工程师可以参考所需的全部诊断信息。

- 有关 SOS 报告的更多信息，请参阅 [什么是 SOS 报告以及如何在 Red Hat Enterprise Linux 上创建它？](#)
- 有关附加日志文件的详情，请参考 [如何给红帽支持提供文件？](#)

## 第 20 章 RHEL 8 虚拟化的功能支持和限制

本文档提供了有关 Red Hat Enterprise Linux 8(RHEL 8)虚拟化中的功能支持和限制的信息。

### 20.1. RHEL 虚拟化支持如何工作

一组支持限制适用于 Red Hat Enterprise Linux 8(RHEL 8)中的虚拟化。这意味着，当您在 RHEL 8 中使用特定的功能或者超过特定分配的资源时，除非有特定的订阅计划，否则红帽将不支持这些虚拟机。

[RHEL 8 虚拟化中建议功能中列出的功能](#) 经过红帽测试并认证，可以在 RHEL 8 系统中使用 KVM 管理程序。因此，它们被完全支持，推荐在 RHEL 8 虚拟化中使用。

[RHEL 8 虚拟化中不支持的功能](#)，但不被支持，且不适用于在 RHEL 8 中使用。因此，红帽强烈建议您不要在 RHEL 8 中使用这些功能。

[RHEL 8 虚拟化中的资源分配限制](#) 列出了 RHEL 8 中 KVM 虚拟机中支持的最大特定资源量。红帽不支持超过这些限制的客户端。

另外，除非有其他说明，RHEL 8 虚拟化文档使用的所有功能和解决方案都被支持。然而，其中有些还没有进行充分测试，因此可能无法完全优化。



#### 重要

这些限制不适用于红帽提供的其他虚拟化解决方案，如 Red Hat Virtualization(RHV)、OpenShift Virtualization 或 Red Hat OpenStack Platform(RHOSP)。

### 20.2. RHEL 8 虚拟化中的推荐功能

建议以下功能与 Red Hat Enterprise Linux 8(RHEL 8)中包含的 KVM hypervisor 一起使用：

#### 主机系统构架

带有 KVM 的 RHEL 8 只有在以下主机架构上被支持：

- AMD64 和 Intel 64
- IBM Z - IBM z13 系统及更新版本
- IBM POWER8
- IBM POWER9

任何其它硬件构架都不支持将 RHEL 8 用作 KVM 虚拟化主机，红帽强烈建议您这样做。值得注意的是，这包括 64 位 ARM 架构(ARM 64)。



#### 注意

RHEL 8 文档主要描述了 AMD64 和 Intel 64 特性和用法。有关在不同构架上使用 RHEL 8 虚拟化的具体信息，请参考：

- [在 IBM POWER 中使用虚拟化](#)
- [在 IBM Z 中使用虚拟化](#)

#### 客户机操作系统

红帽为使用特定客户机操作系统(OS)的 KVM 虚拟机提供支持。有关支持的客户机操作系统的详细列表，请查看红帽知识库中的认证客户机操作系统。

但请注意，默认情况下您的客户机操作系统不使用与您的主机相同的订阅。因此，您必须激活单独的许可或者订阅方可使客户机操作系统正常工作。

## 机器类型

为确保虚拟机与您的主机架构兼容，且客户机操作系统最佳运行，虚拟机必须使用合适的机器类型。

当 [使用命令行创建虚拟机](#) 时，`virt-install` 工具提供了多种设置机器类型的方法。

- 当使用 `--os-variant` 选项时，`virt-install` 会自动选择主机 CPU 推荐的和客户机操作系统支持的机器类型。
- 如果不使用 `--os-variant` 或需要其他机器类型，请使用 `--machine` 选项来明确指定机器类型。
- 如果您指定了一个不受支持或与主机不兼容的 `--machine` 值，`virt-install` 将会失败，并显示一条错误信息。

在支持的架构上推荐的 KVM 虚拟机的机器类型，以及 `--machine` 选项的相应值如下。Y 代表 RHEL 8 的最新次要版本。

- 在 Intel 64 和 AMD64(x86\_64)上：`pc-q35-rhel8.Y.0` → `--machine=q35`
- 在 IBM Z(s390x)上：`s390-ccw-virtio-rhel8.Y.0` → `--machine=s390-ccw-virtio`
- 在 IBM POWER(PPC)上，`pseries-rhel8.Y.0` → `--machine=pseries`

要获得现有虚拟机的机器类型：

```
virsh dumpxml VM-name | grep machine=
```

要查看主机上支持的完整的机器类型列表：

```
/usr/libexec/qemu-kvm -M help
```

## 其它资源

- [RHEL 8 虚拟化不支持的功能](#)
- [RHEL 8 虚拟化中的资源分配限制](#)
- [虚拟机迁移支持的主机](#)

### 20.3. RHEL 8 虚拟化不支持的功能

Red Hat Enterprise Linux 8(RHEL 8)中的 KVM hypervisor 不支持以下特性：



## 重要

这些限制可能不适用于红帽提供的其他虚拟化解决方案，如 Red Hat Virtualization(RHV)、OpenShift Virtualization 或 Red Hat OpenStack Platform(RHOSP)。

RHV 4.2 及更高版本或 RHOSP 13 及更高版本支持的特性，如下所述。

## 主机系统构架

在 RHEL 8 虚拟化中没有包括在推荐功能中的任何主机架构中不支持 KVM 的 RHEL 8。

值得注意的是，红帽不支持在 RHEL 8 中对 KVM 虚拟化使用 64 位 ARM 架构(ARM 64)的系统。

## 客户机操作系统

在 RHEL 8 主机上不支持使用以下客户机操作系统(OS)的 KVM 虚拟机(VM)：

- Microsoft Windows 8.1 及更早版本
- Microsoft Windows Server 2008 R2 及更早版本
- macOS
- 用于 x86 系统的 Solaris
- 2009 年之前发布的所有操作系统

有关 RHEL 主机、RHV 或其他虚拟化解决方案支持的客户机操作系统列表，请参阅 [Red Hat OpenStack Platform](#)、[Red Hat Virtualization](#)、[OpenShift Virtualization](#) 和带有 KVM 的 [Red Hat Enterprise Linux](#) 中的经认证的客户机操作系统。

## 在容器中创建虚拟机

红帽不支持在任意类型的容器中创建 KVM 虚拟机，其中包括 RHEL 8 管理程序元素（如 QEMU 模拟器或 libvirt 软件包）。

要在容器中创建虚拟机，红帽建议您使用 [OpenShift Virtualization](#) 产品。

## QEMU 命令行

QEMU 是 RHEL 8 中虚拟化架构的基本组件，但难以手动管理，而不正确的 QEMU 配置可能会导致安全漏洞。因此，红帽不支持使用 `qemu-*` 命令行工具，如 `qemu-kvm`。反之，使用 `libvirt` 工具，如 `virsh`、`virt-install` 和 `virt-xml`，根据最佳实践来编排 QEMU。

## vCPU 热拔

从正在运行的虚拟机中删除虚拟 CPU(vCPU)，也称为 vCPU 热插拔，在 RHEL 8 中不支持。

请注意 RHV 支持 vCPU 热拔。详情请查看 [热插拔 VCPU](#)。

## 内存热拔

删除附加在正在运行的虚拟机上的内存设备，也称为内存热插拔，在 RHEL 8 中不支持。

请注意 RHV，支持内存热拔，但仅在运行带有特定客户机配置的 RHEL 的虚拟机上进行。详情请参阅[热拔虚拟内存](#)。

## QEMU 端的 I/O 节流

使用 `virsh blkdeviotune` 工具为虚拟磁盘上的操作配置最大输入和输出级别，也称为 QEMU 侧 I/O 节流，在 RHEL 8 中不支持。

要在 RHEL 8 中设置 I/O 节流，请使用 `virsh blkiotune`。这也被称为 libvirt-side I/O 节流。具体步骤请查看 [虚拟机中的磁盘 I/O 节流](#)。

请注意 RHV 支持 QEMU 边的 I/O 节流。详情请参阅[存储服务质量](#)。

RHOSP 还支持 QEMU 边的 I/O 节流。详情请参阅 [RHOSP 存储指南](#) 中的 [在磁盘上设置资源限制](#) 和 [使用服务质量规范](#) 章节。

此外，OpenShift Virtualization 也支持 QEMU 侧 I/O 节流。

## 存储动态迁移

在主机之间迁移正在运行的虚拟机的磁盘镜像在 RHEL 8 中不支持。

请注意 RHV 支持存储实时迁移。详情请查看[实时存储迁移概述](#)。

RHOSP 还支持存储实时迁移，但有一些限制。详情请参阅[迁移卷](#)。

## 实时快照

创建或加载正在运行的虚拟机的快照，也称为实时快照，在 RHEL 8 中不支持。

另外，请注意，非实时虚拟机快照在 RHEL 8 中已弃用。因此，支持创建或载入关闭虚拟机的快照，但红帽建议您不要使用它。

请注意，RHV 支持实时快照。详情请参阅 [Red Hat Virtualization 中的实时快照](#)。

RHOSP 还支持实时快照。详情请参阅[把虚拟机导入到 overcloud](#) 部分。

## vhost-user

RHEL 8 不支持用户空间 vHost 接口的实现。

请注意，RHOSP 中支持 vhost-user，但仅对 virtio-net 接口。详情请查看 [virtio-net 实现](#) 和 [vhost 用户端口](#)。

## S3 和 S4 系统电源状态

不支持将虚拟机挂起到 Suspend to RAM (S3) 或 Suspend to disk (S4) 系统电源状态。请注意，这些功能被默认禁用，启用它们会使虚拟机不受红帽支持。

请注意，红帽提供的任何其他虚拟化解决方案目前还不支持 S3 和 S4 状态。

## 多路径 vDisk 中的 s3-PR

RHEL 8 不支持在多路径 vDisk 上的 SCSI3 持久保留(S3-PR)。因此，RHEL 8 不支持 Windows 集群。

请注意 RHV 支持多路径 vDisk 上的 S3-PR。因此，如果您需要 Windows 集群支持，红帽建议使用 RHV 作为虚拟化解决方案。详情请查看 [RHV 客户端中的集群支持](#)。

## virtio-crypto

不支持在 RHEL 8 中使用 `virtio-crypto` 设备，因此强烈建议不要使用它。

请注意，红帽提供的任何其他虚拟化解决方案不支持 `virtio-crypto` 设备。

## 增量实时备份

在 RHEL 8 中不支持配置只保存自上次备份（也称为增量实时备份）以来虚拟机更改的虚拟机备份，红帽强烈不建议使用它。

请注意，在 RHV 4.4 及更新的版本中，会作为技术预览提供增量实时备份。

### net\_failover

使用 `net_failover` 驱动程序设置自动网络设备故障切换机制，在 RHEL 8 中不支持。

请注意，红帽提供的任何其他虚拟化解决方案目前不支持 `net_failover`。

### 多 FD 迁移

使用多个文件描述符(FD)迁移虚拟机，也称为多 FD 迁移，在 RHEL 8 中不支持。

请注意，红帽提供的任何其他虚拟化解决方案不支持多FD 迁移。

### TPM 透传

在 RHEL 8 主机上不支持使用 `passthrough` 后端为虚拟机分配物理可信平台模块(TPM)设备。反之，使用 `vTPM` 功能，它使用仿真程序后端并被完全支持。

### virtiofs

使用 `virtiofs` 文件系统在主机及其虚拟机间共享文件，在 RHEL 8 中不支持。

但请注意，RHEL 9 支持使用 `virtiofs`。如需更多信息，请参阅在 [RHEL 9 中配置和管理虚拟化](#)。

### NVMe 设备

不支持将 Non-volatile Memory express(NVMe)设备作为带有 PCI-passthrough 的 PCIe 设备附加到虚拟机。

请注意，红帽提供的其它虚拟化解决方案不支持将 NVMe 设备附加到虚拟机。

### TCG

QEMU 和 libvirt 包含使用 QEMU Tiny Code Generator(TCG)的动态转换模式。这个模式不需要硬件虚拟化支持。但是，红帽不支持 TCG。

通过检查其 XML 配置可识别基于 TCG 的客户端，例如使用 "virsh dumpxml" 命令。

- TCG 客户端的配置文件包括以下行：

```
<domain type='qemu'>
```

- KVM 客户端的配置文件包含以下行：

```
<domain type='kvm'>
```

### SR-IOV InfiniBand 网络设备

不支持使用单根 I/O 虚拟化(SR-IOV)将 InfiniBand 网络设备附加到虚拟机。

### 其它资源

- [RHEL 8 虚拟化中的推荐功能](#)
- [RHEL 8 虚拟化中的资源分配限制](#)

## 20.4. RHEL 8 虚拟化中的资源分配限制

以下限制适用于可分配给 Red Hat Enterprise Linux 8(RHEL 8)主机上单个 KVM 虚拟机(VM)的虚拟化资源。



### 重要

这些限制不适用于红帽提供的其他虚拟化解决方案，如 Red Hat Virtualization(RHV)、OpenShift Virtualization 或 Red Hat OpenStack Platform(RHOSP)。

#### 每个虚拟机的最大 vCPU

有关在 RHEL 8 主机上运行的单个虚拟机支持的最大 vCPU 和内存量，请参阅[使用 KVM 的 Red Hat Enterprise Linux 虚拟化限制](#)

#### 每个虚拟机的 PCI 设备

RHEL 8 支持每个虚拟机总线 64 个 PCI 设备插槽，每个设备插槽支持 8 个 PCI 功能。这在虚拟机中启用了多功能时，每个总线理论上最多提供 512 PCI 功能，并且不使用 PCI 网桥。

每个 PCI 网桥添加一个新的总线，就有可能启用其他 512 个设备地址。但是，一些总线并不将所有 512 个设备地址都提供给用户使用；例如，根总线有多个内置设备占据了插槽。

#### 虚拟 IDE 设备

KVM 限制每台虚拟机最多有 4 个虚拟化 IDE 设备。

## 20.5. RHEL 8 支持虚拟化功能概述

下表提供了有关 RHEL 8 中所选的虚拟化特性的支持状态在所选的系统架构中的比较信息。

表 20.1. 设备热插和热拔

	Intel 64 和 AMD64	IBM Z	IBM POWER
CPU 热插	支持	支持	支持
CPU 热拔	不支持	不支持	不支持
内存热插拔	支持	不支持	支持
内存热拔	不支持	不支持	不支持
PCI 热插	支持	支持 [a]	支持
PCI 热拔	支持	支持 [b]	支持

[a] 需要使用 `virtio-*ccw` 设备，而不是 `virtio-*pci`

[b] 需要使用 `virtio-*ccw` 设备，而不是 `virtio-*pci`

表 20.2. 其他选择的功能

	Intel 64 和 AMD64	IBM Z	IBM POWER
NUMA 调整	支持	不支持	支持
SR-IOV 设备	支持	不支持	支持
virt-v2v 和 p2v	支持	不支持	不支持

请注意，一些不受支持的功能在其它红帽产品（如 Red Hat Virtualization 和 Red Hat OpenStack Platform）上被支持。如需更多信息，请参阅 [RHEL 8 虚拟化中的不支持的功能](#)。

## 其他资源

- 有关 RHEL 8 中虚拟机不支持的功能的完整列表，请参阅 [RHEL 8 虚拟化中的不支持的功能](#)。
- 有关 IBM Z 构架虚拟化详情，请参阅 [IBM Z 中的虚拟化与 AMD64 和 Intel 64 的不同](#)。
- 有关 IBM POWER 构架虚拟化详情，请参阅 [IBM POWER 中的虚拟化与 AMD64 和 Intel 64 不同](#)。