



OpenShift Container Platform 4.2

架构

OpenShift Container Platform 4.2 架构概述

OpenShift Container Platform 4.2 架构

OpenShift Container Platform 4.2 架构概述

法律通告

Copyright © 2020 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

本文概述 OpenShift Container Platform 4.2 中的平台和应用架构。

目录

第 1 章 OPENSIFT CONTAINER PLATFORM 架构	3
1.1. OPENSIFT CONTAINER PLATFORM 简介	3
第 2 章 安装和更新	7
2.1. OPENSIFT CONTAINER PLATFORM 安装概述	7
2.2. 关于 OPENSIFT CONTAINER PLATFORM 更新服务	11
2.3. 支持非受管 OPERATOR 的策略	11
第 3 章 OPENSIFT CONTAINER PLATFORM CONTROL PLANE	13
3.1. 了解 OPENSIFT CONTAINER PLATFORM CONTROL PLANE	13
第 4 章 了解 OPENSIFT CONTAINER PLATFORM 开发	17
4.1. 关于容器化应用程序开发	17
4.2. 构建一个简单容器	17
4.3. 为 OPENSIFT CONTAINER PLATFORM 创建 KUBERNETES 清单	20
4.4. 面向 OPERATOR 进行开发	22
第 5 章 RED HAT ENTERPRISE LINUX COREOS (RHCOS)	23
5.1. 关于 RHCOS	23
5.2. 查看 IGNITION 配置文件	25
5.3. 安装后更改 IGNITION 配置	27

第 1 章 OPENSIFT CONTAINER PLATFORM 架构

1.1. OPENSIFT CONTAINER PLATFORM 简介

OpenShift Container Platform 是用于开发和运行容器化应用程序的平台。它旨在允许支持的应用程序和数据中心从少量机器和应用程序扩展到为数百万客户端服务的数千台机器。

OpenShift Container Platform 以 Kubernetes 为基础，为大规模电信、流视频、游戏、银行和其他应用提供引擎技术。借助红帽开放技术中的实现，您可以将容器化应用程序从单一云扩展到内部和多云环境。

1.1.1. 关于 Kubernetes

尽管容器镜像和从中运行的容器是现代应用程序开发的主要构建块，但要大规模运行它们，则需要可靠且灵活的分发系统。Kubernetes 是编配容器的事实标准。

Kubernetes 是一个开源容器编配引擎，用于自动化容器化应用程序的部署、扩展和管理。Kubernetes 的一般概念比较简单：

- 从一个或多个 worker 节点开始，以运行容器工作负载。
- 从一个或多个 master 节点管理这些工作负载的部署。
- 将容器嵌套在称为 Pod 的部署单元中。使用 Pod 可以为容器提供额外的元数据，并可在单个部署实体中对多个容器进行分组。
- 创建特殊种类的资产。例如，服务由一组 Pod 及定义了访问方式的策略来表示。此策略可使容器连接到所需的服务，即便容器没有用于服务的特定 IP 地址。复制控制器（replication controller）是另一种特殊资产，用于指示一次需要运行多少个 Pod 副本。您可以使用此功能来自动扩展应用程序，以适应其当前的需求。

短短数年，Kubernetes 已在大量的云和本地环境中被采用。借助开源开发模型，拥护和可以通过为组件（如网络、存储和身份验证）实施不同的技术来扩展 Kubernetes 的功能。

1.1.2. 容器化应用程序的好处

与使用传统部署方法相比，使用容器化应用程序具有许多优势。过去应用程序要安装到包含所有依赖项的操作系统上，容器能让一个应用程序随身携带自己的依赖项。创建容器化应用程序有很多好处。

1.1.2.1. 操作系统的好处

容器使用不含内核的小型专用 Linux 操作系统。它们的文件系统、网络、cgroups、进程表和命名空间与主机 Linux 系统分开，但容器可以在必要时与主机无缝集成。容器以 Linux 为基础，因此可以利用快速创新的开源开发模型带来的所有优势。

因为每个容器都使用专用的操作系统，所以您能够在同一主机上部署需要冲突软件依赖项的不同应用程序。每个容器都带有各自的依赖软件，并且管理自己的接口，如网络和文件系统，因此应用程序无需争用这些资产。

1.1.2.2. 部署和扩展优势

如果您在应用程序的主要版本之间进行滚动升级，则可以持续改进应用程序，既不会造成停机，又能仍然保持与当前版本的兼容性。

您还可以与现有版本一起部署和测试应用程序的新版本。在部署了当前版本的同时，还部署应用程序的新版本。容器通过测试后，只要部署更多新容器并删除旧容器便可。

由于应用程序的所有软件依赖项都在容器本身内解决，因此数据中心的每台主机上都能使用通用操作系统。您无需逐一为应用主机配置特定的操作系统。当数据中心需要更多容量时，您可以部署另一个通用主机系统。

同样，扩展容器化应用程序也很简单。OpenShift Container Platform 提供了一种简单的、标准方式的容器化服务扩展功能。例如，如果将应用程序构建为一组微服务，而非大型的单体式应用程序，您可以分别扩展各个微服务来满足需求。有了这一能力，您可以只扩展需要的服务，而不是整个应用程序，从而在使用最少资源的前提下满足应用程序需求。

1.1.3. OpenShift Container Platform 概述

OpenShift Container Platform 为 Kubernetes 带来企业级增强，具体包括以下所列：

- 混合云部署。您可以将 OpenShift Container Platform 集群部署到各种公有云平台或数据中心中。
- 集成了红帽技术。OpenShift Container Platform 中的主要组件源自 Red Hat Enterprise Linux 和相关的红帽技术。OpenShift Container Platform 得益于红帽企业级优质软件的严格测试和认证计划。
- 开源开发模型。开发以开放方式完成，源代码可从公共软件存储库中获得。这种开放协作促进了快速创新和开发。

虽然 Kubernetes 擅长管理应用程序，但它并未指定或管理平台级要求或部署过程。强大而灵活的平台管理工具和流程是 OpenShift Container Platform 4.2 具备的重要优势。以下各节介绍 OpenShift Container Platform 的一些独特功能和优势。

1.1.3.1. 定制操作系统

OpenShift Container Platform 使用 Red Hat Enterprise Linux CoreOS (RHCOS)，这是一款面向容器的操作系统，结合了 CoreOS 和 Red Hat Atomic Host 操作系统的一些最佳特性和功能。RHCOS 是专门为从 OpenShift Container Platform 运行容器化应用程序而设计的，能够与新工具配合，提供快速安装、基于 Operator 的管理和简化的升级。

RHCOS 包括：

- Ignition，OpenShift Container Platform 将其用作首次启动系统配置来进行机器的初次上线和配置。
- CRI-O，Kubernetes 的原生容器运行时实现，可与操作系统紧密集成来提供高效和优化的 Kubernetes 体验。CRI-O，提供用于运行、停止和重启容器的工具。它完全取代了 OpenShift Container Platform 3 中使用的 Docker Container Engine。
- Kubelet，Kubernetes 的主要节点代理，负责启动和监视容器。

在 OpenShift Container Platform 4.2 中，所有 control plane 都需要使用 RHCOS。但 compute（计算）机器（也被称为 worker）可以使用 Red Hat Enterprise Linux (RHEL) 做为操作系统。如果选择使用 RHEL worker，与将 RHCOS 用于所有集群机器相比，您必须执行更多的系统维护。

1.1.3.2. 简化的安装和更新流程

使用 OpenShift Container Platform 4.2 时，如果您拥有具有正确权限的帐户，通过运行单个命令并提供几个值，就能在支持的云中部署生产集群。如果使用支持的平台，您还可以自定义云安装或在数据中心中安装集群。

对于将 RHCOS 用于所有机器的集群，不论是更新还是升级，OpenShift Container Platform 都是一个简单又高度自动化的流程。由于 OpenShift Container Platform 可以从中央 control plane 全面控制每台机器上运行的系统和服务（包括操作系统本身），因此升级被设计为一个自动事件。如果集群包含 RHEL worker 机器，则 control plane 可从简化的更新过程中受益，但您还需执行更多任务来升级 RHEL 机器。

1.1.3.3. 其他主要功能

Operator 既是 OpenShift Container Platform 4.2 代码库的基本单元，又是部署供应用程序使用的应用程序和软件组件的便捷方式。在 OpenShift Container Platform 中，Operator 可充当平台的基础，不再需要手动升级操作系统和 control plane 应用程序。OpenShift Container Platform Operator（如 Cluster Version Operator 和 Machine Config Operator）允许对这些关键组件进行简化的集群范围内管理。

Operator Lifecycle Manager (OLM) 和 OperatorHub 提供了相应的工具，可用于存储 Operator 并将其分发给开发和部署应用程序的人员。

Red Hat Quay Container Registry 是一个 Quay.io 容器 registry，为 OpenShift Container Platform 集群提供大多数容器镜像和 Operator。Quay.io 是 Red Hat Quay 的一个公共 registry 版本，可存储数百万镜像和标签。

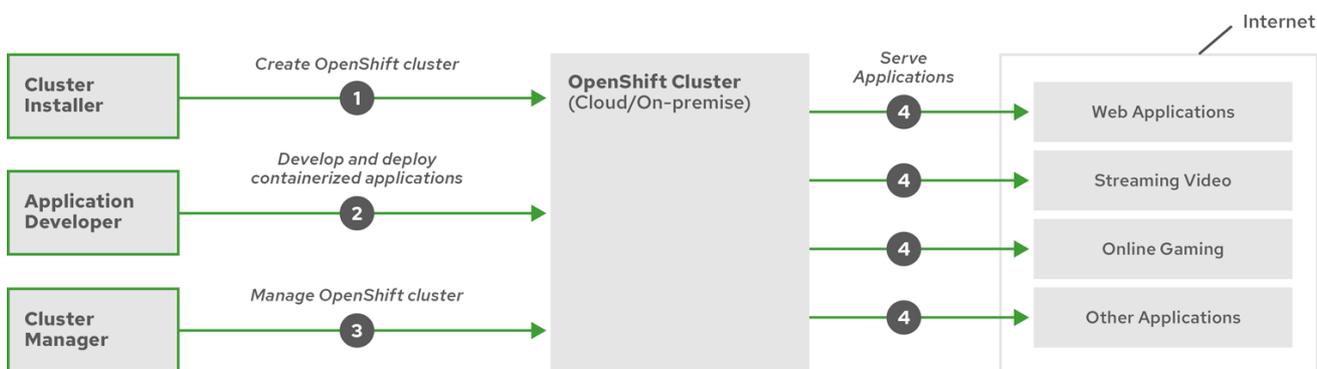
OpenShift Container Platform 中的其他 Kubernetes 增强功能包括软件定义网络 (SDN)、身份验证、日志聚合、监视和路由方面的改进。OpenShift Container Platform 还提供功能齐全的 web 控制台和自定义 OpenShift CLI (**oc**) 界面。

1.1.3.4. OpenShift Container Platform 生命周期

下图显示了 OpenShift Container Platform 的基本生命周期：

- 创建 OpenShift Container Platform 集群
- 管理集群
- 开发和部署应用程序
- 扩展应用程序

图 1.1. OpenShift Container Platform 高级概述



OpenShift_25_0519

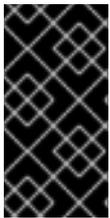
1.1.4. OpenShift Container Platform 对互联网和 Telemetry 的访问

在 OpenShift Container Platform 4.2 中，您需要访问互联网来安装集群。默认运行的 Telemetry 服务提供有关集群健康状况和成功更新的指标，这也需要访问互联网。如果您的集群连接到互联网，Telemetry 会自动运行，而且集群会注册到 [Red Hat OpenShift Cluster Manager](#)（OCM）。

确认 Red Hat OpenShift Cluster Manager 清单正确后，可以由 Telemetry 自动维护，也可以使用 OCM 手动维护，[使用订阅监控](#) 来跟踪帐户或多集群级别的 OpenShift Container Platform 订阅。

您必须具有以下互联网访问权限：

- 访问 [Red Hat OpenShift Cluster Manager](#) 页面，以下载安装程序并执行订阅管理。如果集群可以访问互联网，并且没有禁用 Telemetry，该服务会自动授权您的集群。
- 访问 [Quay.io](#)，以获取安装集群所需的软件包。
- 获取执行集群更新所需的软件包。



重要

如果您的集群无法直接访问互联网，则可以在置备的某些类基础架构上执行受限网络安装。在此过程中，您要下载所需的内容，并使用它在镜像 registry（mirror registry）中填充安装集群并生成安装程序所需的软件包。对于某些安装类型，集群要安装到的环境不需要访问互联网。在更新集群之前，您要更新镜像 registry（mirror registry）的内容。

第 2 章 安装和更新

2.1. OPENSIFT CONTAINER PLATFORM 安装概述

OpenShift Container Platform 安装程序为您提供了灵活性。您可以使用安装程序将集群部署到由安装程序置备并由集群维护的基础架构中，也可以将集群部署到您自己准备和维护的基础架构中。

这两种基本类型的 OpenShift Container Platform 集群通常称为安装程序置备的基础架构集群和用户置备的基础架构集群。

两种类型的集群都具有以下特征：

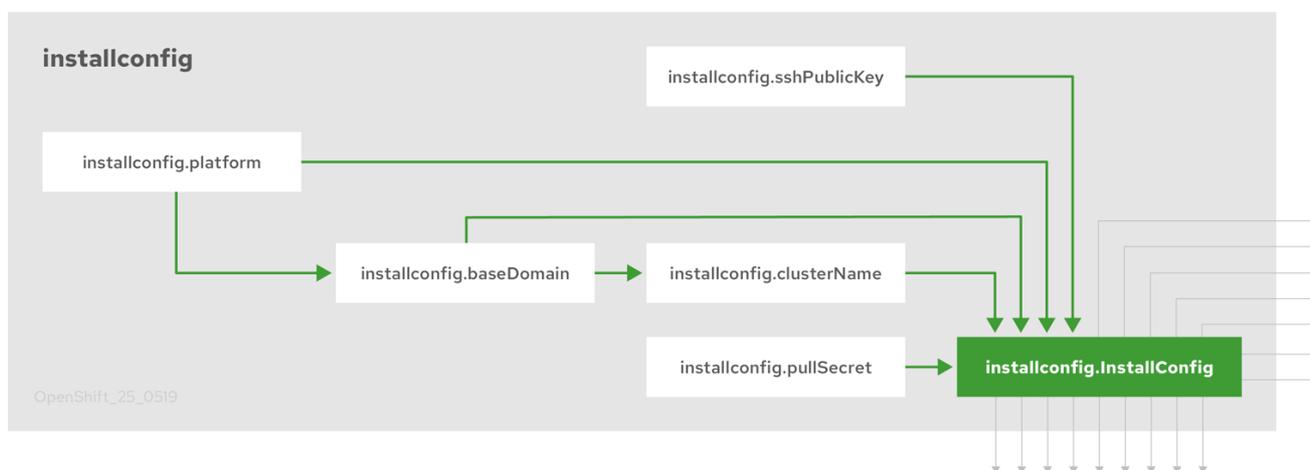
- 默认提供无单点故障的高可用性基础架构
- 管理员可以控制要应用的更新内容和更新的时间

两种类型的集群都使用同一个安装程序来部署。安装程序生成的主要资产是用于 Bootstrap、master 和 worker 机器的 Ignition 配置文件。有了这三个配置和配置得当的基础架构，就能启动 OpenShift Container Platform 集群。

OpenShift Container Platform 安装程序使用一组目标和依赖项来管理集群安装。安装程序具有一组必须实现的目标，并且每个目标都有一组依赖项。因为每个目标仅关注其自己的依赖项，所以安装程序可以采取并行实现多个目标。最终目标是正常运行的集群。通过满足依赖项而不是运行命令，安装程序能够识别和使用现有的组件，而不必运行命令来再次创建它们。

下图显示了安装目标和依赖项的子集：

图 2.1. OpenShift Container Platform 安装目标和依赖项



在安装后，每一个集群机器都将使用 Red Hat Enterprise Linux CoreOS (RHCOS) 作为操作系统。RHCOS 是 Red Hat Enterprise Linux (RHEL) 的不可变容器主机版本，具有默认启用 SELinux 的 RHEL 内核。它包括作为 Kubernetes 节点代理的 **kubelet**，以及为 Kubernetes 优化的 CRI-O 容器运行时。

OpenShift Container Platform 4.2 集群中的每一 control plane 机器都必须使用 RHCOS，其中包括一个关键的首次启动置备工具，称为 Ignition。这一工具让集群能够配置机器。操作系统更新作为嵌入在容器镜像中的 Atomic OSTree 存储库交付，该镜像由 Operator 在整个集群中推广。实际的操作系统的更改通过使用 rpm-ostree 在每台机器上作为原子操作原位进行。通过结合使用这些技术，OpenShift Container Platform 可以像管理集群上的任何其他应用程序一样管理操作系统，通过原位升级使整个平台保持最新状态。这些原位更新可以减轻运维团队的负担。

如果将 RHCOS 用作所有集群机器的操作系统，则集群将管理其组件和机器的所有方面，包括操作系统在内。因此，只有安装程序和 Machine Config Operator 才能更改机器。安装程序使用 Ignition 配置文件设置每台机器的确切状态，安装后则由 Machine Config Operator 完成对机器的更多更改，例如应用新证书或密钥等。

2.1.1. 适用的平台

在 OpenShift Container Platform 4.2 中，您可以在以下平台上安装使用安装程序置备的基础架构集群：

- Amazon Web Services (AWS)
- Google Cloud Platform (GCP)
- Microsoft Azure
- Red Hat OpenStack Platform 版本 13 和 15
 - 最新的 OpenShift Container Platform 版本支持最新的 RHOSP 长生命版本和中间版本。如需完整的 RHOSP 发行版本兼容性信息，请参阅 [RHOSP 上的 OpenShift Container Platform 支持列表](#)。
- Red Hat Virtualization (RHV)

对于所有这些集群，包括用来运行安装过程的计算机在内的所有机器都必须可直接访问互联网，以便为平台容器拉取镜像并向红帽提供 telemetry 数据。

在 OpenShift Container Platform 4.2 中，您可以在以下平台上安装使用用户置备的基础架构集群：

- AWS
- GCP
- VMware vSphere
- 裸机

在用户置备的基础架构上安装，每台机器都可拥有完整的互联网访问能力，您可以将集群放置在代理后面，也可以执行受限网络安装。在受限网络安装中，您可以下载安装集群所需的镜像（image），将它们放在镜像 registry（mirror registry）中，然后使用那些数据安装集群。虽然您需要访问互联网来为平台容器拉取镜像，但在 vSphere 或裸机基础架构上进行受限网络安装，您的集群机器不需要直接访问互联网。

[OpenShift Container Platform 4.x Tested Integrations](#) 页面中提供了有关针对不同平台进行集成测试的详细信息。

2.1.2. 安装过程

当安装 OpenShift Container Platform 集群时，从 Red Hat OpenShift Cluster Manager 站点上的相应的 [Infrastructure Provider](#) 页面下载安装程序。此网站管理以下内容：

- 帐户的 REST API
- registry 令牌，这是用于获取所需组件的 pull secret
- 集群注册，它将集群身份信息与您的红帽帐户相关联，以方便收集使用情况指标

在 OpenShift Container Platform 4.2 中，安装程序是对一组资产执行一系列文件转换的 Go 二进制文件。与安装程序交互的方式因您的安装类型而异。

- 对于具有安装程序置备的基础架构集群，您可以将基础架构启动和置备委派给安装程序，而不是亲自执行。安装程序将创建支持集群所需的所有网络、机器和操作系统。
- 如果亲自为集群置备和管理基础架构，则必须提供所有集群基础架构和资源，包括 Bootstrap 机器、网络、负载均衡、存储和独立的集群机器。您无法使用安装程序置备的基础架构集群所提供的高级机器管理和扩展功能。

安装期间使用三组文件：名为 **install-config.yaml** 的安装配置文件、Kubernetes 清单，以及您的机器类型适用的 Ignition 配置文件。

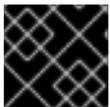


重要

安装期间可以修改控制基础 RHCOS 操作系统的 Kubernetes 和 Ignition 配置文件。但是，没有可用的验证机制来确认您对这些对象所做修改是适当的。如果修改了这些对象，集群可能会无法运行。由于存在这种风险，修改 Kubernetes 和 Ignition 配置文件不受支持，除非您遵循记录的流程或在红帽支持指示下操作。

安装配置文件转换为 Kubernetes 清单，然后清单嵌套到 Ignition 配置文件中。安装程序使用这些 Ignition 配置文件来创建集群。

运行安装程序时，所有配置文件会被修剪，因此请务必备份需要再次使用的所有配置文件。



重要

安装之后，您无法修改在安装过程中设置的参数，但可以修改一些集群属性。

采用安装程序置备的基础架构的安装过程

默认安装类型为使用安装程序置备的基础架构。默认情况下，安装程序充当安装向导，提示您输入它无法自行确定的值，并为其余参数提供合理的默认值。您还可以自定义安装过程来支持高级基础架构场景。安装程序将为集群置备底层基础架构。

您可以安装标准集群或自定义集群。对于标准集群，您要提供安装集群所需的最低限度详细信息。对于自定义集群，您可以指定有关平台的更多详细信息，如 control plane 使用的机器数量、集群部署的虚拟机的类型，或 Kubernetes 服务网络的 CIDR 范围。

若有可能，可以使用此功能来避免置备和维护集群基础架构。在所有其他环境中，可以使用安装程序来生成置备集群基础架构所需的资产。

对于安装程序置备的基础架构的集群，OpenShift Container Platform 可以管理集群的所有方面，包括操作系统本身。每台机器在启动时使用的配置引用其加入的集群中托管的资源。此配置允许集群在应用更新时自行管理。

采用用户置备的基础架构的安装过程

您还可以在自己提供的基础架构上安装 OpenShift Container Platform。您可以使用安装程序来生成置备集群基础架构所需的资产，再创建集群基础架构，然后将集群部署到您提供的基础架构中。

如果不使用安装程序置备的基础架构，您必须自己管理和维护集群资源，包括：

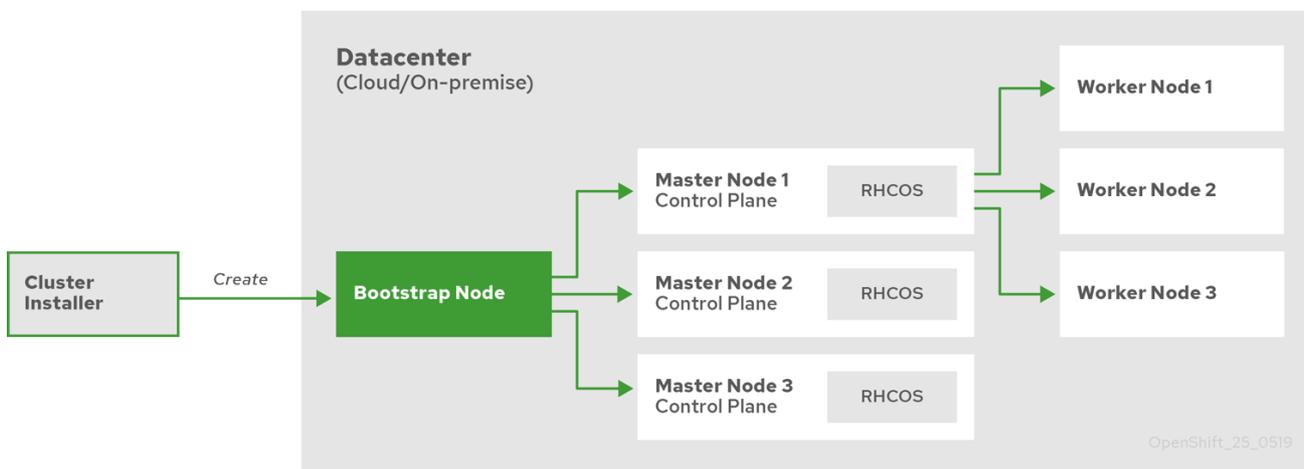
- 组成集群的 control plane 和计算机器
- 负载均衡器
- 集群网络，包括 DNS 记录和所需的子网
- 集群基础架构和应用程序的存储

如果您的集群使用用户置备的基础架构，则可以选择将 RHEL worker 机器添加到集群中。

安装过程详细信息

由于在置备时集群中的每台机器都需要集群的相关信息，因此 OpenShift Container Platform 在初始配置期间会使用临时 *Bootstrap* 机器将所需的信息提供给持久 control plane。通过使用描述如何创建集群的 Ignition 配置文件进行启动。Bootstrap 机器将创建组成 control plane 的 master 机器。然后，control plane 机器创建计算（compute）机器。下图说明了这一过程：

图 2.2. 创建 Bootstrap、master 和 worker 机器



集群机器初始化后，Bootstrap 机器将被销毁。所有集群都使用 Bootstrap 过程来初始化集群，但若您自己置备集群的基础架构，则必须手动完成许多步骤。



重要

安装程序生成的 Ignition 配置文件中所含的证书会在 24 小时后过期。您必须完成集群安装，并使集群以非降级状态运行 24 小时，以确保完成第一次证书轮转。

bootstrap 集群涉及以下步骤：

1. bootstrap 机器启动并开始托管 master 机器启动所需的远程资源。（如果自己配置基础架构，则需要人工干预）
2. master 机器从 bootstrap 机器获取远程资源并完成启动。（如果自己配置基础架构，则需要人工干预）
3. master 机器使用 bootstrap 机器来组成 etcd 集群。
4. bootstrap 机器使用新的 etcd 集群启动临时 Kubernetes control plane。
5. 临时 control plane 将生产环境的 control plane 调度到 master 机器。
6. 临时 control plane 关机，并将控制权交给生产环境 control plane。
7. bootstrap 机器将 OpenShift Container Platform 组件注入生产环境 control plane。
8. 安装程序关闭 bootstrap 机器。（如果自己配置基础架构，则需要人工干预）
9. control plane 设置 worker 节点。
10. control plane 以一组 Operator 的形式安装其他服务。

完成此 bootstrap 过程后，将生成一个全面运作的 OpenShift Container Platform 集群。然后，集群下载并配置日常运作所需的其余组件，包括在受支持的环境中创建 worker 机器。

安装过程详细信息

安装范围

OpenShift Container Platform 安装程序的作用范围特意设计得比较狭窄。它旨在简化操作并确保成功。安装完成后，您可以完成更多的配置任务。

2.2. 关于 OPENSIFT CONTAINER PLATFORM 更新服务

OpenShift Container Platform 更新服务是一种托管服务，为 OpenShift Container Platform 和 Red Hat Enterprise Linux CoreOS (RHCOS) 提供无线更新 (over-the air update)。它提供了一个组件 Operator 图，其中包含各个 *顶点* 及连接它们的 *边*。图中的边显示可以安全更新的版本，顶点是更新有效负载，用于指定托管集群组件的预期状态。

集群中的 Cluster Version Operator (CVO) 会检查 OpenShift Container Platform 更新服务，并根据当前组件版本和图中的信息决定有效的更新和更新路径。当您请求更新时，OpenShift Container Platform CVO 使用该更新的发行镜像来升级您的集群。发行工件 (artifact) 作为容器镜像托管在 Quay 中。

为了使 OpenShift Container Platform 更新服务仅提供兼容的更新，提供了一个版本验证管道来驱动自动化。每个发行工件都会被验证是否与支持的云平台 and 系统架构以及其他组件包兼容。在管道确认有适用的版本后，OpenShift Container Platform 更新服务会通知您可以进行更新。



重要

因为更新服务会显示所有有效的更新，所以不能强制更新到一个更新服务没有显示的版本。

对于连续更新模式，会运行两个控制器。一个控制器不断更新有效负载清单，将它们应用于集群，并输出受控 Operator 部署的状态（可用、正在进行升级或失败）。第二个控制器轮询 OpenShift Container Platform 更新服务以确定更新是否可用。



重要

不支持将集群还原到以前的版本或执行回滚。仅支持升级到较新版本。

2.3. 支持非受管 OPERATOR 的策略

Operator 的 *管理状态* 决定了一个 Operator 是否按设计积极管理集群中其相关组件的资源。如果 Operator 设置为 *非受管 (unmanaged)* 状态，它不会响应配置更改，也不会收到更新。

虽然它可以在非生产环境集群或调试过程中使用，但处于非受管状态的 Operator 不被正式支持，集群管理员需要完全掌控各个组件的配置和升级。

可使用以下方法将 Operator 设置为非受管状态：

- **独立 Operator 配置**

独立 Operator 的配置中具有 **managementState** 参数。这可以通过不同的方法来访问，具体取决于 Operator。例如，Cluster Logging Operator 通过修改它管理的自定义资源 (CR) 来达到此目的，而 Samples Operator 使用了集群范围配置资源。

将 **managementState** 参数更改为 **Unmanaged** 意味着 Operator 不会主动管理它的资源，也不会执行与相关组件相关的操作。一些 Operator 可能不支持此管理状态，因为它可能会损坏集群，需要手动恢复。



警告

将独立 Operator 更改为**非受管**状态会导致不支持该特定组件和功能。报告的问题必须在**受管 (Managed)** 状态中可以重复出现才能继续获得支持。

- **Cluster Version Operator (CVO) 覆盖**

可将 **spec.overrides** 参数添加到 CVO 配置中，以便管理员提供对组件的 CVO 行为覆盖的列表。将一个组件的 **spec.overrides[].unmanaged** 参数设置为 **true** 会阻止集群升级并在设置 CVO 覆盖后提醒管理员：

```
Disabling ownership via cluster version overrides prevents upgrades. Please remove overrides before continuing.
```



警告

设置 CVO 覆盖会使整个集群处于不受支持状态。在删除所有覆盖后，必须可以重现报告的问题方可获得支持。

第 3 章 OPENSIFT CONTAINER PLATFORM CONTROL PLANE

3.1. 了解 OPENSIFT CONTAINER PLATFORM CONTROL PLANE

control plane 由 master 机器组成，负责管理 OpenShift Container Platform 集群。control plane 机器管理计算机器（也被称为 worker）上的工作负载。集群本身通过 Cluster Version Operator、Machine Config Operator 和一组单独 Operator 的操作来管理对机器的所有升级。

3.1.1. OpenShift Container Platform 中的机器角色

OpenShift Container Platform 为主机分配不同的角色。这些角色定义机器在集群内的功能。集群包含标准 master 和 worker 角色类型的定义。



注意

集群还包含 bootstrap 角色的定义。由于 bootstrap 机器仅在集群安装期间使用，因此其功能在集群安装文档中阐述。

3.1.1.1. 集群 worker

在 Kubernetes 集群中，worker 节点是运行和管理 Kubernetes 用户请求的实际工作负载的地方。worker 节点公告其容量，而作为 master 服务一部分的调度程序决定在哪些节点上启动容器和 Pod。重要服务在每个 worker 节点上运行，包括 CRI-O（即容器引擎）、Kubelet（接受并履行运行和停止容器工作负载请求的服务），以及服务代理（管理 worker 之间 Pod 的通信）。

在 OpenShift Container Platform 中，MachineSets 控制 worker 机器。具有 worker 角色的机器驱动计算工作负载，这些负载由自动扩展它们的特定机器池管控。因为 OpenShift Container Platform 具有支持多种机器类型的能力，因此 worker 机器被归类为 *计算 (compute)* 机器。在此版本中，术语“worker 机器”和“计算机器”可互换使用，因为计算机器的唯一默认类型就是 worker 机器。在未来的 OpenShift Container Platform 版本中，默认情况下可能会使用不同类型的计算机器，如基础架构机器。

3.1.1.2. 集群 master

在 Kubernetes 集群中，master 节点运行控制 Kubernetes 集群所需的服务。在 OpenShift Container Platform 中，master 机器是 control plane。它们不仅仅包含用于管理 OpenShift Container Platform 集群的 Kubernetes 服务。因为所有具有 control plane 角色的机器都是 master 机器，所以 *master* 和 *control plane* 是可以互换的术语。master 机器不会被分成一个 MachineSet，而是由一系列独立的机器 API 资源定义。附加控件应用到 master 机器，以防止您删除所有 master 机器并破坏集群。



注意

使用三个 master 节点虽然理论上可以使用任意数量的 master 节点，但由于 master 静态 Pod 和 etcd 静态 Pod 在相同的主机上工作，所以这个数量会受 etcd 仲裁的限制。

master 上属于 Kubernetes 类别的服务包括 API 服务器、etcd、Controller Manager Server 和 HAProxy 服务。

表 3.1. 在 control plane 上运行的 Kubernetes 服务

组件	描述
API Server	Kubernetes API 服务器验证并配置 Pod、服务和复制控制器的数据。它还为集群的共享状态提供一个焦点。
etcd	etcd 存储持久 master 状态，其他组件则监视 etcd 的更改，以使其自身进入指定状态。
Controller Manager Server	Controller Manager Server 监视 etcd 中对象的更改，如复制、命名空间和服务帐户控制器对象，然后使用 API 强制执行指定的状态。多个这样的过程会创建在某个时间点上有一个活跃群首的集群。

master 机器上的某些服务作为 systemd 服务运行，其他服务则作为静态 Pod 运行。

systemd 服务适合需要始终在特定系统启动后不久出现的服务。对于 master 机器，这包括允许远程登录的 sshd。它还包括以下服务：

- CRI-O 容器引擎 (crio)，用于运行和管理容器。OpenShift Container Platform 4.2 使用 CRI-O，而不是 Docker Container Engine。
- Kubelet (kubelet)，从 master 服务接受管理机器上容器的请求。

CRI-O 和 Kubelet 必须作为 systemd 服务直接在主机上运行，因为它们必须先运行，然后您才能运行其他容器。

3.1.2. OpenShift Container Platform 中的 Operator

在 OpenShift Container Platform 中，Operator 是在 control plane 上打包、部署和管理服务的首选方法。它们还为用户运行的应用程序提供了便利。Operator 与 Kubernetes API 和 CLI 工具（如 **kubectl** 和 **oc** 命令）集成。它们提供了各种方式，以监视应用程序、执行健康检查、管理无线更新，以及确保应用程序保持在指定的状态。

因为 CRI-O 和 Kubelet 在每个节点上运行，所以几乎所有其他集群功能都可以通过使用 Operator 在 control plane 上进行管理。Operator 是 OpenShift Container Platform 4.2 中最重要的组件。使用 Operator 添加到 control plane 的组件包括重要的网络服务和凭证服务。

在 OpenShift Container Platform 集群中管理其他 Operator 的 Operator 是 Cluster Version Operator。

OpenShift Container Platform 4.2 使用不同类型的 Operator 来执行集群操作，并在集群上运行各种服务供您的应用程序使用。

3.1.2.1. OpenShift Container Platform 中的平台 Operator

在 OpenShift Container Platform 4.2 中，所有集群功能都划分到一系列平台 Operator 中。平台 Operator 管理集群功能的特定方面，如集群范围的应用程序日志记录、Kubernetes control plane 管理或机器置备系统。

每个 Operator 都为您提供确定集群功能的简单 API。Operator 将管理组件生命周期的细节隐藏起来。Operator 可以管理一个组件或数十个组件，但最终目标始终是通过自动化常见操作来减轻运维负担。

Operator 还提供了更为精细的配置体验。若要配置各个组件，您可以修改 Operator 公开的 API，而不必修改全局配置文件。

3.1.2.2. 由 OLM 管理的 Operator

Cluster Operator Lifecycle Management (OLM) 组件管理可在应用程序中使用的 Operator。它不管理组成 OpenShift Container Platform 的 Operator。OLM 是一个将 Kubernetes 原生应用程序作为 Operator 进行管理的框架。它不管理 Kubernetes 清单，而是管理 Kubernetes Operator。OLM 管理两种 Operator，即 Red Hat Operator 和经认证的 Operator。

一些 Red Hat Operator 用来提供集群功能，如调度程序和问题检测器。其他 Operator 则可供您自助管理并在应用程序中使用，例如 etcd。OpenShift Container Platform 还提供由社区构建和维护并经过认证的 Operator。这些经过认证的 Operator 为传统应用程序提供 API 层，因此您可以通过 Kubernetes 构造来管理应用程序。

3.1.2.3. 关于 OpenShift Container Platform 更新服务

OpenShift Container Platform 更新服务是一种托管服务，为 OpenShift Container Platform 和 Red Hat Enterprise Linux CoreOS (RHCOS) 提供无线更新 (over-the air update)。它提供了一个组件 Operator 图，其中包含各个 *顶点* 及连接它们的 *边*。图中的边显示可以安全更新的版本，顶点是更新有效负载，用于指定托管集群组件的预期状态。

集群中的 Cluster Version Operator (CVO) 会检查 OpenShift Container Platform 更新服务，并根据当前组件版本和图中的信息决定有效的更新和更新路径。当您请求更新时，OpenShift Container Platform CVO 使用该更新的发行镜像来升级您的集群。发行工件 (artifact) 作为容器镜像托管在 Quay 中。

为了使 OpenShift Container Platform 更新服务仅提供兼容的更新，提供了一个版本验证管道来驱动自动化。每个发行工件都会被验证是否与支持的云平台 and 系统架构以及其他组件包兼容。在管道确认有适用的版本后，OpenShift Container Platform 更新服务会通知您可以进行更新。



重要

因为更新服务会显示所有有效的更新，所以不能强制更新到一个更新服务没有显示的版本。

对于连续更新模式，会运行两个控制器。一个控制器不断更新有效负载清单，将它们应用于集群，并输出受控 Operator 部署的状态（可用、正在进行升级或失败）。第二个控制器轮询 OpenShift Container Platform 更新服务以确定更新是否可用。



重要

不支持将集群还原到以前的版本或执行回滚。仅支持升级到较新版本。

3.1.2.4. 了解 Machine Config Operator

OpenShift Container Platform 4.2 集成了操作系统和集群管理。由于集群管理自己的更新，包括集群节点上 Red Hat Enterprise Linux CoreOS (RHCOS) 的更新，因此 OpenShift Container Platform 提供了可靠的生命周期管理体验，能够简化节点升级的编配。

OpenShift Container Platform 使用三个 DaemonSet 和控制器来简化节点管理。这些 DaemonSet 通过使用标准的 Kubernetes 式构造来编配操作系统更新和主机配置更改。它们包括：

- **machine-config-controller**，协调从 control plane 进行的机器升级。它监控所有集群节点并编配其配置更新。

- **machine-config-daemon** DaemonSet, 在集群中的每个节点上运行, 并按照 MachineConfigController 的指示将机器更新为 MachineConfig 定义的配置。当节点看到更改时, 它将排空其 Pod, 应用更新并重启。这些更改以 Ignition 配置文件的形式出现, 这些文件应用指定的机器配置并控制 kubelet 配置。更新本身在容器中交付。此过程是成功管理 OpenShift Container Platform 和 RHCOS 更新的关键。
- **machine-config-server** DaemonSet, 在 master 节点加入集群时为其提供 Ignition 配置文件。

机器配置是 Ignition 配置的子集。**machine-config-daemon** 读取机器配置, 以查看是否需要进行 OSTree 更新, 或者是否必须应用一系列 systemd kubelet 文件更改、配置更改, 或者对操作系统或 OpenShift Container Platform 配置的其他更改。

执行节点管理操作时, 您可以创建或修改 KubeletConfig Custom Resource (CR)。

第 4 章 了解 OPENSIFT CONTAINER PLATFORM 开发

为了在开发和运行企业级品质应用程序时充分利用容器的功能，请确保您的环境受相应工具的支持，让容器能够：

- 创建为离散的服务，可以连接到其他容器化和非容器化服务。例如，您可能希望将应用程序与数据库衔接，或将监控应用程序附加到数据库。
- 具有弹性，因此在服务器崩溃或需要停机维护或退役时，容器可以在另一台机器上启动。
- 实现自动化，以自动获取代码更改，然后启动和部署自身的新版本。
- 得以扩展或复制，在需求增加时为客户端提供更多实例，在需求下降时缩减为更少的实例。
- 以不同的方式运行，具体由应用程序的类型决定。例如，一个应用程序可能每月运行一次来生成报告，然后退出。另一个应用程序可能需要持续运行，并且必须对客户端高度可用。
- 受到管理，以便您可以监视应用程序的状态并在出现问题时做出反应。

容器得到广泛接受，对能让容器适合企业使用的工具和方法的需求也随之诞生，这使得容器有了丰富的选择。

本节的其余部分介绍在 OpenShift Container Platform 中构建和部署容器化 Kubernetes 应用程序时可以选择创建的资产。它还说明您可以采用哪些方法来满足不同类型的应用程序和开发需求。

4.1. 关于容器化应用程序开发

您可以通过多种方式使用容器来进行应用程序开发，每种方法更适合的使用情景也各不相同。为了说明这种多样性，本文在介绍一系列方法时首先从开发单个容器开始，最后将容器部署为面向大型企业的任务关键型应用程序。这些方法展示了不同的工具、格式和方法，供您用于容器化应用程序开发。本主题描述：

- 构建一个简单容器并将其存储在 registry 中
- 创建 Kubernetes 清单并将其保存到 Git 存储库
- 使 Operator 能够与其他项共享您的应用程序

4.2. 构建一个简单容器

您对应用程序有了一个想法，想要对其进行容器化。

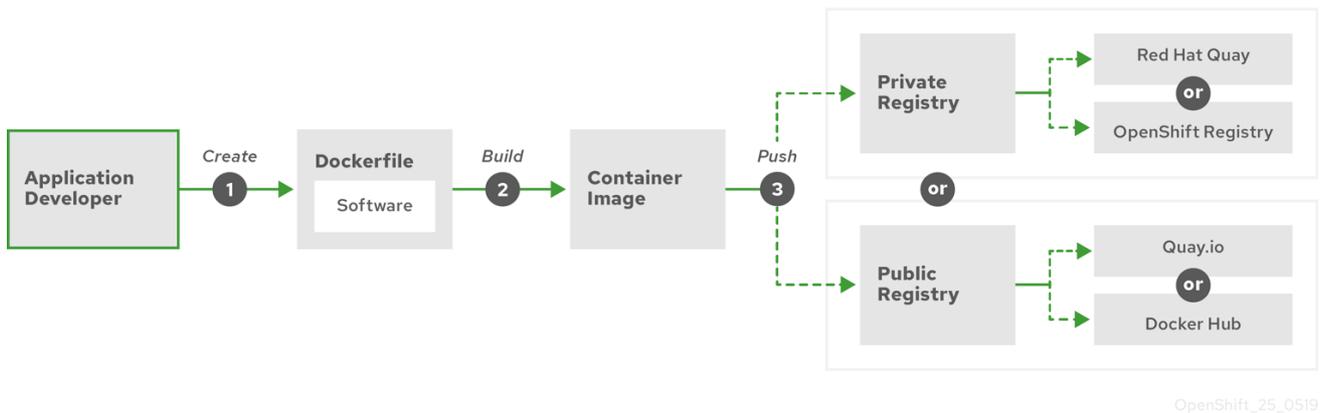
首先，您需要一个用于构建容器的工具，如 buildah 或 docker，还需要一个描述容器中内容的文件，通常是 [Dockerfile](#)。

接下来，您需要一个位置来推送生成的容器镜像，以便可以将它拉取到您想要它运行的位置。这个位置就是容器 registry。

大多数 Linux 操作系统上都会默认安装各个组件的一些示例，但 Dockerfile 除外，需要您自己提供。

下图显示了构建和推送镜像的流程：

图 4.1. 创建简单容器化应用程序并将其推送到 registry



OpenShift_25_0519

如果您使用运行 Red Hat Enterprise Linux (RHEL) 作为操作系统的计算机，则容器化应用程序创建过程需要以下几个步骤：

1. 安装容器构建工具：RHEL 包含一组工具，其中包括用于构建和管理容器的 podman、buildah 和 skopeo。
2. 创建一个 Dockerfile 来组合基础镜像和软件：有关构建容器的信息存放在名为 **Dockerfile** 的文件中。在这个文件中，您要标识从中构建的基本镜像、要安装的软件包，以及要复制到容器中的软件。您还要标识参数值，如公开到容器外部的网络端口和挂载到容器内的卷。将您的 Dockerfile 和要容器化的软件放在 RHEL 系统上的某一目录中。
3. 运行 buildah 或 docker 构建：运行 **buildah build-using-dockerfile** 或 **docker build** 命令，将您选择的基础镜像拉取到本地系统，再创建一个存储在本地的容器镜像。您还可以使用 buildah，在不用 Dockerfile 的前提下构建容器。
4. 标记 (tag) 并推送到 registry：向新容器镜像添加标签 (tag)，以标识要在其中存储和共享容器的 registry 位置。然后，通过运行 **podman push** 或 **docker push** 命令将该镜像推送到 registry。
5. 拉取并运行镜像：从具有 podman 或 docker 等容器客户端工具的任何系统，运行用于标识新镜像的命令。例如，运行 **podman run <image_name>** 或 **docker run <image_name>** 命令。其中，**<image_name>** 是新容器镜像的名称，类似于 **quay.io/myrepo/myapp:latest**。registry 可能需要凭证才能推送和拉取镜像。

如需有关如何构建容器镜像、将其推送到 registry 并运行它们的过程的更多详细信息，请参阅[使用 Buildah 自定义镜像构建](#)。

4.2.1. 容器构建工具选项

尽管 Docker Container Engine 和 **docker** 命令是操作容器的常用工具，但在 RHEL 和许多其他 Linux 系统上，您也可以选择另一组容器工具，包括 podman、skopeo 和 buildah 等。您仍然可以使用 Docker Container Engine 来创建容器，这些容器将在 OpenShift Container Platform 和任何其他容器平台中运行。

使用 buildah、podman 和 skopeo 构建和管理容器会生成行业标准的容器镜像，其包含经过专门调优的功能，以便于这些容器最终部署到 OpenShift Container Platform 或其他 Kubernetes 环境中。这些工具不是作为守护进程运行的，可在没有 root 权限的情况下运行，因此运行它们的开销比较小。

最终在 OpenShift Container Platform 中运行容器时，您要使用 [CRI-O](#) 容器引擎。CRI-O 在 OpenShift Container Platform 集群中的每台 worker 机器和 master 机器上运行，但 CRI-O 尚不支持作为 OpenShift Container Platform 外的独立运行时。

4.2.2. 基础镜像选项

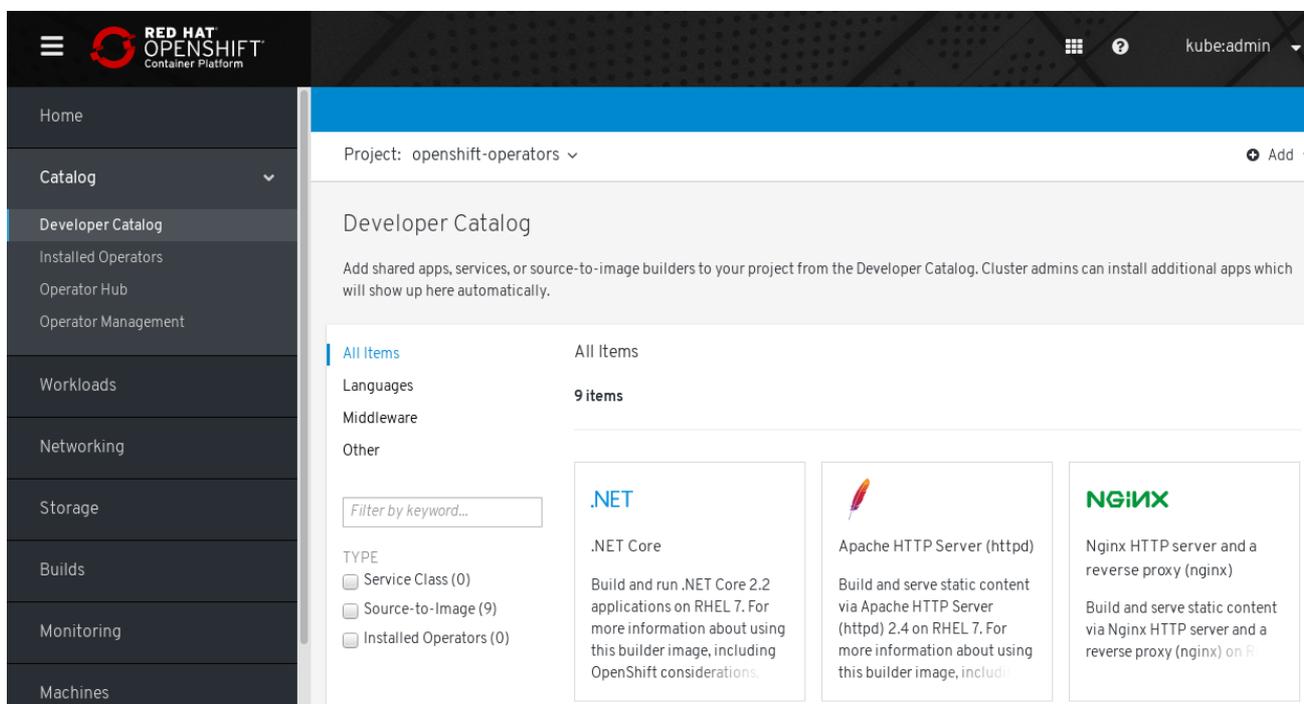
选择用来构建应用程序的基础镜像包含一组软件，这些软件为应用程序提供一个 Linux 系统。在您构建自己的镜像时，您的软件将放置到该文件系统中，可以像对待操作系统一样看待该文件系统。基础镜像的选择会对容器未来的安全性、效率和可升级性产生重大影响。

红帽提供了一组新的基础镜像，称为**红帽通用基础镜像 (UBI)**。这些镜像基于 Red Hat Enterprise Linux，与红帽过去提供的基础镜像相似，但有一个主要区别：无需红帽订阅就能自由重新分发。因此，您可以在 UBI 镜像上构建应用程序，不必担心如何共享它们或需要为不同的环境创建不同的镜像。

这些 UBI 镜像具有标准、初始和最低版本。您还可以将 [Red Hat Software Collections](#) 镜像用作依赖特定运行时环境（如 Node.js、Perl 或 Python）的应用程序的基础。其中一些运行时基础镜像的特殊版本称为 Source-to-Image (S2I) 镜像。使用 S2I 镜像时，您可以将代码插入到可随时运行该代码的基础镜像环境中。

您可以直接从 OpenShift Container Platform Web UI 中使用 S2I 镜像，只需选择 **Catalog → Developer Catalog** 便可，如下图中所示：

图 4.2. 为需要特定运行时的应用选择 S2I 基础镜像



4.2.3. registry 选项

容器 registry 是存储容器镜像的位置，便于与他人共享，并提供给它们最终运行的平台。您可以选择提供免费帐户的大型公共容器 registry，也可选择提供更多存储和特殊功能的高级版本。您还可以安装自己的 registry，供您的组织专用或有选择地共享给他人。

要获取红帽和认证合作伙伴提供的镜像，您可以从 Red Hat Registry 中提取。Red Hat Registry 存在于两个位置：[registry.access.redhat.com](#)（无需身份验证，但已弃用）和 [registry.redhat.io](#)（需要身份验证）。您可从 [Red Hat Container Catalog](#) 了解 Red Hat Registry 中由红帽和合作伙伴提供的镜像。除了列出红帽容器镜像外，它还显示有关这些镜像的内容和质量广泛信息，包括基于已应用安全更新的健康分数。

大型公共 registry 包括 [Docker Hub](#) 和 [Quay.io](#)。Quay.io registry 由红帽所有和管理。OpenShift Container Platform 中使用的许多组件都存储在 Quay.io 中，包括用于部署 OpenShift Container Platform 本身的容器镜像和 Operator。Quay.io 还提供了存储其他类型内容的方法，包括 Helm Charts。

如果需要私有容器 registry，OpenShift Container Platform 本身包括一个私有容器 registry，随 OpenShift Container Platform 一起安装并在其集群上运行。红帽也提供 Quay.io registry 的私有版本，称为 [Red Hat Quay](#)。Red Hat Quay 包括跨地域复制、Git 构建触发器、Clair 镜像扫描和许多其他功能。

此处提到的所有 registry 都可能需要凭证才能从中下载镜像。这些凭证中有些是通过 OpenShift Container Platform 在整个集群范围内提供的，另一些凭证则可以分配给个人。

4.3. 为 OPENSIFT CONTAINER PLATFORM 创建 KUBERNETES 清单

尽管容器镜像是容器化应用程序的基本构建块，但需要更多信息才能在 Kubernetes 环境（如 OpenShift Container Platform）中管理和部署该应用程序。创建镜像后的典型后续步骤：

- 了解 Kubernetes 清单中使用的不同资源
- 就您将运行的应用程序做出一些决策
- 收集支持组件
- 创建清单并将该清单存储到 Git 存储库中，以便您可以将其存储在源版本控制系统中，对其进行审核、跟踪和升级，并将其部署到下一环境中，在必要时回滚到旧版本，以及与他人共享

4.3.1. 关于 Kubernetes Pod 和服务

容器镜像是 docker 的基本单元，而 Kubernetes 使用的基本单元称为 [Pod](#)。Pod 代表构建应用程序的下一步。一个 Pod 可以包含一个或多个容器。关键之处在于，Pod 是您部署、扩展和管理的单个单元。

在决定 Pod 中要放入的内容时，可扩展性和命名空间或许是要考虑的主要项目。为便于部署，您可能需要将容器部署到 Pod 中，并在 Pod 中包含其本身的日志记录和监控容器。以后，在您运行 Pod 并需要扩展额外的实例时，其他那些容器也会随之扩展。对于命名空间，Pod 中的容器共享相同的网络接口、共享存储卷和资源限制（如内存和 CPU）。这样一来，将 Pod 内容作为一个单元进行管理变得更加轻松。Pod 中的容器还可以使用标准的进程间通信（如 System V 信号或 POSIX 共享内存）相互通信。

虽然单个 Pod 代表 Kubernetes 中的一个可扩展单元，但[服务](#)提供了一个途径，能够将一系列 Pod 分组到一起以创造完整且稳定的应用程序，完成诸如负载均衡之类的任务。服务也比 Pod 更持久，因为服务可以从同一 IP 地址使用，直到您删除为止。在使用服务时，可通过名称来请求服务，OpenShift Container Platform 集群则将该名称解析为您可用于访问构成该服务的 Pod 的 IP 地址和端口。

从本质上讲，容器化应用程序与运行时所处的操作系统是隔开的，进而与其用户隔开。Kubernetes 清单的一部分描述了如何通过定义[网络策略](#)将应用程序公开给内部和外部网络，对您的容器化应用的通信进行精细的控制。要将来自集群外部的 HTTP、HTTPS 和其他服务的传入请求连接到集群内部的服务，可以使用 [Ingress](#) 资源。

如果容器需要磁盘存储而不是数据库存储（可以通过服务提供），则可以将[卷](#)添加到清单中，使该存储可供 Pod 使用。您可以配置清单以创建持久性卷 (PV)，或动态创建添加到 Pod 定义中的卷。

在确定了组成应用程序的一组 Pod 后，您可以在 [deployments](#) 和 [deploymentconfigs](#) 中定义这些 Pod。

4.3.2. 应用程序类型

接下来，考虑您的应用程序类型对其运行方式的影响。

Kubernetes 定义了适用于不同类型应用程序的不同工作负载。要确定适合应用程序的工作负载，请考虑该应用程序是否：

- 运行结束后即告完成。例如，启动某个应用程序以生成报告，并在报告完成时退出。之后一个月内可能不会再次运行这个应用程序。对于这些类型的应用程序，合适的 OpenShift Container Platform 对象包括 [Job](#) 和 [CronJob](#) 对象。
- 预计将持续运行。对于长期运行的应用程序，您可以编写 [Deployment](#) 或 [DeploymentConfig](#)。
- 需要高度可用。如果应用程序需要高可用性，那么您需要调整部署的大小，使其包含不止一个实例。[Deployment](#) 或 [DeploymentConfig](#) 可以融合适于该类应用程序的 [ReplicaSet](#)。借助 [ReplicaSet](#)，Pod 可以跨越多个节点运行，确保即使在 worker 中断时该应用程序也始终可用。
- 需要在每个节点上运行。某些类型的 Kubernetes 应用程序设计为在集群中的每个 master 节点或 worker 节点上运行。例如，DNS 和监控应用程序需要在每个节点上持续运行。您可以将这类应用程序作为 [DaemonSet](#) 运行。您还可以基于节点标签 (label)，在节点的一个子集上运行 [DaemonSet](#)。
- 需要生命周期管理。当您要移交应用程序供其他人使用时，请考虑创建 [Operator](#)。Operator 可帮助您构建智能功能，自动处理备份和升级之类的事务。与 [Operator Lifecycle Manager \(OLM\)](#) 相结合，集群管理器可以将 Operator 公开给选定命名空间，以便集群中的用户可以运行它们。
- 具有标识或编号要求。应用程序可能具有标识或编号要求。例如，您可能需要运行应用程序的三个实例，并将这些实例命名为 **0**、**1** 和 **2**。[StatefulSet](#) 适合此应用程序。[StatefulSet](#) 对于需要独立存储的应用程序最有用处，如数据库和 zookeeper 集群。

4.3.3. 可用的支持组件

您编写的应用程序可能需要支持组件，如数据库或日志记录组件。为满足这一需求，可以从 OpenShift Container Platform Web 控制台中提供的以下目录获取所需的组件：

- [OperatorHub](#)，可在每个 OpenShift Container Platform 4.2 集群中使用。借助 [OperatorHub](#)，集群操作员可以使用来自红帽、红帽认证合作伙伴和社区成员的 Operator。集群操作员可以在集群中的所有命名空间或选定命名空间中提供这些 Operator，让开发人员能够通过他们的应用程序启动并配置这些 Operator。
- 服务目录，提供 Operator 的替代方案。尽管部署 Operator 是在 OpenShift Container Platform 中获取打包应用程序的首选方法，但出于某些原因，您可能想要使用服务目录来获取自己应用程序的支持应用程序。如果您是现有的 OpenShift Container Platform 3 客户并且已投入于服务目录应用程序，或者您已拥有 Cloud Foundry 环境，并且您有兴趣使用来自其他生态系统的代理，则可能要使用服务目录。
- 模板，对于一次性类型的应用程序很有用。在该应用程序中，组件的生命周期在安装后并不重要。模板提供了一种简便方式，可以从最小的开销开始开发 Kubernetes 应用程序。模板可以是资源定义的列表，这包括部署、服务、路由或其他对象。如果要更改名称或资源，可通过参数形式在模板中设置这些值。[Template Service Broker Operator](#) 是一个服务代理，可供您用于实例化自己的模板。您也可以直接从命令行安装模板。

您可以根据开发团队的特定需求，配置支持的 Operator、服务目录应用程序和模板，然后在开发人员开展工作的命名空间中提供它们。许多人将共享模板添加到 **openshift** 命名空间中，因为可以从所有其他命名空间访问这个命名空间。

4.3.4. 应用清单

借助 Kubernetes 清单 (manifest)，您可以更加全面地了解组成 Kubernetes 应用程序的组件。您可以将这些清单编写为 YAML 文件，并通过应用到集群来部署它们，例如通过运行 **oc apply** 命令。

4.3.5. 后续步骤

此时，请考虑对容器开发过程进行自动化的方法。理想情况下，您可以使用某种 CI 管道来构建镜像并将其推送到 registry。特别是，GitOps 管道可将容器开发与 Git 存储库集成在一起，您将使用 Git 存储库来存储构建应用程序所需的软件。

到目前为止的工作流程可能如下所示：

- 第 1 天：编写一些 YAML。然后，运行 **oc apply** 命令将 YAML 应用于集群并测试其是否正常工作。
- 第 2 天：将 YAML 容器配置文件放进您的 Git 存储库中。想要安装该应用或协助您改进的人可以从那里拉取 YAML，并应用到他们用于运行应用程序的集群中。
- 第 3 天：考虑为应用程序编写 Operator。

4.4. 面向 OPERATOR 进行开发

如果您的应用程序要提供给他人运行，最好将其打包并部署为 Operator。如前文所述，Operator 向您的应用程序添加了一个生命周期组件。使用它可以实现在安装应用程序后需要进行的维护任务。

在将应用程序创建为 Operator 时，您可以纳入自己有关如何运行和维护应用程序的知识。您可以内置用来升级、备份、扩展应用程序或跟踪其状态的功能。正确配置应用程序后，维护任务（如更新 Operator）可以自动发生，并且不为 Operator 用户所见。

例如，设置为在特定时间自动备份数据的 Operator 就非常实用。让 Operator 在设定的时间管理应用程序备份，可以使系统管理员免于记忆这些事务。

传统上手动完成的任何应用程序维护（如备份数据或轮转证书）都可以借助 Operator 自动完成。

第 5 章 RED HAT ENTERPRISE LINUX COREOS (RHCOS)

5.1. 关于 RHCOS

Red Hat Enterprise Linux CoreOS (RHCOS) 代表了下一代单用途容器操作系统技术。RHCOS 由创建了 Red Hat Enterprise Linux Atomic Host 和 CoreOS Container Linux 的同一开发团队打造，它将 Red Hat Enterprise Linux (RHEL) 的质量标准与 Container Linux 的自动化远程升级功能结合在一起。

对于所有 OpenShift Container Platform 机器，仅支持将 RHCOS 作为 OpenShift Container Platform 4.2 的组件。RHCOS 是唯一受 OpenShift Container Platform control plane 或 master 机器支持的操作系统。虽然 RHCOS 是所有集群机器的默认操作系统，但您仍可以创建使用 RHEL 作为其操作系统的计算（compute）机器（也称为 worker）。

如果将集群安装到集群置备的基础架构上，则要在安装过程中将 RHCOS 镜像下载到目标平台，并且使用控制 RHCOS 配置的合适 Ignition 配置文件来部署机器。如果将集群安装到您自己管理的基础架构上，则必须遵循安装文档来获取 RHCOS 镜像，生成 Ignition 配置文件，并且使用 Ignition 配置文件来置备机器。

5.1.1. RHCOS 主要功能

下表描述了 RHCOS 操作系统的主要功能：

底层操作系统主要由 RHEL 组件构成。支持 RHEL 的相同质量、安全性和控制措施也支持 RHCOS。例如，RHCOS 软件位于 RPM 软件包中，并且每个 RHCOS 系统都以 RHEL 内核以及由 systemd 初始化系统管理的一组服务启动。

尽管 RHCOS 包含 RHEL 组件，但它的管理要比默认的 RHEL 安装更加严格。管理从 OpenShift Container Platform 集群远程执行。设置 RHCOS 机器时，您只能修改一些系统设置。这种受控的不变性使 OpenShift Container Platform 可以存储集群中 RHCOS 系统的最新状态，因而始终都能创建额外的机器并根据最新的 RHCOS 配置执行更新。

尽管 RHCOS 包含运行 Docker 所需的 OCI 和 libcontainer 格式容器的功能，但它融合的是 CRI-O 容器引擎，而非 Docker 容器引擎。通过专注于 Kubernetes 平台（例如 OpenShift Container Platform）所需的功能，CRI-O 可以提供与不同 Kubernetes 版本兼容的特定功能。与提供更大功能集的容器引擎相比，CRI-O 对内存的要求更低，且对安全的攻击面更小。目前，CRI-O 仅可用作 OpenShift Container Platform 集群中的容器引擎。

对于诸如构建、复制和以其他方式管理容器的任务，RHCOS 用一组兼容的容器工具来代替 Docker CLI 工具。podman CLI 工具支持许多容器运行时功能，例如运行、启动、停止、列举和删除容器及容器镜像。skopeo CLI 工具可以复制、身份认证和签名镜像。您可以使用 crictl CLI 工具来处理 CRI-O 容器引擎中的容器和 Pod。虽然不建议在 RHCOS 中直接使用这些工具，但可以把它们用于调试目的。

RHCOS 具有使用 **rpm-ostree** 系统进行事务升级的功能。更新是通过容器镜像交付的，并且是 OpenShift 更新过程的一部分。部署之后，拉取、提取容器镜像并将其写入磁盘，然后修改启动加载程序以启动到新版本。机器将以滚动方式重启并进入更新，确保对集群容量的影响最小。

对于 RHCOS 系统，**rpm-ostree** 文件系统的布局具有以下特征：

- **/usr** 是操作系统二进制文件和库的存储位置，并且是只读的。我们不支持更改此设置。
- **/etc**、**/boot** 和 **/var** 在系统上是可写的，但只能由 Machine Config Operator 更改。
- **/var/lib/containers** 是用于存储容器镜像的图形存储位置。

在 OpenShift Container Platform 中，Machine Config Operator 处理操作系统升级。**rpm-ostree** 以原子

单元形式提供升级，不像 **yum** 升级那样单独升级各个软件包。新的操作系统部署在升级过程中进行，并在下次重启时才会生效。如果升级出现问题，则进行一次回滚并重启就能使系统返回到以前的状态。OpenShift Container Platform 中的 RHCOS 升级是在集群更新期间执行的。

5.1.2. 在 OpenShift Container Platform 中配置 RHCOS

对于 OpenShift Container Platform，RHCOS 镜像最初使用称为 **Ignition** 的功能进行设置，该功能仅在系统首次启动时运行。首次启动后，RHCOS 系统由 OpenShift Container Platform 集群中运行的 Machine Config Operator (MCO) 进行管理。

由于 OpenShift Container Platform 中的 RHCOS 系统设计为可从 OpenShift Container Platform 集群中进行全面管理，因此不建议直接登录 RHCOS 机器。出于调试目的，可以对 OpenShift Container Platform 集群中的 RHCOS 机器进行有限的直接访问。

5.1.3. 关于 Ignition

Ignition 是 RHCOS 在初始配置期间用于操作磁盘的实用程序。它可完成常见的磁盘任务，如分区磁盘、格式化分区、写入文件和配置用户等。首次启动时，Ignition 从安装介质或您指定的位置读取其配置，并将配置应用到机器。

无论您要安装集群还是向其中添加机器，Ignition 始终都执行 OpenShift Container Platform 集群机器的初始配置。实际的系统设置大多都在每台机器上进行。对于每台机器，Ignition 都会获取 RHCOS 镜像并启动 RHCOS 内核。内核命令行上的选项标识部署的类型，以及启用了 Ignition 的初始 Ram 磁盘 (initramfs) 的位置。

OpenShift Container Platform 使用 Ignition 版本 2 和 Ignition 配置版本 2.3

5.1.3.1. Ignition 工作方式

要使用 Ignition 创建机器，需要 Ignition 配置文件。OpenShift Container Platform 安装程序创建部署集群所需的 Ignition 配置文件。这些文件基于您直接提供给安装程序或通过 **install-config.yaml** 文件提供的信息。

Ignition 配置机器的方式类似于 **cloud-init** 或 Linux Anaconda **kickstart** 等工具配置系统的方式，但有一些重要的区别：

- Ignition 从一个初始 RAM 磁盘运行，该磁盘与您要安装到的系统相隔离。因此，Ignition 可以重新分区磁盘、设置文件系统，以及对机器的持久文件系统执行其他更改。与之相反，cloud-init 会在系统启动时作为机器的初始系统的一部分运行，因而不易对磁盘分区之类的事项进行基本的更改。使用 cloud-init 时，难以在节点启动期间重新配置启动过程。
- Ignition 旨在初始化系统，而不是更改现有系统。机器完成初始化且内核在安装的系统上运行之后，OpenShift Container Platform 集群中的 Machine Config Operator 将完成所有后续的机器配置。
- Ignition 不是完成一组定义的操作，而是实施声明性配置。它会在新机器启动之前检查所有分区、文件、服务和其他项目是否就位。然后进行更改，例如将必要的文件复制到磁盘，以便新机器符合指定的配置。
- 在 Ignition 完成机器配置之后，内核将继续运行，但会丢弃初始 RAM 磁盘，并转至磁盘上已安装的系统。所有新的系统服务和其他功能都将启动，无需重启系统。
- 因为 Ignition 会确认所有新机器是否都符合声明的配置，所以不会存在配置不全的机器。如果机器设置失败，则初始化过程不会完成，而且 Ignition 也不会启动新机器。您的集群永不会包含配置不全的机器。如果 Ignition 无法完成，机器就不会添加到集群中。您必须添加新的机器。一些

失败配置任务的结果可能一直不为人所知，直到后来依赖它的某些事物也失败才被发现。对这类问题的故障排除将会非常困难。而 Ignition 配置的工作方式可以防止出现这个问题。

- 如果 Ignition 配置存在问题，而导致一台机器的设置失败，Ignition 不会尝试使用相同的配置来设置另一台机器。例如，故障可能源自于某一个 Ignition 配置，而构成该配置的父级和子级配置都希望创建同一个文件。在这种情况下，出现的故障将导致 Ignition 配置无法再次用于设置其他机器，直到问题解决为止。
- 如果您有多个 Ignition 配置文件，您可获得该组配置的并集。由于 Ignition 是声明性的，配置之间的冲突可能会导致 Ignition 无法设置机器。这些文件中信息的次序无关紧要。Ignition 将以最有效的方式对每项设置进行分类和实施。例如，如果一个文件需要有多个层级深的目录，而另一个文件需要其路径上的某一目录，则首先创建后一个文件。Ignition 按深度排序并创建所有文件、目录和链接。
- 因为 Ignition 可以从全空的硬盘开始，所以它可以做 cloud-init 不能做的任务：从头开始在裸机上设置系统（使用 PXE 启动等功能）。在裸机情形中，Ignition 配置注入启动分区，以便 Ignition 可以找到它并正确配置系统。

5.1.3.2. Ignition 操作序列

OpenShift Container Platform 集群中 RHCOS 机器的 Ignition 过程包括以下步骤：

- 机器获取其 Ignition 配置文件。Master 机器从 bootstrap 机器获取 Ignition 配置文件，worker 机器从 master 机器获取 Ignition 配置文件。
- Ignition 在机器上创建磁盘分区、文件系统、目录和链接。它支持 RAID 阵列，但不支持 LVM 卷。
- Ignition 将持久文件系统的根目录挂载到 initramfs 中的 **/sysroot** 目录，然后开始在 **/sysroot** 中工作。
- Ignition 配置所有定义的文件系统，并将它们设置为在运行时进行相应地挂载。
- Ignition 运行 **systemd** 临时文件，将必要的文件填充到 **/var** 目录。
- Ignition 运行 Ignition 配置文件，以设置用户、systemd 单元文件和其他配置文件。
- Ignition 卸载 initramfs 中挂载的持久系统中的所有组件。
- Ignition 启动新机器的初始化过程，该过程再启动系统启动期间在机器上运行的所有其他服务。

然后，机器便已准备好加入集群，不需要重启。

5.2. 查看 IGNITION 配置文件

要查看用于部署 bootstrap 机器的 Ignition 配置文件，请运行以下命令：

```
$ openshift-install create ignition-configs --dir $HOME/testconfig
```

回答几个问题后，您所在的目录中将出现 **bootstrap.ign**、**master.ign** 和 **worker.ign** 文件。

要查看 **bootstrap.ign** 文件的内容，请通过 **jq** 过滤器对其进行管道传递。以下是该文件的片段：

```
$ cat $HOME/testconfig/bootstrap.ign | jq
```

```

\{
  "ignition": \{
    "config": \{,
...
  "storage": \{
    "files": [
      \{
        "filesystem": "root",
        "path": "/etc/motd",
        "user": \{
          "name": "root"
        },
        "append": true,
        "contents": \{
          "source": "data:text/plain;charset=utf-
8;base64,VGhpcyBpcyB0aGUgYm9vdHN0cmFwIG5vZGU7IGl0IHdpcGwgYmUgZGVzdHJveWVkiHdo
ZW4gdGhlIG1hc3RlciBpcyBmdWxseSB1cC4KCIRoZSBwcmItYXJ5IHNIcnZpY2UgaXMglmJvb3RrdWJl
nNlcnZpY2UiLiBUbyB3YXRjaCBpdHMgc3RhdHVzLCBydW4gZS5nLgoKICBqb3VybmFsY3RslC1iIC1m
IC11IGJvb3RrdWJlLnNlcnZpY2UK",

```

要解码 **bootstrap.ign** 文件中列出的文件内容，请将代表该文件内容的 base64 编码的数据字符串通过管道传递给 **base64 -d** 命令。以下示例使用了上方输出中添加至 bootstrap 机器的 **/etc/motd** 文件的内容：

```

$ echo
VGhpcyBpcyB0aGUgYm9vdHN0cmFwIG5vZGU7IGl0IHdpcGwgYmUgZGVzdHJveWVkiHdoZW4gdG
hlIG1hc3RlciBpcyBmdWxseSB1cC4KCIRoZSBwcmItYXJ5IHNIcnZpY2UgaXMglmJvb3RrdWJlLnNlcnZp
Y2UiLiBUbyB3YXRjaCBpdHMgc3RhdHVzLCBydW4gZS5nLgoKICBqb3VybmFsY3RslC1iIC1mIC11IGJ
vb3RrdWJlLnNlcnZpY2UK | base64 -d

```

这是 bootstrap 机器，将在 master 完全启动后销毁。

主要服务是“bootkube.service”。要查看其状态，请运行如下命令：

```
journalctl -b -f -u bootkube.service
```

对 **master.ign** 和 **worker.ign** 文件重复这些命令，查看每种机器类型的 Ignition 配置文件的来源。对于 **worker.ign**，您应该会看到类似于下面这一行，它确认了如何从 bootstrap 获取 Ignition 配置：

```
"source": "https://api.myign.develcluster.example.com:22623/config/worker",
```

您可以从 **bootstrap.ign** 文件中了解到以下内容：

- 格式：文件的格式在 [Ignition 配置规范](#) 中定义。MCO 稍后使用相同格式的文件，将更改合并到机器的配置中。
- 内容：由于 bootstrap 机器为其他机器提供 Ignition 配置，因此 master 机器和 worker 机器的 Ignition 配置信息都与 bootstrap 机器的配置一起存储在 **bootstrap.ign** 中。
- 大小：文件长度超过 1300 行，包含指向各种资源的路径。
- 要复制到机器的每个文件的内容实际上编码为数据 URL，这往往会使内容读起来有些混乱。（使用前面演示的 jq 和 base64 命令可使内容更易读。）
- 配置：Ignition 配置文件的不同部分通常涵盖刚放入机器文件系统中的文件，而不是用于修改现有文件的命令。例如，不添加与配置该服务的 NFS 相关的一节，是仅添加一个 NFS 配置文件，然后在系统启动时由 init 进程启动该文件。
- 用户：创建一个名为 core 的用户，并将您的 ssh 密钥分配给该用户。这样，您可以使用该用户名和凭证来登录集群。
- 存储：存储部分标识添加到每台机器的文件。一些值得注意的文件包括 **/root/.docker/config.json**（提供集群从容器镜像 registry 表拉取时所需的凭证），以及 **/opt/openshift/manifests** 中用于配置集群的一系列清单文件：
- systemd：systemd 部分包含用于创建 systemd 单元文件的内容。这些文件用于在启动时启动服务，还用于在运行中的系统上管理这些服务。
- 原语：Ignition 还公开低级别原语，其他工具可以此为基础进行构建。

5.3. 安装后更改 IGNITION 配置

机器配置池管理节点集群及其相应的机器配置（Machine Configs）。机器配置包含集群的配置信息。列出所有已知的机器配置池：

```
$ oc get machineconfigpools
```

```
NAME CONFIG                                UPDATED UPDATING DEGRADED
master master-1638c1aea398413bb918e76632f20799 False False False
worker worker-2feef4f8288936489a5a832ca8efe953 False False False
```

列出所有机器配置：

```
$ oc get machineconfig
```

```
NAME                                GENERATEDBYCONTROLLER IGNITIONVERSION  CREATED
OSIMAGEURL
00-master                            4.0.0-0.150.0.0-dirty 2.2.0            16m
00-master-ssh                        4.0.0-0.150.0.0-dirty                16m
00-worker                            4.0.0-0.150.0.0-dirty 2.2.0            16m
00-worker-ssh                        4.0.0-0.150.0.0-dirty                16m
01-master-kubelet                    4.0.0-0.150.0.0-dirty 2.2.0            16m
01-worker-kubelet                    4.0.0-0.150.0.0-dirty 2.2.0            16m
master-1638c1aea398413bb918e76632f20799 4.0.0-0.150.0.0-dirty 2.2.0            16m
worker-2feef4f8288936489a5a832ca8efe953 4.0.0-0.150.0.0-dirty 2.2.0            16m
```

在涉及到应用这些机器配置时，Machine Config Operator 的行为与 Ignition 有些不同。MachineConfig

按顺序读取（从 00* 到 99*）。MachineConfig 中的标签标识每个所用于的节点类型（master 节点或 worker 节点）。如果同一文件出现在多个 MachineConfig 文件中，则以最后一个文件为准。例如，出现在 99* 文件中的任何文件都将替换出现在 00* 文件中的同一文件。输入的 MachineConfig 对象将合并为一个“渲染的”MachineConfig 对象，该对象将被 Operator 用作目标，也是您可以在 MachineConfigPool 中看到的值。

要查看正在通过 MachineConfig 管理的文件，请查找特定 MachineConfig 中的“Path:”。例如：

```
$ oc describe machineconfigs 01-worker-container-runtime | grep Path:
    Path:      /etc/containers/registries.conf
    Path:      /etc/containers/storage.conf
    Path:      /etc/crio/crio.conf
```

如果要更改其中一个文件中的某一设置，例如将 **crio.conf** 文件中的 `pids_limit` 更改为 1500 (`pids_limit = 1500`)，您可以创建一个仅含想要更改的文件的新 MachineConfig。

确保为 MachineConfig 提供更新的名称（例如 10-worker-container-runtime）。请记住，每个文件的内容都是 URL 样式的数据。然后，将新 MachineConfig 应用到集群。