OpenShift Dedicated 4

规划您的环境

Dedicated 4 计划概述

Last Updated: 2024-10-15

OpenShift Dedicated 4 规划您的环境

Dedicated 4 计划概述

法律通告

摘要

本文档为 OpenShift Dedicated 集群部署提供了规划注意事项。

目录

第 1 章 限制和可扩展性

本文档详细介绍了为 OpenShift Dedicated 集群测试的集群最大值，以及用于测试最大测试环境和配置的信息。另外还提供了有关 control plane 和基础架构节点大小和扩展的信息。

1.1. 集群最大限制

在规划 OpenShift Dedicated 集群安装时，请考虑以下测试的对象最大值。表指定 OpenShift Dedicated 集群中每个测试类型的最大限值。

这些指南基于多个可用区配置中的 180 个计算（也称为 worker）节点的集群。对于较小的集群，最大值限制会较低。

表 1.1. 测试的集群最大值

最大类型	4.x 测试的最大值
pod 数量 [1]	25,000
每个节点的 pod 数量	250
每个内核的 pod 数量	没有默认值
命名空间数量 [2]	5,000
每个命名空间的 pod 数量 [3]	25,000
服务数 [4]	10,000
每个命名空间的服务数	5,000
每个服务中的后端数	5,000
每个命名空间的部署数量 [3]	2,000

1. 这里的 pod 数量是 test pod 的数量。pod 的实际数量取决于应用程序的内存、CPU 和存储要求。
2. 当有大量活跃的项目时，如果键空间增长过大并超过空间配额，etcd 的性能将会受到影响。强烈建议您定期维护 etcd 存储（包括整理碎片）来释放 etcd 存储。
3. 系统中有一些控制循环必须迭代给定命名空间中的所有对象，作为对一些状态更改的响应。在单一命名空间中有大量给定类型的对象可使这些循环的运行成本变高，并降低对给定状态变化的处理速度。限制假设系统有足够的 CPU 、内存和磁盘来满足应用程序的要求。
4. 每个服务端口和每个服务后端在 **iptables** 中都有对应条目。给定服务的后端数量会影响端点对象的大小，这会影响到整个系统发送的数据大小。

1.2. OPENSHIFT CONTAINER PLATFORM 测试环境和配置

下表列出了为 AWS 云平台测试集群最大值的 OpenShift Container Platform 环境和配置。

节点	类型	vCPU	RAM(GiB)	磁盘类型	磁盘大小(GiB)/IOPS	数量	区域
control plane/etcd [1]	m5.4xlarge	16	64	gp3	350 / 1,000	3	us-west-2
基础架构节点 [2]	r5.2xlarge	8	64	gp3	300 / 900	3	us-west-2
Workload [3]	m5.2xlarge	8	32	gp3	350 / 900	3	us-west-2
Compute 节点	m5.2xlarge	8	32	gp3	350 / 900	102	us-west-2

1. io1 磁盘用于 4.10 之前版本中的 control plane/etcd 节点。
2. 基础架构节点用于托管监控组件，因为 Prometheus 可以根据使用情况模式声明大量内存。
3. 工作负载节点专用于运行性能和可扩展工作负载生成器。

更大的集群大小和更高的对象数量可能可以被访问。但是，基础架构节点的大小限制 Prometheus 可用的内存量。在创建、修改或删除对象时，Prometheus 会将指标存储在其内存中，时长大约 3 小时，然后再在磁盘上保留指标。如果创建、修改或删除对象的速率过高，Prometheus 可能会因为缺少内存资源而造成问题问题。

1.3. CONTROL PLANE 和基础架构节点大小和扩展

安装 OpenShift Dedicated 集群时，control plane 和基础架构节点的大小由计算节点计数自动决定。

如果您在安装后更改集群中的计算节点数量，Red Hat Site Reliability Engineering (SRE)团队会根据需要扩展 control plane 和基础架构节点，以保持集群稳定性。

1.3.1. 安装过程中的节点大小

在安装过程中，control plane 和基础架构节点的大小会被动态计算。大小计算基于集群中计算节点的数量。

下表列出了在安装过程中应用的 control plane 和基础架构节点大小。

AWS control plane 和基础架构节点大小：

计算节点数量	control plane 大小	基础架构节点大小
1 到 25	m5.2xlarge	r5.xlarge

计算节点数量	control plane 大小	基础架构节点大小
26 到 100	m5.4xlarge	r5.2xlarge
101 到 180	m5.8xlarge	r5.4xlarge

GCP control plane 和基础架构节点大小：

计算节点数量	control plane 大小	基础架构节点大小
1 到 25	custom-8-32768	custom-4-32768-ext
26 到 100	custom-16-65536	custom-8-65536-ext
101 到 180	custom-32-131072	custom-16-131072-ext

GCP control plane 和基础架构节点大小，以便在 2024 年 6 月 21 日创建的集群：

计算节点数量	control plane 大小	基础架构节点大小
1 到 25	n2-standard-8	n2-highmem-4
26 到 100	n2-standard-16	n2-highmem-8
101 到 180	n2-standard-32	n2-highmem-16

注意

OpenShift Dedicated 上的最大计算节点数量为 180。

1.3.2. 安装后的节点扩展

如果您在安装后更改计算节点数量，则 control plane 和基础架构节点会根据需要由 Red Hat Site Reliability Engineering (SRE)团队扩展。节点已扩展以保持平台稳定性。

control plane 和基础架构节点安装后扩展要求会根据具体情况进行评估。考虑使用节点资源消耗和接收的警报。

control plane 节点重新定义警报大小的规则

在以下情况下，会为集群中的 control plane 节点触发重新定义大小警报：

- control plane 节点会保持集群中平均 66% 的利用率。

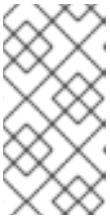
注意

OpenShift Dedicated 上的最大计算节点数量为 180。

基础架构节点大小警报的规则

当具有高 CPU 或内存使用率时，会为集群中的基础架构节点触发重新定义警报的大小。这个高影响的利用率状态为：

- 基础架构节点在具有两个基础架构节点的单一可用区的集群中保持超过 50% 的利用率。
- 基础架构节点在具有 3 个基础架构节点的多个可用区的集群中平均保持 66% 的利用率。



注意

OpenShift Dedicated 上的最大计算节点数量为 180。

调整大小的警报仅在达到高利用率时显示。短期使用量激增（如节点暂时关闭导致其他节点扩展）不会触发这些警报。

SRE 团队可能会因为其他原因扩展 control plane 和基础架构节点，例如管理节点上资源消耗的增加。

1.3.3. 大集群的大小注意事项

对于大集群，基础架构节点大小可能会严重影响可扩展性。很多因素会影响指定的阈值，包括 etcd 版本或者存储数据格式。

超过这些限制并不一定意味着集群将失败。在大多数情况下，超过这些限制会降低整体性能。

第 2 章 AWS 上的客户云订阅

OpenShift Dedicated 提供了一个客户云订阅 (CCS) 模型，允许红帽将集群部署和管理到客户的现有 Amazon Web Service (AWS) 帐户中。

2.1. 了解 AWS 上的客户云订阅

要使用客户云订阅 (CCS) 模型将 OpenShift Dedicated 部署到现有 Amazon Web Services (AWS) 帐户中，红帽需要满足几个先决条件。

红帽建议使用 AWS 机构来管理多个 AWS 帐户。由客户管理的 AWS 机构托管多个 AWS 帐户。机构中有一个 root 帐户，所有帐户都将在帐户层次结构中引用。

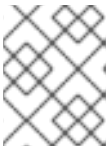
建议在 AWS 机构单元的 AWS 帐户中使用 CCS 模型托管 OpenShift Dedicated 集群。创建服务控制策略 (SCP) 并应用到 AWS 机构单元，后者管理 AWS 子帐户可访问的服务。SCP 仅适用于单个 AWS 帐户内对机构单元中的所有 AWS 子帐户的可用权限。也可以将 SCP 应用到单个 AWS 帐户。客户 AWS 组织中的所有其他帐户以客户要求的任何方式进行管理。红帽站点可靠性工程师 (SRE) 对 AWS 机构中的 SCP 没有任何控制。

2.2. 客户要求

在 Amazon Web Services (AWS) 上使用客户云订阅 (CCS) 模型的 OpenShift Dedicated 集群必须满足几个先决条件，然后才能进行部署。

2.2.1. 帐户

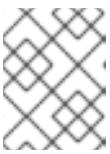
- 客户可确保 AWS 限制 足以支持在客户提供的 AWS 帐户中置备的 OpenShift Dedicated。
- 客户提供的 AWS 帐户应该位于客户的 AWS 机构中，并应用了适用服务控制策略 (SCP)。



注意

不要求客户的帐户位于 AWS 机构内或要应用的 SCP，但红帽必须能够在不限制任何限制的情况下执行 SCP 中列出的所有操作。

- 客户提供的 AWS 帐户不能转移到红帽。
- 客户可能没有对红帽的活动实施 AWS 使用限制。实施限制会严重破坏红帽响应事件的能力。
- 红帽会在 AWS 中部署监控，以便在有高特权的帐户（如 root 帐户）登录到客户提供的 AWS 帐户时提醒红帽。
- 客户可以在同一客户提供的 AWS 帐户内部署原生 AWS 服务。

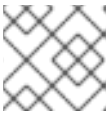


注意

我们鼓励客户在虚拟私有云 (VPC) 中部署资源，并与托管 OpenShift Dedicated 和其他红帽支持服务的 VPC 部署资源。

2.2.2. 访问要求

- 要在 AWS 服务上正确管理 OpenShift Dedicated 服务，红帽始终必须将 **AdministratorAccess** 策略应用到管理员角色。



注意

此政策只为红帽提供了更改客户提供的 AWS 帐户资源的权限和功能。

- 红帽必须具有对客户提供的 AWS 帐户的 AWS 控制台访问权限。此访问权限由红帽保护和管理。
- 客户不得使用 AWS 帐户在 OpenShift Dedicated 集群中提升其权限。
- OpenShift Cluster Manager 中可用的操作不能直接在客户提供的 AWS 帐户中执行。

2.2.3. 支持要求

- 红帽建议客户至少有 AWS 的商业支持（Business Support）。
- 红帽由客户的授权，可以代表他们请求 AWS 支持。
- 红帽的客户授权可以请求对客户账户增加 AWS 资源限制。
- 除非本要求部分中另有指定，否则红帽以相同的方式管理所有 OpenShift Dedicated 集群上的限制、预期和默认值。

2.2.4. 安全要求

- 客户提供的 IAM 凭证对于客户提供的 AWS 帐户来说必须是唯一的，且不得存储在客户提供的 AWS 帐户中的任何位置。
- 卷快照将保留在客户提供的 AWS 帐户和客户指定的区域。
- 红帽必须通过白名单的红帽机器对 EC2 主机和 API 服务器进行入口访问。
- 红帽需要有一个出口，可以将系统和审计日志转发到红帽管理的中央日志记录环境中。

2.3. 所需的客户流程

客户云订阅 (CCS) 模型允许红帽在客户的 Amazon Web Services (AWS) 帐户中部署和管理 OpenShift Dedicated。为了提供这些服务，红帽需要满足几个先决条件。

流程

1. 如果客户使用 AWS 机构，则必须在您的机构中使用 AWS 帐户或创建一个新帐户。
2. 为确保红帽可以执行必要的操作，您必须创建一个服务控制策略 (SCP)，或确保 none 应用到 AWS 帐户。
3. 将 SCP 附加到 AWS 帐户中。
4. 在 AWS 帐户中，您必须创建一个具有以下要求的 **osdCcsAdmin** IAM 用户：
   - 此用户需要最少启用 **Programmatic access**。
   - 此用户必须附加了 **AdministratorAccess** 策略。
5. 向红帽提供 IAM 用户凭证。
   - 您必须在 OpenShift Cluster Manager 中提供 **访问密钥 ID** 和 **secret 访问密钥**。

2.4. 最低需要的服务控制策略 (SCP)

服务控制策略 (SCP) 管理由客户自己负责。这些策略在 AWS 机构中维护，并控制附加的 AWS 帐户中可用服务。

必需/可选	服务	Actions	效果
必需	Amazon EC2	All	Allow
	Amazon EC2 自动扩展	All	Allow
	Amazon S3	All	Allow
	身份和访问管理	All	Allow
	Elastic Load Balancing	All	Allow
	Elastic Load Balancing V2	All	Allow
	Amazon CloudWatch	All	Allow
	Amazon CloudWatch Events	All	Allow
	Amazon CloudWatch Logs	All	Allow
	AWS Support	All	Allow
	AWS 密钥管理服务	All	Allow
	AWS 安全令牌服务	All	Allow
	AWS Resource Tagging	All	Allow
	AWS Route53 DNS	All	Allow
	AWS Service Quotas	ListServices GetRequestedServiceQuotaChange GetServiceQuota RequestServiceQuotaIncrease ListServiceQuotas	Allow

必需/可选	服务	Actions	效果
选填	AWS Billing	ViewAccount Viewbilling ViewUsage	Allow
	AWS 成本和使用量报告	All	Allow
	AWS Cost Explorer Services	All	Allow

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "ec2:*"
            ],
            "Resource": [
                "*"
            ]
        },
        {
            "Effect": "Allow",
            "Action": [
                "autoscaling:*"
            ],
            "Resource": [
                "*"
            ]
        },
        {
            "Effect": "Allow",
            "Action": [
                "s3:*"
            ],
            "Resource": [
                "*"
            ]
        },
        {
            "Effect": "Allow",
            "Action": [
                "iam:*"
            ],
            "Resource": [
                "*"
            ]
        },
        {
```

```
        "Effect": "Allow",
        "Action": [
            "elasticloadbalancing:*"
        ],
        "Resource": [
            "*"
        ]
    },
    {
        "Effect": "Allow",
        "Action": [
            "cloudwatch:*"
        ],
        "Resource": [
            "*"
        ]
    },
    {
        "Effect": "Allow",
        "Action": [
            "events:*"
        ],
        "Resource": [
            "*"
        ]
    },
    {
        "Effect": "Allow",
        "Action": [
            "logs:*"
        ],
        "Resource": [
            "*"
        ]
    },
    {
        "Effect": "Allow",
        "Action": [
            "support:*"
        ],
        "Resource": [
            "*"
        ]
    },
    {
        "Effect": "Allow",
        "Action": [
            "kms:*"
        ],
        "Resource": [
            "*"
        ]
    },
    {
        "Effect": "Allow",
        "Action": [
```

```
            "sts:*"
        ],
        "Resource": [
            "*"
        ]
    },
    {
        "Effect": "Allow",
        "Action": [
            "tag:*"
        ],
        "Resource": [
            "*"
        ]
    },
    {
        "Effect": "Allow",
        "Action": [
            "route53:*"
        ],
        "Resource": [
            "*"
        ]
    },
    {
        "Effect": "Allow",
        "Action": [
            "servicequotas:ListServices",
            "servicequotas:GetRequestedServiceQuotaChange",
            "servicequotas:GetServiceQuota",
            "servicequotas:RequestServiceQuotaIncrease",
            "servicequotas:ListServiceQuotas"
        ],
        "Resource": [
            "*"
        ]
    }
  ]
}
```

2.5. RED HAT MANAGED IAM 参考

红帽负责创建和管理以下 Amazon Web Services (AWS)资源：IAM 策略、IAM 用户和 IAM 角色。

2.5.1. IAM 策略

注意

IAM 策略会随着 OpenShift Dedicated 更改的功能而进行修改。

- **AdministratorAccess** 策略由管理角色使用。此政策提供了在客户提供的 AWS 帐户中管理 OpenShift Dedicated 集群所需的访问权限。

```
{
```

```
    "Version": "2012-10-17",
    "Statement": [
      {
        "Action": "*",
        "Resource": "*",
        "Effect": "Allow"
      }
    ]
}
```

- **CustomerAdministratorAccess** 角色为客户提供管理 AWS 帐户中的服务子集的访问权限。目前，允许以下内容：
  - VPC Peering
  - VPN 设置
  - 直接连接（仅在通过服务控制策略授予时才可用）

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "ec2:AttachVpnGateway",
                "ec2:DescribeVpnConnections",
                "ec2:AcceptVpcPeeringConnection",
                "ec2:DeleteVpcPeeringConnection",
                "ec2:DescribeVpcPeeringConnections",
                "ec2:CreateVpnConnectionRoute",
                "ec2:RejectVpcPeeringConnection",
                "ec2:DetachVpnGateway",
                "ec2:DeleteVpnConnectionRoute",
                "ec2:DeleteVpnGateway",
                "ec2:DescribeVpcs",
                "ec2:CreateVpnGateway",
                "ec2:ModifyVpcPeeringConnectionOptions",
                "ec2:DeleteVpnConnection",
                "ec2:CreateVpcPeeringConnection",
                "ec2:DescribeVpnGateways",
                "ec2:CreateVpnConnection",
                "ec2:DescribeRouteTables",
                "ec2:CreateTags",
                "ec2:CreateRoute",
            "directconnect:*"
            ],
            "Resource": "*"
        }
    ]
}
```

- 如果启用，**BillingReadOnlyAccess** 角色会提供只读访问权限，以查看帐户的计费和使用信息。只有 AWS 机构中的 root 帐户启用了它时，才会授予计费和使用访问权限。这是客户必须执行的可选步骤，才能启用只读账单和使用访问，不会影响创建此配置集及其使用的角色。如果没有启用此角色，用户将不会看到计费和使用信息。请参阅本教程，了解如何启用对计费数据的访问。

```
{
   "Version": "2012-10-17",
   "Statement": [
      {
         "Effect": "Allow",
         "Action": [
            "aws-portal:ViewAccount",
            "aws-portal:ViewBilling"
         ],
         "Resource": "*"
      }
   ]
}
```

2.5.2. IAM 用户

在控制客户提供的 AWS 帐户后，将立即创建 **osdManagedAdmin** 用户。这是将执行 OpenShift Dedicated 集群安装的用户。

2.5.3. IAM 角色

- **network-mgmt** 角色通过单独的 AWS 帐户提供对 AWS 帐户的管理访问权限。它还具有与只读角色相同的访问权限。**network-mgmt** 角色只适用于非自定义云订阅 (CCS) 集群。以下策略附加到角色中：
  - AmazonEC2ReadOnlyAccess
  - CustomerAdministratorAccess
- **只读** 角色通过单独的 AWS 帐户提供对 AWS 帐户的只读访问权限。以下策略附加到角色中：
  - AWSAccountUsageReportAccess
  - AmazonEC2ReadOnlyAccess
  - AmazonS3ReadOnlyAccess
  - IAMReadOnlyAccess
  - BillingReadOnlyAccess

2.6. 置备的 AWS 基础架构

这是在部署的 OpenShift Dedicated 中置备的 Amazon Web Services (AWS) 组件的概述。有关所有置备的 AWS 组件的详细列表，请参阅 OpenShift Container Platform 文档。

2.6.1. AWS Elastic Computing (EC2) 实例

在 AWS 公有云中部署 OpenShift Dedicated 的 control plane 和 data plane 功能需要 AWS EC2 实例。根据 worker 节点数，实例类型可能会因 control plane 和基础架构节点而异。

- 单个可用区
  - 3 m5.2xlarge minimum (control plane 节点)

- 2 r5.xlarge minimum (基础架构节点)
- 2 m5.xlarge minimum 但会有很大不同 (worker 节点)

- 多个可用区
  - 3 m5.2xlarge minimum (control plane 节点)
  - 3 r5.xlarge minimum (基础架构节点)
  - 3 m5.xlarge minimum 但会有很大不同 (worker 节点)

2.6.2. AWS Elastic Block Store (EBS) 存储

Amazon EBS 块存储用于本地节点存储和持久性卷存储。

每个 EC2 实例的卷要求：

- control plane 卷
  - 大小：350 GB
  - 类型： io1
  - 每秒输入/输出操作：1000
- 基础架构卷
  - 大小：300 GB
  - 类型： gp2
  - 每秒输入/输出操作：900
- Worker 卷
  - 大小：300 GB
  - 类型： gp2
  - 每秒输入/输出操作：900

2.6.3. Elastic Load Balancing (ELB)负载均衡器

最多两个用于 API 的 Network Load Balancers，最多两个用于应用程序路由器的 Classic Load Balancers。如需更多信息，请参阅 AWS 的 ELB 文档。

2.6.4. S3 存储

镜像 registry 和 Elastic Block Store (EBS) 卷快照由 AWS S3 存储支持。定期修剪资源以优化 S3 使用量和集群性能。

注意

需要两个存储桶，每个存储桶典型的大小为 2 TB。

2.6.5. VPC

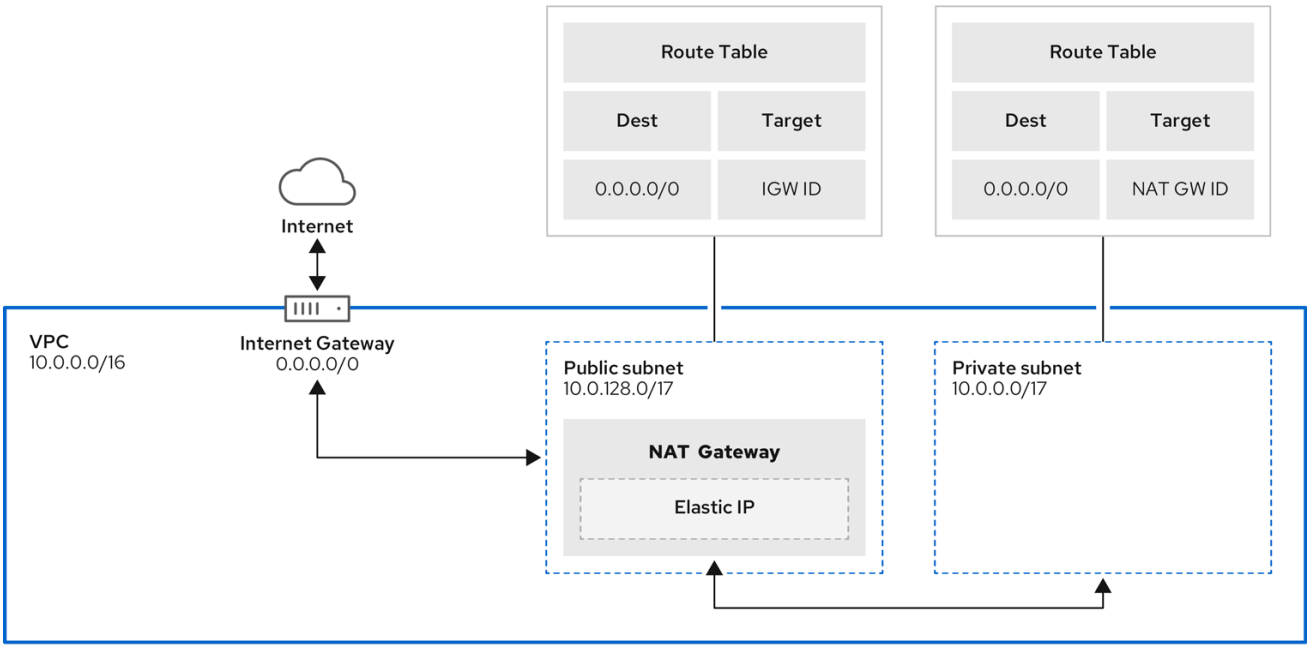
客户应该希望看到每个集群一个 VPC。另外，VPC 需要以下配置：

- **子网** ：一个具有单一可用区的集群的两个子网，或具有多个可用区的集群 6 个子网。

> **注意**
>
> **公共子网**通过互联网网关直接连接到互联网。**专用子网**通过网络地址转换 (NAT) 网关连接到互联网。

- **路由器表** ：每个专用子网一个路由器表，每个集群一个额外表。
- **Internet 网关** ：每个集群一个互联网网关。
- **NAT 网关** ：每个公共子网一个 NAT 网关。

2.6.5.1. VPC 架构示例



2.6.6. 安全组

AWS 安全组在协议和端口访问级别提供安全；它们与 EC2 实例和 Elastic Load Balancing 关联。每个安全组包含一组规则，这些规则过滤进出 EC2 实例的流量。您必须确保 OpenShift Container Platform 安装在网络上打开所需的端口，并配置为允许主机间的访问。

2.6.6.1. 其他自定义安全组

当使用非管理的 VPC 创建集群时，您可以在集群安装过程中添加自定义安全组。自定义安全组受以下限制：

- 在创建集群时，您必须在 AWS 中创建自定义安全组。如需更多信息，请参阅适用于 Linux 实例的 Amazon EC2 安全组。

- 您必须将自定义安全组与集群要安装的 VPC 关联。您的自定义安全组不能与另一个 VPC 关联。
- 如果要添加额外的自定义安全组，您可能需要为 VPC 请求额外的配额。有关请求 AWS 配额增加的详情，请参阅请求配额增加。

2.7. AWS 防火墙先决条件

如果使用防火墙来控制 OpenShift Dedicated 的出口流量，您必须配置防火墙，以授予以下特定域和端口组合的访问权限。OpenShift Dedicated 需要此访问权限来提供完全托管的 OpenShift 服务。

先决条件

- 您已在 AWS Virtual Private Cloud (VPC)中配置了 Amazon S3 网关端点。需要此端点才能完成从集群到 Amazon S3 服务的请求。

流程

1. 允许列出用于安装和下载软件包和工具的以下 URL：

域	端口	功能
registry.redhat.io	443	提供核心容器镜像。
quay.io	443	提供核心容器镜像。
cdn01.quay.io	443	提供核心容器镜像。
cdn02.quay.io	443	提供核心容器镜像。
cdn03.quay.io	443	提供核心容器镜像。
sso.redhat.com	443	必需。**https://console.redhat.com/openshift** 站点使用来自 **sso.redhat.com** 的身份验证下载 pull secret，并使用 Red Hat SaaS 解决方案来简化订阅、集群清单、计费报告等的监控。
quay-registry.s3.amazonaws.com	443	提供核心容器镜像。
quayio-production-s3.s3.amazonaws.com	443	提供核心容器镜像。
openshift.org	443	提供 Red Hat Enterprise Linux CoreOS (RHCOS) 镜像。
registry.access.redhat.com	443	托管存储在 Red Hat Ecosytem Catalog 中的所有容器镜像。另外，registry 提供了对 **odo** CLI 工具的访问，可帮助开发人员在 OpenShift 和 Kubernetes 上进行构建。

域	端口	功能
access.redhat.com	443	必需。托管容器客户端在从 **registry.access.redhat.com** 中拉取镜像时验证镜像所需的签名存储。
registry.connect.redhat.com	443	所有第三方镜像和认证 Operator 都需要。
console.redhat.com	443	必需。允许集群和 OpenShift Console Manager 之间的交互以启用功能，如调度升级。
sso.redhat.com	443	**https://console.redhat.com/openshift** 站点使用来自 **sso.redhat.com**的身份验证
pull.q1w2.quay.rhcloud.com	443	当 quay.io 不可用时，提供核心容器镜像作为回退。
.q1w2.quay.rhcloud.com	443	当 quay.io 不可用时，提供核心容器镜像作为回退。
www.okd.io	443	**openshift.org** 站点通过 **www.okd.io** 重定向。
www.redhat.com	443	**sso.redhat.com** 站点通过 **www.redhat.com** 重定向。
aws.amazon.com	443	**iam.amazonaws.com** 和 **sts.amazonaws.com** 站点通过 **aws.amazon.com** 重定向。
catalog.redhat.com	443	**registry.access.redhat.com** 和 **https://registry.redhat.io** 站点通过 **catalog.redhat.com** 重定向。
dvbwgdztaeq9o.cloudfront.net[1]	443	ROSA 用于带有管理的 OIDC 配置的 STS 实现。

1. 如果 **cloudfront.net** 前面有一个主要云前端中断需要重定向资源，则字母数字字符的字符串可能会改变。

2. 将以下遥测 URL 列入允许列表：

域	端口	功能
cert-api.access.redhat.com	443	遥测是必需的。
api.access.redhat.com	443	遥测是必需的。

域	端口	功能
infogw.api.openshift.com	443	遥测是必需的。
console.redhat.com	443	遥测和 Red Hat Insights 需要。
cloud.redhat.com/api/ingress	443	遥测和 Red Hat Insights 需要。
observatorium-mst.api.openshift.com	443	受管 OpenShift 遥测的需要。
observatorium.api.openshift.com	443	受管 OpenShift 遥测的需要。

受管集群需要启用遥测功能，以便红帽可以更快地对问题做出反应，更好地支持客户，并更好地了解产品升级对集群的影响。有关红帽如何使用远程健康监控数据的更多信息，*请参阅附加资源部分 关于远程健康监控* 的信息。

3. 允许以下 Amazon Web Services (AWS) API URl：

域	端口	功能
.amazonaws.com	443	需要此项以访问 AWS 服务和资源。

或者，如果您选择不为 Amazon Web Services (AWS) API 使用通配符，则必须允许列出以下 URL：

域	端口	功能
ec2.amazonaws.com	443	用于在 AWS 环境中安装和管理集群。
events.<aws_region>.amazonaws.com	443	用于在 AWS 环境中安装和管理集群。
iam.amazonaws.com	443	用于在 AWS 环境中安装和管理集群。
route53.amazonaws.com	443	用于在 AWS 环境中安装和管理集群。
sts.amazonaws.com	443	用于在 AWS 环境中安装和管理集群，用于配置为使用 AWS STS 的全局端点。
sts.<aws_region>.amazonaws.com	443	用于在 AWS 环境中安装和管理集群，用于配置为使用 AWS STS 的区域端点的集群。如需更多信息，请参阅 AWS STS 区域端点。
tagging.us-east-1.amazonaws.com	443	用于在 AWS 环境中安装和管理集群。此端点始终为 us-east-1，无论集群要部署到的区域。
ec2.<aws_region>.amazonaws.com	443	用于在 AWS 环境中安装和管理集群。

域	端口	功能
elasticloadbalancing.<aws_region>.amazonaws.com	443	用于在 AWS 环境中安装和管理集群。
servicequotas.<aws_region>.amazonaws.com	443	必需。用于确认用于部署该服务的配额。
tagging.<aws_region>.amazonaws.com	443	允许以标签的形式分配 AWS 资源的元数据。

4. 将以下 OpenShift URL 列入允许列表：

域	端口	功能
mirror.openshift.com	443	用于访问镜像安装内容和镜像。此站点也是发行版本镜像签名的来源，但 Cluster Version Operator (CVO) 只需要一个可正常工作的源。
storage.googleapis.com/openshift-release (推荐)	443	mirror.openshift.com/ 的替代站点。用于下载集群用来从 quay.io 中拉取哪些镜像的平台发行版本签名。
api.openshift.com	443	用于检查集群是否有可用的更新。

5. 将以下站点可靠性工程 (SRE) 和管理 URL 列入允许：

域	端口	功能
api.pagerduty.com	443	此警报服务由 in-cluster alertmanager 用来发送通知 Red Hat SRE 的事件来执行操作的警报。
events.pagerduty.com	443	此警报服务由 in-cluster alertmanager 用来发送通知 Red Hat SRE 的事件来执行操作的警报。
api.deadmanssnitch.com	443	OpenShift Dedicated 用来发送定期 ping 的警报服务，以指示集群是否可用并在运行。
nosnch.in	443	OpenShift Dedicated 用来发送定期 ping 的警报服务，以指示集群是否可用并在运行。

域	端口	功能
.osdsecuritylogs.splunkcloud.com 或 inputs1.osdsecuritylogs.splunkcloud.cominputs2.osdsecuritylogs.splunkcloud.cominputs4.osdsecuritylogs.splunkcloud.cominputs5.osdsecuritylogs.splunkcloud.cominputs6.osdsecuritylogs.splunkcloud.cominputs7.osdsecuritylogs.splunkcloud.cominputs8.osdsecuritylogs.splunkcloud.cominputs9.osdsecuritylogs.splunkcloud.cominputs10.osdsecuritylogs.splunkcloud.cominputs11.osdsecuritylogs.splunkcloud.cominputs12.osdsecuritylogs.splunkcloud.cominputs13.osdsecuritylogs.splunkcloud.cominputs14.osdsecuritylogs.splunkcloud.cominputs15.osdsecuritylogs.splunkcloud.com	9997	mvapich **-forwarder-operator** 使用为一个日志转发端点，供 Red Hat SRE 用于基于日志的警报。
http-inputs-osdsecuritylogs.splunkcloud.com	443	必需。mvapich **-forwarder-operator** 使用为一个日志转发端点，供 Red Hat SRE 用于基于日志的警报。
sftp.access.redhat.com (推荐)	22	**must-gather-operator** 使用的 SFTP 服务器上传诊断日志，以帮助排除集群中的问题。

6. 将以下 URL 列入允许的可选第三方内容：

域	端口	功能
registry.connect.redhat.com	443	所有第三方镜像和认证操作器是必需的。
rhc4tp-prod-z8cxf-image-registry-us-east-1-evenkyleffocxqvofrk.s3.dualstack.us-east-1.amazonaws.com	443	提供对托管在 **registry.connect.redhat.com**上的容器镜像的访问
oso-rhc4tp-docker-registry.s3-us-west-2.amazonaws.com	443	对于 Sonatype Nexus, F5 Big IP operator 是必需的。

7. 将提供构建所需语言或框架资源的任何站点列入允许列表。

8. 允许允许任何依赖于 OpenShift 中使用的语言和框架的出站 URL。如需防火墙或代理上允许的推荐 URL 列表，请参阅 允许的 OpenShift 出站 URL 。

其他资源

- 关于远程健康监控

2.8. AWS 帐户限值

OpenShift Dedicated 集群使用诸多 Amazon Web Services (AWS) 组件，默认的服务限值会影响您安装 OpenShift Dedicated 集群的能力。如果您使用特定的集群配置，在某些 AWS 区域部署集群，或者从您的帐户运行多个集群，您可能需要为 AWS 帐户请求其他资源。

下表总结了 AWS 组件，它们的限值可能会影响您安装和运行 OpenShift Dedicated 集群的能力。

组件	默认可用的集群数	默认 AWS 限值	描述
实例限值	可变	可变	每个集群至少会创建以下实例： • 一台 Bootstrap 机器，在安装后删除 • 三个 control plane 节点 • 单个可用区的两个基础架构节点：用于多可用区的三个基础架构节点 • 单个可用区有两个 worker 节点：用于多可用区的三个 worker 节点 这些实例类型数量在新帐户的默认限值之内。要部署更多 worker 节点、部署大型工作负载或使用不同的实例类型，请查看您的帐户限制，以确保集群可以部署您需要的机器。 在大多数区域中，bootstrap 和 worker 机器使用 **m4.large** 机器，control plane 机器使用 **m4.xlarge** 实例。在一些区域，包括所有不支持这些实例类型的区域，则使用 **m5.large** 和 **m5.xlarge** 实例。
弹性 IP (EIP)	0 到 1	每个帐户 5 个 EIP	要在高可用性配置中置备集群，安装程序将为区域中的每个可用区创建一个公共和专用子网。每个专用子网都需要 NAT 网关，每个 NAT 网关需要单独的弹性 IP。查看 AWS 区域图来确定每个区域有多少个可用区。要利用默认高可用性，请在至少含有三个可用区的区域安装集群。要在有超过五个可用区的区域安装集群，您必须提高 EIP 限值。  **重要** 要使用 **us-east-1** 区域，必须提高您帐户的 EIP 限值。
虚拟私有云 (VPC)	5	每个区域 5 个 VPC	每个集群创建自己的 VPC。

组件	默认可用的集群数	默认 AWS 限值	描述
Elastic Load Balancing (ELB)	3	每个区域 20 个	默认情况下，每个集群为主 API 服务器创建内部和外部网络负载均衡器，并为路由器创建一个 Classic Load Balancer。部署更多 Kubernetes LoadBalancer Service 对象将生成额外的负载均衡器。
NAT 网关	5	每个可用区 5 个	集群在每个可用区中部署一个 NAT 网关。
弹性网络接口 (ENI)	至少 12 个	每个区域 350 个	默认安装创建 21 个 ENI，并为区域中的每个可用区创建一个 ENI。例如，**us-east-1** 区域包含六个可用区，因此在该区部署的集群将使用 27 个 ENI。查看 AWS 区域图来确定每个区域有多少个可用区。 为使用和部署的工作负载创建的额外机器和负载均衡器创建额外的 ENI。
VPC 网关	20	每个帐户 20 个	每个集群创建一个 VPC 网关来访问 S3。
S3 存储桶	99	每个帐户有 100 个存储桶	因为安装过程会创建一个临时存储桶，并且每个集群中的 registry 组件会创建一个存储桶，所以您只能为每个 AWS 帐户创建 99 个 OpenShift Dedicated 集群。
安全组	250	每个帐户 2,500 个	每个集群创建 10 个不同的安全组。

第 3 章 GCP 上的客户云订阅

OpenShift Dedicated 提供了一个客户云订阅(CCS)模型，允许红帽在客户的现有 Google Cloud Platform (GCP)帐户中部署和管理集群。

3.1. 了解 GCP 上的客户云订阅

Red Hat OpenShift Dedicated 提供了一个客户云订阅(CCS) 模型，允许红帽将 OpenShift Dedicated 部署和管理到客户的现有 Google Cloud Platform (GCP) 帐户中。为了提供此服务，红帽需要满足几个前提条件。

红帽建议使用由客户管理的 GCP 项目来组织所有 GCP 资源。项目由一组用户和 API 组成，以及这些 API 的计费、身份验证和监控设置。

建议在 GCP 机构的 GCP 项目中托管使用 CCS 模型的 OpenShift Dedicated 集群。Organization 资源是 GCP 资源层次结构的根节点，属于某个机构的所有资源都分组到机构节点下。已创建一个具有某些角色的 IAM 服务帐户，并应用到 GCP 项目。当您调用 API 时，您通常提供服务帐户密钥进行身份验证。每个服务帐户归特定的项目所有，但服务帐户可以提供角色来访问其他项目的资源。

3.2. 客户要求

在 Google Cloud Platform (GCP)上使用客户云订阅 (CCS) 模型的 OpenShift Dedicated 集群必须满足几个先决条件，然后才能进行部署。

3.2.1. 帐户

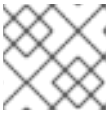
- 客户确保 Google Cloud 限制 足以支持在客户提供的 GCP 帐户中置备的 OpenShift Dedicated。
- 客户提供的 GCP 帐户应该位于客户的 Google Cloud 机构中，并应用了适用服务帐户。
- 客户提供的 GCP 帐户不能转移到红帽。
- 客户可能没有对红帽的活动实施 GCP 使用限制。实施限制会严重破坏红帽响应事件的能力。
- 红帽会在 GCP 中部署监控，以便在有高特权的帐户（如 root 帐户）登录到客户提供的 GCP 帐户时提醒红帽。
- 客户可以在同一客户提供的 GCP 帐户内部署原生 GCP 服务。

注意

我们鼓励客户在虚拟私有云 (VPC) 中部署资源，并与托管 OpenShift Dedicated 和其他红帽支持服务的 VPC 部署资源。

3.2.2. 访问要求

- 要在 AWS 服务上正确管理 OpenShift Dedicated 服务，红帽始终必须将 **AdministratorAccess** 策略应用到管理员角色。



注意

此政策只为红帽提供了更改客户提供的 GCP 帐户资源的权限和功能。

- 红帽必须具有对客户提供的 GCP 帐户的 GCP 控制台访问权限。此访问权限由红帽保护和管理。
- 客户不得使用 GCP 帐户在 OpenShift Dedicated 集群中提升其权限。
- OpenShift Cluster Manager 中可用的操作不能直接在客户提供的 GCP 帐户中执行。

3.2.3. 支持要求

- 红帽建议客户至少具有 GCP 的增强支持。
- 红帽的客户授权可以代表它们请求 GCP 支持。
- 红帽的客户授权可以请求 GCP 资源限制来增加客户提供的帐户。
- 除非本要求部分中另有指定，否则红帽以相同的方式管理所有 OpenShift Dedicated 集群上的限制、预期和默认值。

3.2.4. 安全要求

- 客户提供的 IAM 凭证对于客户提供的 GCP 帐户来说必须是唯一的，且不得存储在客户提供的 GCP 帐户中的任何位置。
- 卷快照将保留在客户提供的 GCP 帐户和客户指定的区域。
- 红帽必须通过允许列表 IP 地址对 API 服务器进行入口访问。

注意

有关允许列表 IP 地址的详情，请参考附加资源。

- 红帽需要有一个出口，可以将系统和审计日志转发到红帽管理的中央日志记录环境中。

3.3. 所需的客户流程

客户云订阅 (CCS) 模型允许红帽在客户的 a customer's Google Cloud Platform (GCP) 项目中部署和管理 OpenShift Dedicated。红帽需要几个先决条件来提供这些服务。

警告

要在 GCP 项目中使用 OpenShift Dedicated，无法放置以下 GCP 组织策略限制：

- **constraints/iam.allowedPolicyMemberDomains** (这个策略约束的支持只限于红帽的 **DIRECTORY_CUSTOMER_ID C02k0l5e8** 包括在允许列表中的情况。请谨慎使用此策略约束）。
- **constraints/compute.restrictLoadBalancerCreationForTypes**
- **constraints/compute.requireShieldedVm** （只有集群安装了在初始集群创建过程中选择的 "Enable Secure Boot support for Shielded VM" 时，才支持此策略约束。
- **constraints/compute.vmExternalIpAccess** （此策略约束在安装后才被支持）。

流程

1. 创建 Google Cloud 项目 来托管 OpenShift Dedicated 集群。
2. 在托管 OpenShift Dedicated 集群的项目中启用以下所需的 API：

表 3.1. 所需的 API 服务

API 服务	控制台服务名称
Cloud Deployment Manager V2 API	**deploymentmanager.googleapis.com**
Compute Engine API	**compute.googleapis.com**
Google Cloud API	**cloudapis.googleapis.com**
Cloud Resource Manager API	**cloudresourcemanager.googleapis.com**
Google DNS API	**dns.googleapis.com**
网络安全 API	**networksecurity.googleapis.com**
IAM Service Account Credentials API	**iamcredentials.googleapis.com**
Identity and Access Management(IAM)API	**iam.googleapis.com**
服务管理 API	**servicemanagement.googleapis.com**
Service Usage API	**serviceusage.googleapis.com**

API 服务	控制台服务名称
Google Cloud Storage JSON API	**storage-api.googleapis.com**
Cloud Storage	**storage-component.googleapis.com**
机构策略 API	**orgpolicy.googleapis.com**

3. 为确保红帽可以执行必要的操作，您必须在 GCP 项目中创建 **osd-ccs-admin** IAM 服务帐户 用户。
以下角色必须授予服务帐户 ：

表 3.2. 所需角色

角色	控制台角色名称
Compute Admin	**roles/compute.admin**
DNS Administrator	**roles/dns.admin**
机构策略查看器	**roles/orgpolicy.policyViewer**
服务管理管理员	**roles/servicemanagement.admin**
Service Usage Admin	**roles/serviceusage.serviceUsageAdmin**
Storage Admin	**roles/storage.admin**
Compute Load Balancer Admin	**roles/compute.loadBalancerAdmin**
角色查看器	**roles/viewer**
Role Administrator	**roles/iam.roleAdmin**
Security Admin	**roles/iam.securityAdmin**
Service Account Key Admin	**roles/iam.serviceAccountKeyAdmin**
Service Account Admin	**roles/iam.serviceAccountAdmin**
Service Account User	**roles/iam.serviceAccountUser**

4. 为 **osd-ccs-admin** IAM 服务帐户创建服务帐户密钥。将密钥导出到名为 **osServiceAccount.json** 的文件；创建集群时，此 JSON 文件将在 Red Hat OpenShift Cluster Manager 中上传。

3.4. RED HAT 管理的 GOOGLE CLOUD 资源

红帽负责创建和管理以下 IAM Google Cloud Platform (GCP) 资源。

3.4.1. IAM 服务帐户和角色

在控制客户提供的 GCP 帐户后，**osd-managed-admin** IAM 服务帐户会立即创建。这是将执行 OpenShift Dedicated 集群安装的用户。

以下角色附加到服务帐户：

表 3.3. osd-managed-admin 的 IAM 角色

角色	控制台角色名称	描述
Compute Admin	**roles/compute.admin**	提供对所有 Compute Engine 资源的完整控制。
DNS Administrator	**roles/dns.admin**	提供对所有云 DNS 资源的读写访问。
Security Admin	**roles/iam.securityAdmin**	安全管理员角色，具有获取和设置任何 IAM 策略的权限。
Storage Admin	**roles/storage.admin**	赋予对象和存储桶的完全控制。 当应用到单个**存储桶**时，控制仅适用于存储桶中指定的存储桶和对象。
Service Account Admin	**roles/iam.serviceAccountAdmin**	创建和管理服务帐户。
Service Account Key Admin	**roles/iam.serviceAccountKeyAdmin**	创建和管理（及轮转）服务帐户密钥。
Service Account User	**roles/iam.serviceAccountUser**	以服务帐户身份运行操作。
Role Administrator	**roles/iam.roleAdmin**	提供对项目中的所有自定义角色的访问权限。

3.4.2. IAM 组和角色

sd-sre-platform-gcp-access Google 组被授予 GCP 项目的访问权限，以允许 Red Hat Site Reliability Engineering (SRE) 访问控制台以实现紧急故障排除目的。

以下角色附加到组中：

表 3.4. sd-sre-platform-gcp-access 的 IAM 角色

角色	控制台角色名称	描述
Compute Admin	**roles/compute.admin**	提供对所有 Compute Engine 资源的完整控制。
Editor	**roles/editor**	提供所有查看权限，以及修改状态的操作的权限。
机构策略查看器	**roles/orgpolicy.policyViewer**	提供查看资源的机构策略的访问权限。
项目 IAM Admin	**roles/resourcemanager.projectIamAdmin**	提供管理项目的 IAM 策略的权限。
配额管理员	**roles/servicemanagement.quotaAdmin**	提供管理服务配额的访问权限。
Role Administrator	**roles/iam.roleAdmin**	提供对项目中的所有自定义角色的访问权限。
Service Account Admin	**roles/iam.serviceAccountAdmin**	创建和管理服务帐户。
Service Usage Admin	**roles/serviceusage.serviceUsageAdmin**	能够为消费者项目启用、禁用和检查服务状态、检查操作以及消耗配额和计费。
技术支持编辑器	**roles/cloudsupport.techSupportEditor**	提供对技术支持问题单的完整读写访问。

3.5. 置备 GCP 基础架构

这是已部署 OpenShift Dedicated 集群中置备的 Google Cloud Platform (GCP) 组件概述。有关所有置备的 GCP 组件的详细列表，请参阅OpenShift Container Platform 文档。

3.5.1. 计算实例

在 GCP 中部署 OpenShift Dedicated 的 control plane 和 data plane 功能需要 GCP 计算实例。根据 worker 节点数，实例类型可能会因 control plane 和基础架构节点而异。

- 单个可用区
  - 2 个 infra 节点（自定义机器类型：4 个 vCPU 和 32 GB RAM）
  - 3 个 control plane 节点（自定义机器类型：8 个 vCPU 和 32 GB RAM）
  - 2 个 worker 节点（自定义机器类型：4 个 vCPU 和 16 GB RAM）
- 多个可用区
  - 3 个 infra 节点（自定义机器类型：4 个 vCPU 和 32 GB RAM）

- 3 个 control plane 节点（自定义机器类型：8 个 vCPU 和 32 GB RAM）
- 3 个 worker 节点（自定义机器类型：4 个 vCPU 和 16 GB RAM）

3.5.2. 存储

- 基础架构卷：
  - 300 GB SSD 持久磁盘（在实例删除时删除）
  - 110 GB 标准持久磁盘（实例删除偏移）
- Worker 卷：
  - 300 GB SSD 持久磁盘（在实例删除时删除）
- control plane 卷：
  - 350 GB SSD 持久磁盘（在实例删除时删除）

3.5.3. VPC

- **子网：** 一个用于 control plane 工作负载的 master 子网，以及所有其他 worker 子网。
- **路由器表：**每个 VPC 一个全局路由表。
- **互联网网关：**每个集群一个互联网网关。
- **NAT 网关：**每个集群一个 master NAT 网关和一个 worker NAT 网关。

3.5.4. 服务

在 GCP CCS 集群中必须启用以下服务：

- **deploymentmanager**
- **Compute**
- **cloudapis**
- **cloudresourcemanager**
- **dns**
- **iamcredentials**
- **iam**
- **servicemanagement**
- **serviceusage**
- **storage-api**
- **storage-component**
- **orgpolicy**

- **networksecurity**

3.6. GCP 帐户限值

OpenShift Dedicated 集群使用多个 Google Cloud Platform (GCP) 组件，但默认配额不会影响您安装 OpenShift Dedicated 集群的能力。

标准 OpenShift Dedicated 集群使用以下资源。请注意，有些资源只在 bootstrap 过程中需要，并在集群部署后删除。

表 3.5. 默认集群中使用的 GCP 资源

service	组件	位置	所需的资源总数	bootstrap 后删除的资源
服务帐户	IAM	全局	5	0
防火墙规则	Compute	全局	11	1
转发规则	Compute	全局	2	0
使用的全局 IP 地址	Compute	全局	4	1
健康检查	Compute	全局	3	0
镜像	Compute	全局	1	0
网络	Compute	全局	2	0
静态 IP 地址	Compute	区域	4	1
路由器	Compute	全局	1	0
Routes	Compute	全局	2	0
子网	Compute	全局	2	0
目标池	Compute	全局	3	0
CPU	Compute	区域	28	4
持久性磁盘 SSD (GB)	Compute	区域	896	128

注意

如果在安装过程中任何配额不足，安装程序会显示一个错误信息，包括超过哪个配额，以及显示区域。

请考虑您的集群的实际大小、预定的集群增长以及来自与您的帐户关联的其它集群的使用情况。CPU、静态 IP 地址和持久性磁盘 SSD (Storage) 配额是最可能不足的。

如果您计划在以下区域之一部署集群，您将超过最大存储配额，并可能会超过 CPU 配额限制：

- asia-east2
- asia-northeast2
- asia-south1
- domain-southeast1
- europe-north1
- europe-west2
- europe-west3
- europe-west6
- northamerica-northeast1
- southamerica-east1
- us-west2

您可以从 GCP 控制台 增加资源配额，但可能需要提交一个支持问题单。务必提前规划集群大小，以便在安装 OpenShift Dedicated 集群前有足够的时间解决支持问题单。

3.7. 其他资源

- SRE 访问所需的允许列表 IP 地址