OpenShift Container Platform 4.6

Installing on AWS

Installing OpenShift Container Platform AWS clusters
Installing OpenShift Container Platform AWS clusters
Legal Notice

Copyright © 2022 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.

Abstract

This document provides instructions for installing and uninstalling OpenShift Container Platform clusters on Amazon Web Services.
1.4.10. Telemetry access for OpenShift Container Platform
1.4.11. Next steps

1.5. INSTALLING A CLUSTER ON AWS WITH NETWORK CUSTOMIZATIONS

1.5.1. Prerequisites
1.5.2. Internet access for OpenShift Container Platform
1.5.3. Generating an SSH private key and adding it to the agent
1.5.4. Obtaining the installation program
1.5.5. Network configuration phases
1.5.6. Creating the installation configuration file
1.5.6.1. Installation configuration parameters
1.5.6.1.1. Required configuration parameters
1.5.6.1.2. Network configuration parameters
1.5.6.1.3. Optional configuration parameters
1.5.6.1.4. Optional AWS configuration parameters
1.5.6.2. Sample customized install-config.yaml file for AWS
1.5.6.3. Configuring the cluster-wide proxy during installation
1.5.7. Cluster Network Operator configuration
1.5.7.1. Cluster Network Operator configuration object
   defaultNetwork object configuration
   Configuration for the OpenShift SDN CNI cluster network provider
   Configuration for the OVN-Kubernetes CNI cluster network provider
   kubeProxyConfig object configuration
1.5.8. Specifying advanced network configuration
1.5.9. Configuring an Ingress Controller Network Load Balancer on a new AWS cluster
1.5.10. Configuring hybrid networking with OVN-Kubernetes
1.5.11. Deploying the cluster
1.5.12. Installing the OpenShift CLI by downloading the binary
1.5.12.1. Installing the OpenShift CLI on Linux
1.5.12.2. Installing the OpenShift CLI on Windows
1.5.12.3. Installing the OpenShift CLI on macOS
1.5.13. Logging in to the cluster by using the CLI
1.5.14. Logging in to the cluster by using the web console
1.5.15. Telemetry access for OpenShift Container Platform
1.5.16. Next steps

1.6. INSTALLING A CLUSTER ON AWS INTO AN EXISTING VPC

1.6.1. Prerequisites
1.6.2. About using a custom VPC
1.6.2.1. Requirements for using your VPC
1.6.2.2. VPC validation
1.6.2.3. Division of permissions
1.6.2.4. Isolation between clusters
1.6.3. Internet access for OpenShift Container Platform
1.6.4. Generating an SSH private key and adding it to the agent
1.6.5. Obtaining the installation program
1.6.6. Creating the installation configuration file
1.6.6.1. Installation configuration parameters
1.6.6.1.1. Required configuration parameters
1.6.6.1.2. Network configuration parameters
1.6.6.1.3. Optional configuration parameters
1.6.6.1.4. Optional AWS configuration parameters
1.6.6.2. Sample customized install-config.yaml file for AWS
1.6.6.3. Configuring the cluster-wide proxy during installation
1.6.7. Deploying the cluster
1.8.8.1. Installation configuration parameters
1.8.8.1.1. Required configuration parameters
1.8.8.1.2. Network configuration parameters
1.8.8.1.3. Optional configuration parameters
1.8.8.1.4. Optional AWS configuration parameters
1.8.8.2. Sample customized install-config.yaml file for AWS
1.8.8.3. AWS regions without a published RHCOS AMI
1.8.8.4. Uploading a custom RHCOS AMI in AWS
1.8.8.5. Configuring the cluster-wide proxy during installation
1.8.9. Deploying the cluster
1.8.10. Installing the OpenShift CLI by downloading the binary
1.8.10.1. Installing the OpenShift CLI on Linux
1.8.10.2. Installing the OpenShift CLI on Windows
1.8.10.3. Installing the OpenShift CLI on macOS
1.8.11. Logging in to the cluster by using the CLI
1.8.12. Logging in to the cluster by using the web console
1.8.13. Telemetry access for OpenShift Container Platform
1.8.14. Next steps

1.9. INSTALLING A CLUSTER ON AWS IN A RESTRICTED NETWORK WITH USER-PROVISIONED INFRASTRUCTURE
1.9.1. Prerequisites
1.9.2. About installations in restricted networks
1.9.2.1. Additional limits
1.9.3. Internet access for OpenShift Container Platform
1.9.4. Required AWS infrastructure components
1.9.4.1. Cluster machines
1.9.4.2. Other infrastructure components
1.9.4.3. Certificate signing requests management
1.9.4.4. Required AWS permissions
1.9.5. Generating an SSH private key and adding it to the agent
1.9.6. Creating the installation files for AWS
1.9.6.1. Optional: Creating a separate /var partition
1.9.6.2. Creating the installation configuration file
1.9.6.3. Configuring the cluster-wide proxy during installation
1.9.6.4. Creating the Kubernetes manifest and Ignition config files
1.9.7. Extracting the infrastructure name
1.9.8. Creating a VPC in AWS
1.9.8.1. CloudFormation template for the VPC
1.9.9. Creating networking and load balancing components in AWS
1.9.9.1. CloudFormation template for the network and load balancers
1.9.10. Creating security group and roles in AWS
1.9.10.1. CloudFormation template for security objects
1.9.11. RHCOS AMIs for the AWS infrastructure
1.9.12. Creating the bootstrap node in AWS
1.9.12.1. CloudFormation template for the bootstrap machine
1.9.13. Creating the control plane machines in AWS
1.9.13.1. CloudFormation template for control plane machines
1.9.14. Creating the worker nodes in AWS
1.9.15. Initializing the bootstrap sequence on AWS with user-provisioned infrastructure
1.9.16. Logging in to the cluster by using the CLI
1.9.17. Approving the certificate signing requests for your machines
1.9.18. Initial Operator configuration
1.9.18.1. Disabling the default OperatorHub sources
1.9.18.2. Image registry storage configuration
  1.9.18.2.1. Configuring registry storage for AWS with user-provisioned infrastructure
  1.9.18.2.2. Configuring storage for the image registry in non-production clusters
1.9.19. Deleting the bootstrap resources
1.9.20. Creating the Ingress DNS Records
1.9.21. Completing an AWS installation on user-provisioned infrastructure
1.9.22. Logging in to the cluster by using the web console
1.9.23. Telemetry access for OpenShift Container Platform
1.9.24. Additional resources
1.9.25. Next steps

1.10. UNINSTALLING A CLUSTER ON AWS
  1.10.1. Removing a cluster that uses installer-provisioned infrastructure
1. CONFIGURING AN AWS ACCOUNT

Before you can install OpenShift Container Platform, you must configure an Amazon Web Services (AWS) account.

1.1. Configuring Route 53

To install OpenShift Container Platform, the Amazon Web Services (AWS) account you use must have a dedicated public hosted zone in your Route 53 service. This zone must be authoritative for the domain. The Route 53 service provides cluster DNS resolution and name lookup for external connections to the cluster.

Procedure

1. Identify your domain, or subdomain, and registrar. You can transfer an existing domain and registrar or obtain a new one through AWS or another source.

   NOTE
   If you purchase a new domain through AWS, it takes time for the relevant DNS changes to propagate. For more information about purchasing domains through AWS, see Registering Domain Names Using Amazon Route 53 in the AWS documentation.

2. If you are using an existing domain and registrar, migrate its DNS to AWS. See Making Amazon Route 53 the DNS Service for an Existing Domain in the AWS documentation.

3. Create a public hosted zone for your domain or subdomain. See Creating a Public Hosted Zone in the AWS documentation.
   Use an appropriate root domain, such as openshiftcorp.com, or subdomain, such as clusters.openshiftcorp.com.

4. Extract the new authoritative name servers from the hosted zone records. See Getting the Name Servers for a Public Hosted Zone in the AWS documentation.

5. Update the registrar records for the AWS Route 53 name servers that your domain uses. For example, if you registered your domain to a Route 53 service in a different accounts, see the following topic in the AWS documentation: Adding or Changing Name Servers or Glue Records.

6. If you are using a subdomain, add its delegation records to the parent domain. This gives Amazon Route 53 responsibility for the subdomain. Follow the delegation procedure outlined by the DNS provider of the parent domain. See Creating a subdomain that uses Amazon Route 53 as the DNS service without migrating the parent domain in the AWS documentation for an example high level procedure.

1.1.1. Ingress Operator endpoint configuration for AWS Route 53

If you install in either Amazon Web Services (AWS) GovCloud (US) US-West or US-East region, the Ingress Operator uses us-gov-west-1 region for Route 53 and tagging API clients.

The Ingress Operator uses https://tagging.us-gov-west-1.amazonaws.com as the tagging API endpoint if a tagging custom endpoint is configured that includes the string ‘us-gov-east-1’.
For more information on AWS GovCloud (US) endpoints, see the Service Endpoints in the AWS documentation about GovCloud (US).

**IMPORTANT**

Private, disconnected installations are not supported for AWS GovCloud when you install in the **us-gov-east-1** region.

**Example Route 53 configuration**

```yaml
platform:
  aws:
    region: us-gov-west-1
    serviceEndpoints:
      - name: ec2
        url: https://ec2.us-gov-west-1.amazonaws.com
      - name: elasticloadbalancing
        url: https://elasticloadbalancing.us-gov-west-1.amazonaws.com
      - name: route53
        url: https://route53.us-gov.amazonaws.com
      - name: tagging
        url: https://tagging.us-gov-west-1.amazonaws.com
```

1. Route 53 defaults to [https://route53.us-gov.amazonaws.com](https://route53.us-gov.amazonaws.com) for both AWS GovCloud (US) regions.

2. Only the US-West region has endpoints for tagging. Omit this parameter if your cluster is in another region.

**1.1.2. AWS account limits**

The OpenShift Container Platform cluster uses a number of Amazon Web Services (AWS) components, and the default Service Limits affect your ability to install OpenShift Container Platform clusters. If you use certain cluster configurations, deploy your cluster in certain AWS regions, or run multiple clusters from your account, you might need to request additional resources for your AWS account.

The following table summarizes the AWS components whose limits can impact your ability to install and run OpenShift Container Platform clusters.

<table>
<thead>
<tr>
<th>Component</th>
<th>Number of clusters available by default</th>
<th>Default AWS limit</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>platform:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>aws:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>region:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>serviceEndpoints:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ec2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>url:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>elasticloadbalancing:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>url:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>route53:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>url:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tagging:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>url:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Component</td>
<td>Number of clusters available by default</td>
<td>Default AWS limit</td>
<td>Description</td>
</tr>
<tr>
<td>--------------------------</td>
<td>----------------------------------------</td>
<td>-------------------</td>
<td>------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
</tbody>
</table>
| Instance Limits          | Varies                                 | Varies            | By default, each cluster creates the following instances:  
- One bootstrap machine, which is removed after installation  
- Three control plane nodes (also known as the master nodes)  
- Three worker nodes  
These instance type counts are within a new account’s default limit. To deploy more worker nodes, enable autoscaling, deploy large workloads, or use a different instance type, review your account limits to ensure that your cluster can deploy the machines that you need.  
In most regions, the bootstrap and worker machines uses an **m4.large** machines and the control plane machines use **m4.xlarge** instances. In some regions, including all regions that do not support these instance types, **m5.large** and **m5.xlarge** instances are used instead. |
| Elastic IPs (EIPs)       | 0 to 1                                 | 5 EIPs per account | To provision the cluster in a highly available configuration, the installation program creates a public and private subnet for each availability zone within a region. Each private subnet requires a NAT Gateway, and each NAT gateway requires a separate elastic IP. Review the AWS region map to determine how many availability zones are in each region. To take advantage of the default high availability, install the cluster in a region with at least three availability zones. To install a cluster in a region with more than five availability zones, you must increase the EIP limit.  
**IMPORTANT**  
To use the **us-east-1** region, you must increase the EIP limit for your account. |
<p>| Virtual Private Clouds (VPCs) | 5                                      | 5 VPCs per region | Each cluster creates its own VPC. |</p>
<table>
<thead>
<tr>
<th>Component</th>
<th>Number of clusters available by default</th>
<th>Default AWS limit</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elastic Load Balancing (ELB/NLB)</td>
<td>3</td>
<td>20 per region</td>
<td>By default, each cluster creates internal and external network load balancers for the master API server and a single classic elastic load balancer for the router. Deploying more Kubernetes Service objects with type LoadBalancer will create additional load balancers.</td>
</tr>
<tr>
<td>NAT Gateways</td>
<td>5</td>
<td>5 per availability zone</td>
<td>The cluster deploys one NAT gateway in each availability zone.</td>
</tr>
<tr>
<td>Elastic Network Interfaces (ENIs)</td>
<td>At least 12</td>
<td>350 per region</td>
<td>The default installation creates 21 ENIs and an ENI for each availability zone in your region. For example, the us-east-1 region contains six availability zones, so a cluster that is deployed in that zone uses 27 ENIs. Review the AWS region map to determine how many availability zones are in each region. Additional ENIs are created for additional machines and elastic load balancers that are created by cluster usage and deployed workloads.</td>
</tr>
<tr>
<td>VPC Gateway</td>
<td>20</td>
<td>20 per account</td>
<td>Each cluster creates a single VPC Gateway for S3 access.</td>
</tr>
<tr>
<td>S3 buckets</td>
<td>99</td>
<td>100 buckets per account</td>
<td>Because the installation process creates a temporary bucket and the registry component in each cluster creates a bucket, you can create only 99 OpenShift Container Platform clusters per AWS account.</td>
</tr>
<tr>
<td>Security Groups</td>
<td>250</td>
<td>2,500 per account</td>
<td>Each cluster creates 10 distinct security groups.</td>
</tr>
</tbody>
</table>

1.1.3. Required AWS permissions

When you attach the AdministratorAccess policy to the IAM user that you create in Amazon Web Services (AWS), you grant that user all of the required permissions. To deploy all components of an OpenShift Container Platform cluster, the IAM user requires the following permissions:

**Example 1.1. Required EC2 permissions for installation**

- tag:TagResources
- tag:UntagResources
- ec2:AllocateAddress
- ec2:AssociateAddress
- ec2:AuthorizeSecurityGroupEgress
- ec2:AuthorizeSecurityGroupIngress
- ec2:CopyImage
- ec2:CreateNetworkInterface
- ec2:AttachNetworkInterface
- ec2:CreateSecurityGroup
- ec2:CreateTags
- ec2:CreateVolume
- ec2:DeleteSecurityGroup
- ec2:DeleteSnapshot
- ec2:DeleteTags
- ec2:DeregisterImage
- ec2:DescribeAccountAttributes
- ec2:DescribeAddresses
- ec2:DescribeAvailabilityZones
- ec2:DescribeDhcpOptions
- ec2:DescribeImages
- ec2:DescribeInstanceAttribute
- ec2:DescribeInstanceCreditSpecifications
- ec2:DescribeInstances
- ec2:DescribeInternetGateways
- ec2:DescribeKeyPairs
- ec2:DescribeNatGateways
- ec2:DescribeNetworkAcls
- ec2:DescribeNetworkInterfaces
- ec2:DescribePrefixLists
- ec2:DescribeRegions
Example 1.2. Required permissions for creating network resources during installation

- ec2:AssociateDhcpOptions
- ec2:AssociateRouteTable
- ec2:AttachInternetGateway
- ec2:CreateDhcpOptions
- ec2:CreateInternetGateway
- ec2:CreateNatGateway
- ec2:CreateRoute
- ec2:CreateRouteTable
- ec2:CreateSubnet
- ec2:CreateVpc
- `ec2:CreateVpcEndpoint`
- `ec2:ModifySubnetAttribute`
- `ec2:ModifyVpcAttribute`

**NOTE**

If you use an existing VPC, your account does not require these permissions for creating network resources.

**Example 1.3. Required Elastic Load Balancing permissions (ELB) for installation**

- `elasticloadbalancing:AddTags`
- `elasticloadbalancing:ApplySecurityGroupsToLoadBalancer`
- `elasticloadbalancing:AttachLoadBalancerToSubnets`
- `elasticloadbalancing:ConfigureHealthCheck`
- `elasticloadbalancing:CreateLoadBalancer`
- `elasticloadbalancing:CreateLoadBalancerListeners`
- `elasticloadbalancing:DeleteLoadBalancer`
- `elasticloadbalancing:DeregisterInstancesFromLoadBalancer`
- `elasticloadbalancing:DescribeInstanceHealth`
- `elasticloadbalancing:DescribeLoadBalancerAttributes`
- `elasticloadbalancing:DescribeLoadBalancers`
- `elasticloadbalancing:DescribeTags`
- `elasticloadbalancing:ModifyLoadBalancerAttributes`
- `elasticloadbalancing:RegisterInstancesWithLoadBalancer`
- `elasticloadbalancing:SetLoadBalancerPoliciesOfListener`

**Example 1.4. Required Elastic Load Balancing permissions (ELBv2) for installation**

- `elasticloadbalancing:AddTags`
- `elasticloadbalancing:CreateListener`
- `elasticloadbalancing:CreateLoadBalancer`
- `elasticloadbalancing:CreateTargetGroup`
- `elasticloadbalancing:DeleteLoadBalancer`
- elasticloadbalancing:DeregisterTargets
- elasticloadbalancing:DescribeListeners
- elasticloadbalancing:DescribeLoadBalancerAttributes
- elasticloadbalancing:DescribeLoadBalancers
- elasticloadbalancing:DescribeTargetGroupAttributes
- elasticloadbalancing:DescribeTargetHealth
- elasticloadbalancing:ModifyLoadBalancerAttributes
- elasticloadbalancing:ModifyTargetGroup
- elasticloadbalancing:ModifyTargetGroupAttributes
- elasticloadbalancing:RegisterTargets

Example 1.5. Required IAM permissions for installation

- iam:AddRoleTo InstanceProfile
- iam:CreateInstanceProfile
- iam:CreateRole
- iam:DeleteInstanceProfile
- iam:DeleteRole
- iam:DeleteRolePolicy
- iam:GetInstanceProfile
- iam:GetRole
- iam:GetRolePolicy
- iam:GetUser
- iam:ListInstanceProfilesForRole
- iam:ListRoles
- iam:ListUsers
- iam:PassRole
- iam:PutRolePolicy
- iam:RemoveRoleFromInstanceProfile
- iam:SimulatePrincipalPolicy
- iam:TagRole
NOTE

If you have not created an elastic load balancer (ELB) in your AWS account, the IAM user also requires the `iam:CreateServiceLinkedRole` permission.

**Example 1.6. Required Route 53 permissions for installation**
- `route53:ChangeResourceRecordSets`
- `route53:ChangeTagsForResource`
- `route53:CreateHostedZone`
- `route53:DeleteHostedZone`
- `route53:GetChange`
- `route53:GetHostedZone`
- `route53:ListHostedZones`
- `route53:ListHostedZonesByName`
- `route53:ListResourceRecordSets`
- `route53:ListTagsForResource`
- `route53:UpdateHostedZoneComment`

**Example 1.7. Required S3 permissions for installation**
- `s3:CreateBucket`
- `s3:DeleteBucket`
- `s3:GetAccelerateConfiguration`
- `s3:GetBucketAcl`
- `s3:GetBucketCors`
- `s3:GetBucketLocation`
- `s3:GetBucketLogging`
- `s3:GetBucketObjectLockConfiguration`
- `s3:GetBucketReplication`
- `s3:GetBucketRequestPayment`
- `s3:GetBucketTagging`
- `s3:GetBucketVersioning`
• s3:GetBucketWebsite
• s3:GetEncryptionConfiguration
• s3:GetLifecycleConfiguration
• s3:GetReplicationConfiguration
• s3:ListBucket
• s3:PutBucketAcl
• s3:PutBucketTagging
• s3:PutEncryptionConfiguration

Example 1.8. S3 permissions that cluster Operators require

• s3:DeleteObject
• s3:GetObject
• s3:GetObjectAcl
• s3:GetObjectTagging
• s3:GetObjectVersion
• s3:PutObject
• s3:PutObjectAcl
• s3:PutObjectTagging

Example 1.9. Required permissions to delete base cluster resources

• autoscaling:DescribeAutoScalingGroups
• ec2:DeleteNetworkInterface
• ec2:DeleteVolume
• elasticloadbalancing:DeleteTargetGroup
• elasticloadbalancing:DescribeTargetGroups
• iam:DeleteAccessKey
• iam:DeleteUser
• iam:ListAttachedRolePolicies
• iam:ListInstanceProfiles
• iam:ListRolePolicies
Example 1.10. Required permissions to delete network resources

- ec2:DeleteDhcpOptions
- ec2:DeleteInternetGateway
- ec2:DeleteNatGateway
- ec2:DeleteRoute
- ec2:DeleteRouteTable
- ec2:DeleteSubnet
- ec2:DeleteVpc
- ec2:DeleteVpcEndpoints
- ec2:DetachInternetGateway
- ec2:DisassociateRouteTable
- ec2:ReplaceRouteTableAssociation

NOTE
If you use an existing VPC, your account does not require these permissions to delete network resources.

Example 1.11. Additional IAM and S3 permissions that are required to create manifests

- iam:DeleteAccessKey
- iam:DeleteUser
- iam:DeleteUserPolicy
- iam:GetUserPolicy
- iam:ListAccessKeys
- iam:PutUserPolicy
- iam:TagUser
- iam:GetUserPolicy
• `iam:ListAccessKeys`
• `s3:PutBucketPublicAccessBlock`
• `s3:GetBucketPublicAccessBlock`
• `s3:PutLifecycleConfiguration`
• `s3:HeadBucket`
• `s3:ListBucketMultipartUploads`
• `s3:AbortMultipartUpload`

**NOTE**

If you are managing your cloud provider credentials with mint mode, the IAM user also requires the `iam:CreateAccessKey` and `iam:CreateUser` permissions.

**Example 1.12. Optional permission for quota checks for installation**

• `servicequotas:ListAWSDefaultServiceQuotas`

### 1.1.4. Creating an IAM user

Each Amazon Web Services (AWS) account contains a root user account that is based on the email address you used to create the account. This is a highly-privileged account, and it is recommended to use it for only initial account and billing configuration, creating an initial set of users, and securing the account.

Before you install OpenShift Container Platform, create a secondary IAM administrative user. As you complete the Creating an IAM User in Your AWS Account procedure in the AWS documentation, set the following options:

**Procedure**

1. Specify the IAM user name and select **Programmatic access**.

2. Attach the **AdministratorAccess** policy to ensure that the account has sufficient permission to create the cluster. This policy provides the cluster with the ability to grant credentials to each OpenShift Container Platform component. The cluster grants the components only the credentials that they require.

**NOTE**

While it is possible to create a policy that grants the all of the required AWS permissions and attach it to the user, this is not the preferred option. The cluster will not have the ability to grant additional credentials to individual components, so the same credentials are used by all components.

3. Optional: Add metadata to the user by attaching tags.
4. Confirm that the user name that you specified is granted the **AdministratorAccess** policy.

5. Record the access key ID and secret access key values. You must use these values when you configure your local machine to run the installation program.

**IMPORTANT**

You cannot use a temporary session token that you generated while using a multi-factor authentication device to authenticate to AWS when you deploy a cluster. The cluster continues to use your current AWS credentials to create AWS resources for the entire life of the cluster, so you must use key-based, long-lived credentials.

Additional resources

- See [Manually creating IAM for AWS](#) for steps to set the Cloud Credential Operator (CCO) to manual mode prior to installation. Use this mode in environments where the cloud identity and access management (IAM) APIs are not reachable, or if you prefer not to store an administrator-level credential secret in the cluster `kube-system` project.

### 1.1.5. Supported AWS regions

You can deploy an OpenShift Container Platform cluster to the following public regions:

- **af-south-1** (Cape Town)
- **ap-east-1** (Hong Kong)
- **ap-northeast-1** (Tokyo)
- **ap-northeast-2** (Seoul)
- **ap-northeast-3** (Osaka)
- **ap-south-1** (Mumbai)
- **ap-southeast-1** (Singapore)
- **ap-southeast-2** (Sydney)
- **ca-central-1** (Central)
- **eu-central-1** (Frankfurt)
- **eu-north-1** (Stockholm)
- **eu-south-1** (Milan)
- **eu-west-1** (Ireland)
- **eu-west-2** (London)
- **eu-west-3** (Paris)
- **me-south-1** (Bahrain)
The following AWS GovCloud regions are supported:

- us-gov-west-1
- us-gov-east-1

1.1.6. Next steps

- Install an OpenShift Container Platform cluster:
  - Quickly install a cluster with default options on installer-provisioned infrastructure
  - Install a cluster with cloud customizations on installer-provisioned infrastructure
  - Install a cluster with network customizations on installer-provisioned infrastructure
  - Installing a cluster on user-provisioned infrastructure in AWS by using CloudFormation templates

1.2. MANUALLY CREATING IAM FOR AWS

In environments where the cloud identity and access management (IAM) APIs are not reachable, or the administrator prefers not to store an administrator-level credential secret in the cluster kube-system namespace, you can put the Cloud Credential Operator (CCO) into manual mode before you install the cluster.

1.2.1. Alternatives to storing administrator-level secrets in the kube-system project

The Cloud Credential Operator (CCO) manages cloud provider credentials as Kubernetes custom resource definitions (CRDs). You can configure the CCO to suit the security requirements of your organization by setting different values for the credentialsMode parameter in the install-config.yaml file.

If you prefer not to store an administrator-level credential secret in the cluster kube-system project, you can choose one of the following options when installing OpenShift Container Platform:

- Manage cloud credentials manually
  You can set the credentialsMode parameter for the CCO to Manual to manage cloud credentials manually. Using manual mode allows each cluster component to have only the permissions it requires, without storing an administrator-level credential in the cluster. You can also use this mode if your environment does not have connectivity to the cloud provider public IAM endpoint. However, you must manually reconcile permissions with new release images for every upgrade. You must also manually supply credentials for every component that requests them.
Remove the administrator-level credential secret after installing OpenShift Container Platform with mint mode:

If you are using the CCO with the credentialsMode parameter set to Mint, you can remove or rotate the administrator-level credential after installing OpenShift Container Platform. Mint mode is the default configuration for the CCO. This option requires the presence of the administrator-level credential during an installation. The administrator-level credential is used during the installation to mint other credentials with some permissions granted. The original credential secret is not stored in the cluster permanently.

**NOTE**

Prior to a non z-stream upgrade, you must reinstate the credential secret with the administrator-level credential. If the credential is not present, the upgrade might be blocked.

**Additional resources**

To learn how to rotate or remove the administrator-level credential secret after installing OpenShift Container Platform, see Rotating or removing cloud provider credentials.

For a detailed description of all available CCO credential modes and their supported platforms, see the Cloud Credential Operator reference.

**1.2.2. Manually create IAM**

The Cloud Credential Operator (CCO) can be put into manual mode prior to installation in environments where the cloud identity and access management (IAM) APIs are not reachable, or the administrator prefers not to store an administrator-level credential secret in the cluster kube-system namespace.

**Procedure**

1. Change to the directory that contains the installation program and create the install-config.yaml file:

   ```bash
   $ openshift-install create install-config --dir <installation_directory>
   ```

2. Edit the install-config.yaml configuration file so that it contains the credentialsMode parameter set to Manual.

**Example install-config.yaml configuration file**

```yaml
apiVersion: v1
baseDomain: cluster1.example.com
credentialsMode: Manual
compute:
  - architecture: amd64
    hyperthreading: Enabled
...
```

1 This line is added to set the credentialsMode parameter to Manual.
3. To generate the manifests, run the following command from the directory that contains the installation program:

```bash
$ openshift-install create manifests --dir <installation_directory>
```

For `<installation_directory>`, specify the directory name to store the files that the installation program creates.

4. Remove the `admin` credential secret created using your local cloud credentials. This removal prevents your `admin` credential from being stored in the cluster:

```bash
$ rm mycluster/openshift/99_cloud-creds-secret.yaml
```

5. From the directory that contains the installation program, obtain details of the OpenShift Container Platform release image that your `openshift-install` binary is built to use:

```bash
$ openshift-install version
```

**Example output**

```
release image quay.io/openshift-release-dev/ocp-release:4.y.z-x86_64
```

6. Locate all `CredentialsRequest` objects in this release image that target the cloud you are deploying on:

```bash
$ oc adm release extract quay.io/openshift-release-dev/ocp-release:4.y.z-x86_64 --credentials-requests --cloud=aws
```

This displays the details for each request.

**Sample `CredentialsRequest` object**

```yaml
apiVersion: cloudcredential.openshift.io/v1
kind: CredentialsRequest
metadata:
  name: cloud-credential-operator-iam-ro
  namespace: openshift-cloud-credential-operator
spec:
  secretRef:
    name: cloud-credential-operator-iam-ro-creds
    namespace: openshift-cloud-credential-operator
  providerSpec:
    apiVersion: cloudcredential.openshift.io/v1
    kind: AWSProviderSpec
    statementEntries:
      - effect: Allow
        action: - iam:GetUser
        - iam:GetUserPolicy
        - iam:ListAccessKeys
        resource: "*"
```

7. Create YAML files for secrets in the `openshift-install` manifests directory that you generated
previously. The secrets must be stored using the namespace and secret name defined in the `spec.secretRef` for each `credentialsRequest`. The format for the secret data varies for each cloud provider.

8. From the directory that contains the installation program, proceed with your cluster creation:

   ```
   $ openshift-install create cluster --dir <installation_directory>
   ```

**IMPORTANT**

Before upgrading a cluster that uses manually maintained credentials, you must ensure that the CCO is in an upgradeable state. For details, see the Upgrading clusters with manually maintained credentials section of the installation content for your cloud provider.

### 1.2.3. Admin credentials root secret format

Each cloud provider uses a credentials root secret in the `kube-system` namespace by convention, which is then used to satisfy all credentials requests and create their respective secrets. This is done either by minting new credentials, with `mint mode`, or by copying the credentials root secret, with `passthrough mode`.

The format for the secret varies by cloud, and is also used for each `CredentialsRequest` secret.

**Amazon Web Services (AWS) secret format**

```yaml
apiVersion: v1
kind: Secret
metadata:
  namespace: kube-system
  name: aws-creds
stringData:
  aws_access_key_id: <AccessKeyId>
  aws_secret_access_key: <SecretAccessKey>
```

### 1.2.4. Upgrading clusters with manually maintained credentials

If credentials are added in a future release, the Cloud Credential Operator (CCO) `upgradable` status for a cluster with manually maintained credentials changes to `false`. For minor release, for example, from 4.5 to 4.6, this status prevents you from upgrading until you have addressed any updated permissions. For z-stream releases, for example, from 4.5.10 to 4.5.11, the upgrade is not blocked, but the credentials must still be updated for the new release.

Use the Administrator perspective of the web console to determine if the CCO is upgradeable.

1. Navigate to Administration → Cluster Settings.

2. To view the CCO status details, click cloud-credential in the Cluster Operators list.

3. If the Upgradeable status in the Conditions section is False, examine the credentialsRequests for the new release and update the manually maintained credentials on your cluster to match before upgrading.

In addition to creating new credentials for the release image that you are upgrading to, you must review
the required permissions for existing credentials and accommodate any new permissions requirements for existing components in the new release. The CCO cannot detect these mismatches and will not set `upgradable` to `false` in this case.

The *Manually creating IAM* section of the installation content for your cloud provider explains how to obtain and use the credentials required for your cloud.

### 1.2.5. Mint mode

Mint mode is the default and recommended Cloud Credential Operator (CCO) credentials mode for OpenShift Container Platform. In this mode, the CCO uses the provided administrator-level cloud credential to run the cluster. Mint mode is supported for AWS, GCP, and Azure.

In mint mode, the `admin` credential is stored in the `kube-system` namespace and then used by the CCO to process the `CredentialsRequest` objects in the cluster and create users for each with specific permissions.

The benefits of mint mode include:

- Each cluster component has only the permissions it requires
- Automatic, on-going reconciliation for cloud credentials, including additional credentials or permissions that might be required for upgrades

One drawback is that mint mode requires `admin` credential storage in a cluster `kube-system` secret.

### 1.2.6. Mint Mode with removal or rotation of the admin credential

Currently, this mode is only supported on AWS.

In this mode, a user installs OpenShift Container Platform with an `admin` credential just like the normal mint mode. However, this mode removes the `admin` credential secret from the cluster post-installation.

The administrator can have the Cloud Credential Operator make its own request for a read-only credential that allows it to verify if all `CredentialsRequest` objects have their required permissions, thus the `admin` credential is not required unless something needs to be changed. After the associated credential is removed, it can be destroyed on the underlying cloud, if desired.

Prior to upgrade, the `admin` credential should be restored. In the future, upgrade might be blocked if the credential is not present.

The `admin` credential is not stored in the cluster permanently.

This mode still requires the `admin` credential in the cluster for brief periods of time. It also requires manually re-instating the secret with `admin` credentials for each upgrade.

### 1.2.7. Next steps

- Install an OpenShift Container Platform cluster:
  - Installing a cluster quickly on AWS with default options on installer-provisioned infrastructure
  - Install a cluster with cloud customizations on installer-provisioned infrastructure
  - Install a cluster with network customizations on installer-provisioned infrastructure
1.3. INSTALLING A CLUSTER QUICKLY ON AWS

In OpenShift Container Platform version 4.6, you can install a cluster on Amazon Web Services (AWS) that uses the default configuration options.

1.3.1. Prerequisites

- Review details about the OpenShift Container Platform installation and update processes.
- Configure an AWS account to host the cluster.

**IMPORTANT**

If you have an AWS profile stored on your computer, it must not use a temporary session token that you generated while using a multi-factor authentication device. The cluster continues to use your current AWS credentials to create AWS resources for the entire life of the cluster, so you must use key-based, long-lived credentials. To generate appropriate keys, see Managing Access Keys for IAM Users in the AWS documentation. You can supply the keys when you run the installation program.

- If you use a firewall, you must configure it to allow the sites that your cluster requires access to.
- If you do not allow the system to manage identity and access management (IAM), then a cluster administrator can manually create and maintain IAM credentials. Manual mode can also be used in environments where the cloud IAM APIs are not reachable.

1.3.2. Internet access for OpenShift Container Platform

In OpenShift Container Platform 4.6, you require access to the Internet to install your cluster.

You must have Internet access to:

- Access the Red Hat OpenShift Cluster Manager page to download the installation program and perform subscription management. If the cluster has Internet access and you do not disable Telemetry, that service automatically entitles your cluster.

- Access Quay.io to obtain the packages that are required to install your cluster.

- Obtain the packages that are required to perform cluster updates.

**IMPORTANT**

If your cluster cannot have direct Internet access, you can perform a restricted network installation on some types of infrastructure that you provision. During that process, you download the content that is required and use it to populate a mirror registry with the packages that you need to install a cluster and generate the installation program. With some installation types, the environment that you install your cluster in will not require Internet access. Before you update the cluster, you update the content of the mirror registry.
1.3.3. Generating an SSH private key and adding it to the agent

If you want to perform installation debugging or disaster recovery on your cluster, you must provide an SSH key to both your ssh-agent and the installation program. You can use this key to access the bootstrap machine in a public cluster to troubleshoot installation issues.

**NOTE**

In a production environment, you require disaster recovery and debugging.

You can use this key to SSH into the master nodes as the user core. When you deploy the cluster, the key is added to the core user’s ~/.ssh/authorized_keys list.

**NOTE**

You must use a local key, not one that you configured with platform-specific approaches such as AWS key pairs.

**Procedure**

1. If you do not have an SSH key that is configured for password-less authentication on your computer, create one. For example, on a computer that uses a Linux operating system, run the following command:

   ```bash
   $ ssh-keygen -t ed25519 -N '' -f <path>/<file_name>
   ```

   Specifying the path and file name, such as ~/.ssh/id_rsa, of the new SSH key. If you have an existing key pair, ensure your public key is in the your ~/.ssh directory.

   Running this command generates an SSH key that does not require a password in the location that you specified.

   **NOTE**

   If you plan to install an OpenShift Container Platform cluster that uses FIPS Validated / Modules in Process cryptographic libraries on the x86_64 architecture, do not create a key that uses the ed25519 algorithm. Instead, create a key that uses the rsa or ecdsa algorithm.

2. Start the ssh-agent process as a background task:

   ```bash
   $ eval "$(ssh-agent -s)"
   ```

**Example output**

Agent pid 31874
NOTE

If your cluster is in FIPS mode, only use FIPS-compliant algorithms to generate the SSH key. The key must be either RSA or ECDSA.

3. Add your SSH private key to the **ssh-agent**:

\[
$ \text{ssh-add <path>/<file_name>}
\]

**Example output**

```
Identity added: /home/<you>/<path>/<file_name> (<computer_name>)
```

Specify the path and file name for your SSH private key, such as `~/.ssh/id_rsa`

**Next steps**

- When you install OpenShift Container Platform, provide the SSH public key to the installation program.

### 1.3.4. Obtaining the installation program

Before you install OpenShift Container Platform, download the installation file on a local computer.

**Prerequisites**

- You have a computer that runs Linux or macOS, with 500 MB of local disk space

**Procedure**

1. Access the **Infrastructure Provider** page on the Red Hat OpenShift Cluster Manager site. If you have a Red Hat account, log in with your credentials. If you do not, create an account.

2. Select your infrastructure provider.

3. Navigate to the page for your installation type, download the installation program for your operating system, and place the file in the directory where you will store the installation configuration files.

**IMPORTANT**

The installation program creates several files on the computer that you use to install your cluster. You must keep the installation program and the files that the installation program creates after you finish installing the cluster. Both files are required to delete the cluster.

**IMPORTANT**

Deleting the files created by the installation program does not remove your cluster, even if the cluster failed during installation. To remove your cluster, complete the OpenShift Container Platform uninstallation procedures for your specific cloud provider.
4. Extract the installation program. For example, on a computer that uses a Linux operating system, run the following command:

   ```
   $ tar xvf openshift-install-linux.tar.gz
   ```

5. From the Pull Secret page on the Red Hat OpenShift Cluster Manager site, download your installation pull secret as a `.txt` file. This pull secret allows you to authenticate with the services that are provided by the included authorities, including Quay.io, which serves the container images for OpenShift Container Platform components.

### 1.3.5. Deploying the cluster

You can install OpenShift Container Platform on a compatible cloud platform.

**IMPORTANT**

You can run the `create cluster` command of the installation program only once, during initial installation.

**Prerequisites**

- Configure an account with the cloud platform that hosts your cluster.
- Obtain the OpenShift Container Platform installation program and the pull secret for your cluster.

**Procedure**

1. Change to the directory that contains the installation program and initialize the cluster deployment:

   ```
   $ ./openshift-install create cluster --dir <installation_directory> \  
   --log-level=info
   ```

   **1** For `<installation_directory>`, specify the directory name to store the files that the installation program creates.
   
   **2** To view different installation details, specify `warn`, `debug`, or `error` instead of `info`.

   **IMPORTANT**

   Specify an empty directory. Some installation assets, like bootstrap X.509 certificates have short expiration intervals, so you must not reuse an installation directory. If you want to reuse individual files from another cluster installation, you can copy them into your directory. However, the file names for the installation assets might change between releases. Use caution when copying installation files from an earlier OpenShift Container Platform version.

   Provide values at the prompts:
   
   a. Optional: Select an SSH key to use to access your cluster machines.
b. Select **aws** as the platform to target.

c. If you do not have an Amazon Web Services (AWS) profile stored on your computer, enter the AWS access key ID and secret access key for the user that you configured to run the installation program.

**NOTE**

The AWS access key ID and secret access key are stored in `~/.aws/credentials` in the home directory of the current user on the installation host. You are prompted for the credentials by the installation program if the credentials for the exported profile are not present in the file. Any credentials that you provide to the installation program are stored in the file.

d. Select the AWS region to deploy the cluster to.

e. Select the base domain for the Route 53 service that you configured for your cluster.

f. Enter a descriptive name for your cluster.

g. Paste the pull secret that you obtained from the **Pull Secret** page on the Red Hat OpenShift Cluster Manager site.

**NOTE**

If the cloud provider account that you configured on your host does not have sufficient permissions to deploy the cluster, the installation process stops, and the missing permissions are displayed.

When the cluster deployment completes, directions for accessing your cluster, including a link to its web console and credentials for the **kubeadmin** user, display in your terminal.

**Example output**

```
... INFO Install complete!
INFO To access the cluster as the system:admin user when using 'oc', run 'export KUBECONFIG=/home/myuser/install_dir/auth/kubeconfig'
INFO Access the OpenShift web-console here: https://console-openshift-console.apps.mycluster.example.com
INFO Login to the console with user: "kubeadmin", and password: "4vYBz-Ee6gm-ymBZj-Wt5AL"
INFO Time elapsed: 36m22s
```
The cluster access and credential information also outputs to `<installation_directory>/openshift_install.log` when an installation succeeds.

**IMPORTANT**

The Ignition config files that the installation program generates contain certificates that expire after 24 hours, which are then renewed at that time. If the cluster is shut down before renewing the certificates and the cluster is later restarted after the 24 hours have elapsed, the cluster automatically recovers the expired certificates. The exception is that you must manually approve the pending `node-bootstrapper` certificate signing requests (CSRs) to recover kubelet certificates. See the documentation for *Recovering from expired control plane certificates* for more information.

**IMPORTANT**

You must not delete the installation program or the files that the installation program creates. Both are required to delete the cluster.

2. Optional: Remove or disable the `AdministratorAccess` policy from the IAM account that you used to install the cluster.

**NOTE**

The elevated permissions provided by the `AdministratorAccess` policy are required only during installation.

**Additional resources**

- See *Configuration and credential file settings* in the AWS documentation for more information about AWS profile and credential configuration.

### 1.3.6. Installing the OpenShift CLI by downloading the binary

You can install the OpenShift CLI (oc) in order to interact with OpenShift Container Platform from a command-line interface. You can install oc on Linux, Windows, or macOS.

**IMPORTANT**

If you installed an earlier version of oc, you cannot use it to complete all of the commands in OpenShift Container Platform 4.6. Download and install the new version of oc.

#### 1.3.6.1. Installing the OpenShift CLI on Linux

You can install the OpenShift CLI (oc) binary on Linux by using the following procedure.

**Procedure**

2. Select the appropriate version in the Version drop-down menu.

3. Click Download Now next to the OpenShift v4.6 Linux Client entry and save the file.

4. Unpack the archive:
   
   ```
   $ tar xvzf <file>
   ```

5. Place the oc binary in a directory that is on your PATH.
   To check your PATH, execute the following command:
   
   ```
   $ echo $PATH
   ```

After you install the OpenShift CLI, it is available using the oc command:

```
$ oc <command>
```

1.3.6.2. Installing the OpenShift CLI on Windows

You can install the OpenShift CLI (oc) binary on Windows by using the following procedure.

Procedure


2. Select the appropriate version in the Version drop-down menu.

3. Click Download Now next to the OpenShift v4.6 Windows Client entry and save the file.

4. Unzip the archive with a ZIP program.

5. Move the oc binary to a directory that is on your PATH.
   To check your PATH, open the command prompt and execute the following command:
   
   ```
   C:\> path
   ```

After you install the OpenShift CLI, it is available using the oc command:

```
C:\> oc <command>
```

1.3.6.3. Installing the OpenShift CLI on macOS

You can install the OpenShift CLI (oc) binary on macOS by using the following procedure.

Procedure


2. Select the appropriate version in the Version drop-down menu.

3. Click Download Now next to the OpenShift v4.6 MacOSX Client entry and save the file.
4. Unpack and unzip the archive.

5. Move the `oc` binary to a directory on your PATH.
   To check your PATH, open a terminal and execute the following command:

   ```
   $ echo $PATH
   ```

   After you install the OpenShift CLI, it is available using the `oc` command:

   ```
   $ oc <command>
   ```

### 1.3.7. Logging in to the cluster by using the CLI

You can log in to your cluster as a default system user by exporting the cluster `kubeconfig` file. The `kubeconfig` file contains information about the cluster that is used by the CLI to connect a client to the correct cluster and API server. The file is specific to a cluster and is created during OpenShift Container Platform installation.

**Prerequisites**

- You deployed an OpenShift Container Platform cluster.
- You installed the `oc` CLI.

**Procedure**

1. Export the `kubeadmin` credentials:

   ```
   $ export KUBECONFIG=<installation_directory>/auth/kubeconfig
   ```

   [1] For `<installation_directory>`, specify the path to the directory that you stored the installation files in.

2. Verify you can run `oc` commands successfully using the exported configuration:

   ```
   $ oc whoami
   ```

   **Example output**

   ```
   system:admin
   ```

### 1.3.8. Logging in to the cluster by using the web console

The `kubeadmin` user exists by default after an OpenShift Container Platform installation. You can log in to your cluster as the `kubeadmin` user by using the OpenShift Container Platform web console.

**Prerequisites**

- You have access to the installation host.
- You completed a cluster installation and all cluster Operators are available.
Procedure

1. Obtain the password for the **kubeadmin** user from the **kubeadmin-password** file on the installation host:

   ```bash
   $ cat <installation_directory>/auth/kubeadmin-password
   ``

   **NOTE**
   Alternatively, you can obtain the **kubeadmin** password from the `<installation_directory>/openshift_install.log` log file on the installation host.

2. List the OpenShift Container Platform web console route:

   ```bash
   $ oc get routes -n openshift-console | grep 'console-openshift'
   ``

   **NOTE**
   Alternatively, you can obtain the OpenShift Container Platform route from the `<installation_directory>/openshift_install.log` log file on the installation host.

   **Example output**

   ```bash
   console   console-openshift-console.apps.<cluster_name>.<base_domain>    console
   https   reencrypt/Redirect   None
   ```

3. Navigate to the route detailed in the output of the preceding command in a web browser and log in as the **kubeadmin** user.

Additional resources

- See [Accessing the web console](#) for more details about accessing and understanding the OpenShift Container Platform web console.

### 1.3.9. Telemetry access for OpenShift Container Platform

In OpenShift Container Platform 4.6, the Telemetry service, which runs by default to provide metrics about cluster health and the success of updates, requires Internet access. If your cluster is connected to the Internet, Telemetry runs automatically, and your cluster is registered to the **Red Hat OpenShift Cluster Manager (OCM)**.

After you confirm that your Red Hat OpenShift Cluster Manager inventory is correct, either maintained automatically by Telemetry or manually by using OCM, **use subscription watch** to track your OpenShift Container Platform subscriptions at the account or multi-cluster level.

Additional resources

- See [About remote health monitoring](#) for more information about the Telemetry service.

### 1.3.10. Next steps

- Validating an installation.
- Customize your cluster.
- If necessary, you can opt out of remote health reporting.
- If necessary, you can remove cloud provider credentials.

1.4. INSTALLING A CLUSTER ON AWS WITH CUSTOMIZATIONS

In OpenShift Container Platform version 4.6, you can install a customized cluster on infrastructure that the installation program provisions on Amazon Web Services (AWS). To customize the installation, you modify parameters in the `install-config.yaml` file before you install the cluster.

1.4.1. Prerequisites

- Review details about the OpenShift Container Platform installation and update processes.
- Configure an AWS account to host the cluster.

**IMPORTANT**

If you have an AWS profile stored on your computer, it must not use a temporary session token that you generated while using a multi-factor authentication device. The cluster continues to use your current AWS credentials to create AWS resources for the entire life of the cluster, so you must use long-lived credentials. To generate appropriate keys, see Managing Access Keys for IAM Users in the AWS documentation. You can supply the keys when you run the installation program.

- If you use a firewall, you must configure it to allow the sites that your cluster requires access to.
- If you do not allow the system to manage identity and access management (IAM), then a cluster administrator can manually create and maintain IAM credentials. Manual mode can also be used in environments where the cloud IAM APIs are not reachable.

1.4.2. Internet access for OpenShift Container Platform

In OpenShift Container Platform 4.6, you require access to the Internet to install your cluster.

You must have Internet access to:

- Access the Red Hat OpenShift Cluster Manager page to download the installation program and perform subscription management. If the cluster has Internet access and you do not disable Telemetry, that service automatically entitles your cluster.
- Access Quay.io to obtain the packages that are required to install your cluster.
- Obtain the packages that are required to perform cluster updates.
IMPORTANT

If your cluster cannot have direct Internet access, you can perform a restricted network installation on some types of infrastructure that you provision. During that process, you download the content that is required and use it to populate a mirror registry with the packages that you need to install a cluster and generate the installation program. With some installation types, the environment that you install your cluster in will not require Internet access. Before you update the cluster, you update the content of the mirror registry.

1.4.3. Generating an SSH private key and adding it to the agent

If you want to perform installation debugging or disaster recovery on your cluster, you must provide an SSH key to both your ssh-agent and the installation program. You can use this key to access the bootstrap machine in a public cluster to troubleshoot installation issues.

NOTE

In a production environment, you require disaster recovery and debugging.

You can use this key to SSH into the master nodes as the user core. When you deploy the cluster, the key is added to the core user’s ~/.ssh/authorized_keys list.

NOTE

You must use a local key, not one that you configured with platform-specific approaches such as AWS key pairs.

Procedure

1. If you do not have an SSH key that is configured for password-less authentication on your computer, create one. For example, on a computer that uses a Linux operating system, run the following command:

   ```
   $ ssh-keygen -t ed25519 -N "\n   -f <path>/<file_name>
   ```

   Specify the path and file name, such as ~/.ssh/id_rsa, of the new SSH key. If you have an existing key pair, ensure your public key is in the your ~/.ssh directory.

   Running this command generates an SSH key that does not require a password in the location that you specified.

   NOTE

   If you plan to install an OpenShift Container Platform cluster that uses FIPS Validated / Modules in Process cryptographic libraries on the x86_64 architecture, do not create a key that uses the ed25519 algorithm. Instead, create a key that uses the rsa or ecdsa algorithm.

2. Start the ssh-agent process as a background task:
If your cluster is in FIPS mode, only use FIPS-compliant algorithms to generate the SSH key. The key must be either RSA or ECDSA.

3. Add your SSH private key to the `ssh-agent`:

   ```bash
   $ ssh-add <path>/<file_name>
   ```

   **Example output**

   ```
   Identity added: /home/<you>/<path>/<file_name> (<computer_name>)
   ```

   Specify the path and file name for your SSH private key, such as `~/.ssh/id_rsa`

**Next steps**

- When you install OpenShift Container Platform, provide the SSH public key to the installation program.

### 1.4.4. Obtaining the installation program

Before you install OpenShift Container Platform, download the installation file on a local computer.

**Prerequisites**

- You have a computer that runs Linux or macOS, with 500 MB of local disk space

**Procedure**

1. Access the [Infrastructure Provider](#) page on the Red Hat OpenShift Cluster Manager site. If you have a Red Hat account, log in with your credentials. If you do not, create an account.

2. Select your infrastructure provider.

3. Navigate to the page for your installation type, download the installation program for your operating system, and place the file in the directory where you will store the installation configuration files.

**IMPORTANT**

The installation program creates several files on the computer that you use to install your cluster. You must keep the installation program and the files that the installation program creates after you finish installing the cluster. Both files are required to delete the cluster.
Deleting the files created by the installation program does not remove your cluster, even if the cluster failed during installation. To remove your cluster, complete the OpenShift Container Platform uninstallation procedures for your specific cloud provider.

4. Extract the installation program. For example, on a computer that uses a Linux operating system, run the following command:

$ tar xvf openshift-install-linux.tar.gz

5. From the Pull Secret page on the Red Hat OpenShift Cluster Manager site, download your installation pull secret as a .txt file. This pull secret allows you to authenticate with the services that are provided by the included authorities, including Quay.io, which serves the container images for OpenShift Container Platform components.

1.4.5. Creating the installation configuration file

You can customize the OpenShift Container Platform cluster you install on Amazon Web Services (AWS).

Prerequisites

- Obtain the OpenShift Container Platform installation program and the pull secret for your cluster.

Procedure

1. Create the install-config.yaml file.

   a. Change to the directory that contains the installation program and run the following command:

   $ ./openshift-install create install-config --dir <installation_directory>

   1 For <installation_directory>, specify the directory name to store the files that the installation program creates.

   IMPORTANT

   Specify an empty directory. Some installation assets, like bootstrap X.509 certificates have short expiration intervals, so you must not reuse an installation directory. If you want to reuse individual files from another cluster installation, you can copy them into your directory. However, the file names for the installation assets might change between releases. Use caution when copying installation files from an earlier OpenShift Container Platform version.

   b. At the prompts, provide the configuration details for your cloud:

      i. Optional: Select an SSH key to use to access your cluster machines.
For production OpenShift Container Platform clusters on which you want to perform installation debugging or disaster recovery, specify an SSH key that your ssh-agent process uses.

ii. Select AWS as the platform to target.

iii. If you do not have an Amazon Web Services (AWS) profile stored on your computer, enter the AWS access key ID and secret access key for the user that you configured to run the installation program.

iv. Select the AWS region to deploy the cluster to.

v. Select the base domain for the Route 53 service that you configured for your cluster.

vi. Enter a descriptive name for your cluster.

vii. Paste the pull secret that you obtained from the Pull Secret page on the Red Hat OpenShift Cluster Manager site.

2. Modify the install-config.yaml file. You can find more information about the available parameters in the Installation configuration parameters section.

3. Back up the install-config.yaml file so that you can use it to install multiple clusters.

IMPORTANT

The install-config.yaml file is consumed during the installation process. If you want to reuse the file, you must back it up now.

1.4.5.1. Installation configuration parameters

Before you deploy an OpenShift Container Platform cluster, you provide parameter values to describe your account on the cloud platform that hosts your cluster and optionally customize your cluster’s platform. When you create the install-config.yaml installation configuration file, you provide values for the required parameters through the command line. If you customize your cluster, you can modify the install-config.yaml file to provide more details about the platform.

NOTE

After installation, you cannot modify these parameters in the install-config.yaml file.

IMPORTANT

The openshift-install command does not validate field names for parameters. If an incorrect name is specified, the related file or object is not created, and no error is reported. Ensure that the field names for any parameters that are specified are correct.

1.4.5.1.1. Required configuration parameters

Required installation configuration parameters are described in the following table:

Table 1.1. Required parameters
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>apiVersion</td>
<td>The API version for the install-config.yaml content. The current version is v1. The installer may also support older API versions.</td>
<td>String</td>
</tr>
<tr>
<td>baseDomain</td>
<td>The base domain of your cloud provider. The base domain is used to create routes to your OpenShift Container Platform cluster components. The full DNS name for your cluster is a combination of the baseDomain and metadata.name parameter values that uses the &lt;metadata.name&gt;.&lt;baseDomain&gt; format.</td>
<td>A fully-qualified domain or subdomain name, such as example.com.</td>
</tr>
<tr>
<td>metadata</td>
<td>Kubernetes resource ObjectMeta, from which only the name parameter is consumed.</td>
<td>Object</td>
</tr>
<tr>
<td>metadata.name</td>
<td>The name of the cluster. DNS records for the cluster are all subdomains of {{.metadata.name}}{{.baseDomain}}.</td>
<td>String of lowercase letters, hyphens (-), and periods (.), such as dev.</td>
</tr>
<tr>
<td>platform</td>
<td>The configuration for the specific platform upon which to perform the installation: aws, baremetal, azure, openstack, ovirt, vsphere. For additional information about platform.&lt;platform&gt; parameters, consult the following table for your specific platform.</td>
<td>Object</td>
</tr>
</tbody>
</table>
### Pull Secret

Get a pull secret from https://console.redhat.com/openshift/install/pull-secret to authenticate downloading container images for OpenShift Container Platform components from services such as Quay.io.

```json
{
  "auths":{
    "cloud.openshift.com":{
      "auth": "b3Blb=",
      "email": "you@example.com"
    },
    "quay.io":{
      "auth": "b3Blb=",
      "email": "you@example.com"
    }
  }
}
```

### 1.4.5.1.2. Network configuration parameters

You can customize your installation configuration based on the requirements of your existing network infrastructure. For example, you can expand the IP address block for the cluster network or provide different IP address blocks than the defaults.

Only IPv4 addresses are supported.

### Table 1.2. Network parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>networking</strong></td>
<td>The configuration for the cluster network.</td>
<td>Object</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NOTE</td>
</tr>
<tr>
<td></td>
<td></td>
<td>You cannot modify parameters specified by the <code>networking</code> object after installation.</td>
</tr>
<tr>
<td>networking.network</td>
<td>The cluster network provider Container Network Interface (CNI) plug-in to install.</td>
<td>Either OpenShiftSDN or OVNKubernetes. The default value is OpenShiftSDN.</td>
</tr>
<tr>
<td>networking.clusterNetwork</td>
<td>The IP address blocks for pods.</td>
<td>An array of objects. For example:</td>
</tr>
<tr>
<td></td>
<td>The default value is 10.128.0.0/14 with a host prefix of /23.</td>
<td>networking:</td>
</tr>
<tr>
<td></td>
<td>If you specify multiple IP address blocks, the blocks must not overlap.</td>
<td>clusterNetwork:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- cidr: 10.128.0.0/14 hostPrefix: 23</td>
</tr>
<tr>
<td>Parameter</td>
<td>Description</td>
<td>Values</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>networking.clusterNetwork.cidr</td>
<td>Required if you use <em>networking.clusterNetwork</em>. An IP address block.</td>
<td>An IP address block in Classless Inter-Domain Routing (CIDR) notation. The prefix length for an IPv4 block is between 0 and 32.</td>
</tr>
<tr>
<td>networking.clusterNetwork.hostPrefix</td>
<td>The subnet prefix length to assign to each individual node. For example, if <em>hostPrefix</em> is set to 23 then each node is assigned a /23 subnet out of the given <em>cidr</em>. A <em>hostPrefix</em> value of 23 provides 510 (2^(32 - 23) - 2) pod IP addresses.</td>
<td>A subnet prefix.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The default value is 23.</td>
</tr>
<tr>
<td>networking.serviceNetwork</td>
<td>The IP address block for services. The default value is 172.30.0.0/16.</td>
<td>An array with an IP address block in CIDR format. For example:</td>
</tr>
<tr>
<td></td>
<td>The OpenShift SDN and OVN-Kubernetes network providers support only a single IP address block for the service network.</td>
<td>networking:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>serviceNetwork:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- 172.30.0.0/16</td>
</tr>
<tr>
<td>networking.machineNetwork</td>
<td>The IP address blocks for machines.</td>
<td>An array of objects. For example:</td>
</tr>
<tr>
<td></td>
<td>If you specify multiple IP address blocks, the blocks must not overlap.</td>
<td>networking:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>machineNetwork:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- cidr: 10.0.0.0/16</td>
</tr>
<tr>
<td>networking.machineNetwork.cidr</td>
<td>Required if you use <em>networking.machineNetwork</em>. An IP address block. The default value is 10.0.0.0/16 for all platforms other than libvirt. For libvirt, the default value is 192.168.126.0/24.</td>
<td>An IP network block in CIDR notation.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>For example, 10.0.0.0/16.</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>NOTE</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Set the <em>networking.machineNetwork</em> to match the CIDR that the preferred NIC resides in.</td>
</tr>
</tbody>
</table>

**1.4.5.1.3. Optional configuration parameters**

Optional installation configuration parameters are described in the following table:

**Table 1.3. Optional parameters**
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>additionalTrustBundle</td>
<td>A PEM-encoded X.509 certificate bundle that is added to the nodes’ trusted certificate store. This trust bundle may also be used when a proxy has been configured.</td>
<td>String</td>
</tr>
<tr>
<td>compute</td>
<td>The configuration for the machines that comprise the compute nodes.</td>
<td>Array of machine-pool objects. For details, see the following &quot;Machine-pool&quot; table.</td>
</tr>
<tr>
<td>compute.architecture</td>
<td>Determines the instruction set architecture of the machines in the pool. Currently, heterogeneous clusters are not supported, so all pools must specify the same architecture. Valid values are <code>amd64</code> (the default).</td>
<td>String</td>
</tr>
<tr>
<td>compute.hyperthreading</td>
<td>Whether to enable or disable simultaneous multithreading, or <strong>hyperthreading</strong>, on compute machines. By default, simultaneous multithreading is enabled to increase the performance of your machines’ cores.</td>
<td><strong>Enabled</strong> or <strong>Disabled</strong></td>
</tr>
<tr>
<td></td>
<td><strong>IMPORTANT</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>If you disable simultaneous multithreading, ensure that your capacity planning accounts for the dramatically decreased machine performance.</td>
<td></td>
</tr>
<tr>
<td>compute.name</td>
<td>Required if you use compute. The name of the machine pool.</td>
<td>worker</td>
</tr>
<tr>
<td>compute.platform</td>
<td>Required if you use compute. Use this parameter to specify the cloud provider to host the worker machines. This parameter value must match the <code>controlPlane.platform</code> parameter value.</td>
<td><code>aws</code>, <code>azure</code>, <code>gcp</code>, <code>openstack</code>, <code>ovirt</code>, <code>vsphere</code>, or <code>{}</code></td>
</tr>
<tr>
<td>compute.replicas</td>
<td>The number of compute machines, which are also known as worker machines, to provision.</td>
<td>A positive integer greater than or equal to 2. The default value is 3.</td>
</tr>
<tr>
<td>Parameter</td>
<td>Description</td>
<td>Values</td>
</tr>
<tr>
<td>--------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>controlPlane</td>
<td>The configuration for the machines that comprise the control plane.</td>
<td>Array of MachinePool objects. For details, see the following &quot;Machine-pool&quot; table.</td>
</tr>
<tr>
<td>controlPlane.architecture</td>
<td>Determines the instruction set architecture of the machines in the pool.</td>
<td>String</td>
</tr>
<tr>
<td></td>
<td>Currently, heterogeneous clusters are not supported, so all pools must</td>
<td></td>
</tr>
<tr>
<td></td>
<td>specify the same architecture. Valid values are <strong>amd64</strong> (the default).</td>
<td></td>
</tr>
<tr>
<td>controlPlane.hypertreading</td>
<td>Whether to enable or disable simultaneous multithreading, or hyperthreading,</td>
<td><strong>Enabled</strong> or <strong>Disabled</strong></td>
</tr>
<tr>
<td></td>
<td>on control plane machines. By default, simultaneous multithreading is</td>
<td></td>
</tr>
<tr>
<td></td>
<td>enabled to increase the performance of your machines' cores.</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>IMPORTANT</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>If you disable simultaneous multithreading, ensure that your capacity</td>
<td></td>
</tr>
<tr>
<td></td>
<td>planning accounts for the dramatically decreased machine performance.</td>
<td></td>
</tr>
<tr>
<td>controlPlane.name</td>
<td>Required if you use controlPlane.</td>
<td><strong>master</strong></td>
</tr>
<tr>
<td></td>
<td>The name of the machine pool.</td>
<td></td>
</tr>
<tr>
<td>controlPlane.platform</td>
<td>Required if you use controlPlane.</td>
<td>aws, azure, gcp, openstack, ovirt, vsphere, or {}</td>
</tr>
<tr>
<td></td>
<td>Use this parameter to specify the cloud provider that hosts the control</td>
<td></td>
</tr>
<tr>
<td></td>
<td>plane machines. This parameter value must match the compute.platform</td>
<td></td>
</tr>
<tr>
<td></td>
<td>parameter value.</td>
<td></td>
</tr>
<tr>
<td>controlPlane.replicas</td>
<td>The number of control plane machines to provision.</td>
<td>The only supported value is 3, which is the default value.</td>
</tr>
</tbody>
</table>
## credentialsMode

The Cloud Credential Operator (CCO) mode. If no mode is specified, the CCO dynamically tries to determine the capabilities of the provided credentials, with a preference for mint mode on the platforms where multiple modes are supported.

### NOTE

Not all CCO modes are supported for all cloud providers. For more information on CCO modes, see the Cloud Credential Operator entry in the Red Hat Operators reference content.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>credentialsMode</td>
<td>The Cloud Credential Operator (CCO) mode. If no mode is specified, the CCO dynamically tries to determine the capabilities of the provided credentials, with a preference for mint mode on the platforms where multiple modes are supported.</td>
<td>Mint, Passthrough, Manual, or an empty string (&quot;&quot;).</td>
</tr>
<tr>
<td>Parameter</td>
<td>Description</td>
<td>Values</td>
</tr>
<tr>
<td>-------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>-------------------------</td>
</tr>
<tr>
<td><strong>fips</strong></td>
<td>Enable or disable FIPS mode. The default is <strong>false</strong> (disabled). If FIPS mode is enabled, the Red Hat Enterprise Linux CoreOS (RHCOS) machines that OpenShift Container Platform runs on bypass the default Kubernetes cryptography suite and use the cryptography modules that are provided with RHCOS instead.</td>
<td><strong>false</strong> or <strong>true</strong></td>
</tr>
<tr>
<td><strong>imageContentSources</strong></td>
<td>Sources and repositories for the release-image content.</td>
<td>Array of objects. Includes a <strong>source</strong> and, optionally, <strong>mirrors</strong>, as described in the following rows of this table.</td>
</tr>
<tr>
<td><strong>imageContentSources.source</strong></td>
<td>Required if you use <strong>imageContentSources</strong>. Specify the repository that users refer to, for example, in image pull specifications.</td>
<td>String</td>
</tr>
<tr>
<td><strong>imageContentSources.mirrors</strong></td>
<td>Specify one or more repositories that may also contain the same images.</td>
<td>Array of strings</td>
</tr>
<tr>
<td><strong>publish</strong></td>
<td>How to publish or expose the user-facing endpoints of your cluster, such as the Kubernetes API, OpenShift routes.</td>
<td><strong>Internal</strong> or <strong>External</strong>. To deploy a private cluster, which cannot be accessed from the internet, set <strong>publish</strong> to <strong>Internal</strong>. The default value is <strong>External</strong>.</td>
</tr>
</tbody>
</table>
### One or more keys. For example:

- `sshKey: <key1>`
- `<key2>`
- `<key3>`

#### 1.4.5.1.4. Optional AWS configuration parameters

Optional AWS configuration parameters are described in the following table:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>compute.platform.aws.amiID</code></td>
<td>The AWS AMI used to boot compute machines for the cluster. This is required for regions that require a custom RHCOS AMI.</td>
<td>Any published or custom RHCOS AMI that belongs to the set AWS region.</td>
</tr>
<tr>
<td><code>compute.platform.aws.rootVolume.iops</code></td>
<td>The Input/Output Operations Per Second (IOPS) that is reserved for the root volume.</td>
<td>Integer, for example 4000.</td>
</tr>
<tr>
<td><code>compute.platform.aws.rootVolume.size</code></td>
<td>The size in GiB of the root volume.</td>
<td>Integer, for example 500.</td>
</tr>
<tr>
<td><code>compute.platform.aws.rootVolume.type</code></td>
<td>The type of the root volume.</td>
<td>Valid AWS EBS volume type, such as <code>io1</code>.</td>
</tr>
<tr>
<td><code>compute.platform.aws.type</code></td>
<td>The EC2 instance type for the compute machines.</td>
<td>Valid AWS instance type, such as <code>c5.9xlarge</code>.</td>
</tr>
<tr>
<td>Parameter</td>
<td>Description</td>
<td>Values</td>
</tr>
<tr>
<td>-----------</td>
<td>-------------</td>
<td>--------</td>
</tr>
<tr>
<td><code>compute.platform.aws.zones</code></td>
<td>The availability zones where the installation program creates machines for the compute machine pool. If you provide your own VPC, you must provide a subnet in that availability zone.</td>
<td>A list of valid AWS availability zones, such as <code>us-east-1c</code>, in a YAML sequence.</td>
</tr>
<tr>
<td><code>compute.aws.region</code></td>
<td>The AWS region that the installation program creates compute resources in.</td>
<td>Any valid AWS region, such as <code>us-east-1</code>.</td>
</tr>
<tr>
<td><code>controlPlane.platform.aws.amiID</code></td>
<td>The AWS AMI used to boot control plane machines for the cluster. This is required for regions that require a custom RH COS AMI.</td>
<td>Any published or custom RH COS AMI that belongs to the set AWS region.</td>
</tr>
<tr>
<td><code>controlPlane.platform.aws.type</code></td>
<td>The EC2 instance type for the control plane machines.</td>
<td>Valid AWS instance type, such as <code>c5.9xlarge</code>.</td>
</tr>
<tr>
<td><code>controlPlane.platform.aws.zone</code></td>
<td>The availability zones where the installation program creates machines for the control plane machine pool.</td>
<td>A list of valid AWS availability zones, such as <code>us-east-1c</code>, in a YAML sequence.</td>
</tr>
<tr>
<td><code>controlPlane.aws.region</code></td>
<td>The AWS region that the installation program creates control plane resources in.</td>
<td>Valid AWS region, such as <code>us-east-1</code>.</td>
</tr>
<tr>
<td><code>platform.aws.amiID</code></td>
<td>The AWS AMI used to boot all machines for the cluster. If set, the AMI must belong to the same region as the cluster. This is required for regions that require a custom RH COS AMI.</td>
<td>Any published or custom RH COS AMI that belongs to the set AWS region.</td>
</tr>
<tr>
<td><code>platform.aws.serviceEndpoints.name</code></td>
<td>The AWS service endpoint name. Custom endpoints are only required for cases where alternative AWS endpoints, like FIPS, must be used. Custom API endpoints can be specified for EC2, S3, IAM, Elastic Load Balancing, Tagging, Route 53, and STS AWS services.</td>
<td>Valid AWS service endpoint name.</td>
</tr>
</tbody>
</table>
### Parameter Description Values

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>platform.aws.serviceEndpoints.url</code></td>
<td>The AWS service endpoint URL. The URL must use the <code>https</code> protocol and the host must trust the certificate.</td>
<td>Valid AWS service endpoint URL.</td>
</tr>
<tr>
<td><code>platform.aws.userTags</code></td>
<td>A map of keys and values that the installation program adds as tags to all resources that it creates.</td>
<td>Any valid YAML map, such as key value pairs in the <code>&lt;key&gt;: &lt;value&gt;</code> format. For more information about AWS tags, see <a href="https://docs.aws.amazon.com/efs/latest/efsv1/tagging.html">Tagging Your Amazon EC2 Resources</a> in the AWS documentation.</td>
</tr>
<tr>
<td><code>platform.aws.subnets</code></td>
<td>If you provide the VPC instead of allowing the installation program to create the VPC for you, specify the subnet for the cluster to use. The subnet must be part of the same <code>machineNetwork[].cidr</code> ranges that you specify. For a standard cluster, specify a public and a private subnet for each availability zone. For a private cluster, specify a private subnet for each availability zone.</td>
<td>Valid subnet IDs.</td>
</tr>
</tbody>
</table>

### 1.4.5.2. Sample customized `install-config.yaml` file for AWS

You can customize the `install-config.yaml` file to specify more details about your OpenShift Container Platform cluster’s platform or modify the values of the required parameters.

**IMPORTANT**

This sample YAML file is provided for reference only. You must obtain your `install-config.yaml` file by using the installation program and modify it.

```yaml
apiVersion: v1
baseDomain: example.com
credentialsMode: Mint
controlPlane:
  name: master
  platform:
    aws:
      zones:
        - us-west-2a
        - us-west-2b
      rootVolume:
        iops: 4000
```
size: 500
  type: io1
  type: m5.xlarge
replicas: 3
compute:
  - hyperthreading: Enabled
name: worker
platform:
  aws:
    rootVolume:
      iops: 2000
      size: 500
      type: io1
      type: c5.xlarge
    zones:
      - us-west-2c
replicas: 3
metadata:
  name: test-cluster
networking:
  clusterNetwork:
    - cidr: 10.128.0.0/14
      hostPrefix: 23
  machineNetwork:
    - cidr: 10.0.0.0/16
      networkType: OpenShiftSDN
      serviceNetwork:
        - 172.30.0.0/16
platform:
  aws:
    region: us-west-2
    userTags:
      adminContact: jdoe
      costCenter: 7536
      amiID: ami-96c6f8f7
      serviceEndpoints:
        - name: ec2
          url: https://vpce-id.ec2.us-west-2.vpce.amazonaws.com
      fips: false
      sshKey: ssh-ed25519 AAAA...
pullSecret: '{"auths": ...}'

1 Required. The installation program prompts you for this value.

2 Optional: Add this parameter to force the Cloud Credential Operator (CCO) to use the specified mode, instead of having the CCO dynamically try to determine the capabilities of the credentials. For details about CCO modes, see the Cloud Credential Operator entry in the Red Hat Operators reference content.

3 If you do not provide these parameters and values, the installation program provides the default value.

4 The controlPlane section is a single mapping, but the compute section is a sequence of mappings. To meet the requirements of the different data structures, the first line of the compute section must begin with a hyphen, - , and the first line of the controlPlane section must not. Although both
sections currently define a single machine pool, it is possible that future versions of OpenShift Container Platform will support defining multiple compute pools during installation. Only one control plane pool is used.

Whether to enable or disable simultaneous multithreading, or **hyperthreading**. By default, simultaneous multithreading is enabled to increase the performance of your machines’ cores. You can disable it by setting the parameter value to **Disabled**. If you disable simultaneous multithreading in some cluster machines, you must disable it in all cluster machines.

**IMPORTANT**
If you disable simultaneous multithreading, ensure that your capacity planning accounts for the dramatically decreased machine performance. Use larger instance types, such as **m4.2xlarge** or **m5.2xlarge**, for your machines if you disable simultaneous multithreading.

To configure faster storage for etcd, especially for larger clusters, set the storage type as **io1** and set **iops** to **2000**.

The ID of the AMI used to boot machines for the cluster. If set, the AMI must belong to the same region as the cluster.

The AWS service endpoints. Custom endpoints are required when installing to an unknown AWS region. The endpoint URL must use the **https** protocol and the host must trust the certificate.

Whether to enable or disable FIPS mode. By default, FIPS mode is not enabled. If FIPS mode is enabled, the Red Hat Enterprise Linux CoreOS (RHCOS) machines that OpenShift Container Platform runs on bypass the default Kubernetes cryptography suite and use the cryptography modules that are provided with RHCOS instead.

**IMPORTANT**
The use of FIPS Validated / Modules in Process cryptographic libraries is only supported on OpenShift Container Platform deployments on the **x86_64** architecture.

You can optionally provide the **sshKey** value that you use to access the machines in your cluster.

**NOTE**
For production OpenShift Container Platform clusters on which you want to perform installation debugging or disaster recovery, specify an SSH key that your **ssh-agent** process uses.

### 1.4.5.3. Configuring the cluster-wide proxy during installation

Production environments can deny direct access to the Internet and instead have an HTTP or HTTPS proxy available. You can configure a new OpenShift Container Platform cluster to use a proxy by configuring the proxy settings in the **install-config.yaml** file.

**Prerequisites**
- You have an existing **install-config.yaml** file.
You reviewed the sites that your cluster requires access to and determined whether any of them need to bypass the proxy. By default, all cluster egress traffic is proxied, including calls to hosting cloud provider APIs. You added sites to the *Proxy* object’s `spec.noProxy` field to bypass the proxy if necessary.

**NOTE**

The *Proxy* object `status.noProxy` field is populated with the values of the `networking.machineNetwork[].cidr`, `networking.clusterNetwork[].cidr`, and `networking.serviceNetwork[]` fields from your installation configuration.

For installations on Amazon Web Services (AWS), Google Cloud Platform (GCP), Microsoft Azure, and Red Hat OpenStack Platform (RHOSP), the *Proxy* object `status.noProxy` field is also populated with the instance metadata endpoint (169.254.169.254).

If your cluster is on AWS, you added the `ec2.<region>.amazonaws.com`, `elasticloadbalancing.<region>.amazonaws.com`, and `s3.<region>.amazonaws.com` endpoints to your VPC endpoint. These endpoints are required to complete requests from the nodes to the AWS EC2 API. Because the proxy works on the container level, not the node level, you must route these requests to the AWS EC2 API through the AWS private network. Adding the public IP address of the EC2 API to your allowlist in your proxy server is not sufficient.

Procedure

1. Edit your *install-config.yaml* file and add the proxy settings. For example:

```yaml
apiVersion: v1
baseDomain: my.domain.com
proxy:
  httpProxy: http://<username>:<pswd>@<ip>:<port>
  httpsProxy: https://<username>:<pswd>@<ip>:<port>
  noProxy: example.com
additionalTrustBundle: |
  -----BEGIN CERTIFICATE-----
  <MY_TRUSTED_CA_CERT>
  -----END CERTIFICATE-----
...
```

1. A proxy URL to use for creating HTTP connections outside the cluster. The URL scheme must be `http`.
2. A proxy URL to use for creating HTTPS connections outside the cluster.
3. A comma-separated list of destination domain names, IP addresses, or other network CIDRs to exclude from proxying. Preface a domain with . to match subdomains only. For example, `.y.com` matches `x.y.com`, but not `y.com`. Use `*` to bypass the proxy for all destinations.
4. If provided, the installation program generates a config map that is named `user-ca-bundle` in the *openshift-config* namespace to hold the additional CA certificates. If you provide `additionalTrustBundle` and at least one proxy setting, the *Proxy* object is configured to reference the `user-ca-bundle` config map in the `trustedCA` field. The Cluster Network Operator then creates a `trusted-ca-bundle` config map that merges the contents specified for the `trustedCA` parameter with the RHCOS trust bundle. The
**NOTE**

The installation program does not support the proxy `readinessEndpoints` field.

2. Save the file and reference it when installing OpenShift Container Platform.

The installation program creates a cluster-wide proxy that is named `cluster` that uses the proxy settings in the provided `install-config.yaml` file. If no proxy settings are provided, a `cluster Proxy` object is still created, but it will have a nil `spec`.

**NOTE**

Only the `Proxy` object named `cluster` is supported, and no additional proxies can be created.

### 1.4.6. Deploying the cluster

You can install OpenShift Container Platform on a compatible cloud platform.

**IMPORTANT**

You can run the `create cluster` command of the installation program only once, during initial installation.

**Prerequisites**

- Configure an account with the cloud platform that hosts your cluster.
- Obtain the OpenShift Container Platform installation program and the pull secret for your cluster.

**Procedure**

1. Change to the directory that contains the installation program and initialize the cluster deployment:

   ```bash
   $ ./openshift-install create cluster --dir <installation_directory> \[
   --log-level=info
   ```

   1. For `<installation_directory>`, specify the location of your customized `./install-config.yaml` file.
   2. To view different installation details, specify `warn`, `debug`, or `error` instead of `info`.  

   1. For `<installation_directory>`, specify the location of your customized `./install-config.yaml` file.
   2. To view different installation details, specify `warn`, `debug`, or `error` instead of `info`.  

   ```bash
   ```
NOTE

If the cloud provider account that you configured on your host does not have sufficient permissions to deploy the cluster, the installation process stops, and the missing permissions are displayed.

When the cluster deployment completes, directions for accessing your cluster, including a link to its web console and credentials for the kubeadmin user, display in your terminal.

Example output

```
INFO Install complete!
INFO To access the cluster as the system:admin user when using 'oc', run 'export KUBECONFIG=/home/myuser/install_dir/auth/kubeconfig'
INFO Access the OpenShift web-console here: https://console-openshift-console.apps.mycluster.example.com
INFO Login to the console with user: "kubeadmin", and password: "4vYBz-Ee6gm-ymBZj-Wl5AL"
INFO Time elapsed: 36m22s
```

NOTE

The cluster access and credential information also outputs to `<installation_directory>/openshift_install.log` when an installation succeeds.

IMPORTANT

The Ignition config files that the installation program generates contain certificates that expire after 24 hours, which are then renewed at that time. If the cluster is shut down before renewing the certificates and the cluster is later restarted after the 24 hours have elapsed, the cluster automatically recovers the expired certificates. The exception is that you must manually approve the pending node-bootstrapper certificate signing requests (CSRs) to recover kubelet certificates. See the documentation for Recovering from expired control plane certificates for more information.

IMPORTANT

You must not delete the installation program or the files that the installation program creates. Both are required to delete the cluster.

2. Optional: Remove or disable the AdministratorAccess policy from the IAM account that you used to install the cluster.

NOTE

The elevated permissions provided by the AdministratorAccess policy are required only during installation.

1.4.7. Installing the OpenShift CLI by downloading the binary
You can install the OpenShift CLI (oc) in order to interact with OpenShift Container Platform from a command-line interface. You can install oc on Linux, Windows, or macOS.

**IMPORTANT**

If you installed an earlier version of oc, you cannot use it to complete all of the commands in OpenShift Container Platform 4.6. Download and install the new version of oc.

### 1.4.7.1. Installing the OpenShift CLI on Linux

You can install the OpenShift CLI (oc) binary on Linux by using the following procedure.

**Procedure**

2. Select the appropriate version in the Version drop-down menu.
3. Click Download Now next to the OpenShift v4.6 Linux Client entry and save the file.
4. Unpack the archive:
   ```
   $ tar xvzf <file>
   ```
5. Place the oc binary in a directory that is on your PATH.
   To check your PATH, execute the following command:
   ```
   $ echo $PATH
   ```

After you install the OpenShift CLI, it is available using the oc command:

```
$ oc <command>
```

### 1.4.7.2. Installing the OpenShift CLI on Windows

You can install the OpenShift CLI (oc) binary on Windows by using the following procedure.

**Procedure**

2. Select the appropriate version in the Version drop-down menu.
3. Click Download Now next to the OpenShift v4.6 Windows Client entry and save the file.
4. Unzip the archive with a ZIP program.
5. Move the oc binary to a directory that is on your PATH.
   To check your PATH, open the command prompt and execute the following command:
   ```
   C:\> path
   ```
After you install the OpenShift CLI, it is available using the `oc` command:

```
C:\> oc <command>
```

### 1.4.7.3. Installing the OpenShift CLI on macOS

You can install the OpenShift CLI (`oc`) binary on macOS by using the following procedure.

**Procedure**

2. Select the appropriate version in the **Version** drop-down menu.
3. Click **Download Now** next to the **OpenShift v4.6 MacOSX Client** entry and save the file.
4. Unpack and unzip the archive.
5. Move the `oc` binary to a directory on your PATH.
   - To check your **PATH**, open a terminal and execute the following command:

```
$ echo $PATH
```

After you install the OpenShift CLI, it is available using the `oc` command:

```
$ oc <command>
```

### 1.4.8. Logging in to the cluster by using the CLI

You can log in to your cluster as a default system user by exporting the cluster `kubeconfig` file. The `kubeconfig` file contains information about the cluster that is used by the CLI to connect a client to the correct cluster and API server. The file is specific to a cluster and is created during OpenShift Container Platform installation.

**Prerequisites**

- You deployed an OpenShift Container Platform cluster.
- You installed the `oc` CLI.

**Procedure**

1. Export the `kubeadmin` credentials:

```
$ export KUBECONFIG=<installation_directory>/auth/kubeconfig
```

   For `<installation_directory>`, specify the path to the directory that you stored the installation files in.

2. Verify you can run `oc` commands successfully using the exported configuration:
1.4.9. Logging in to the cluster by using the web console

The `kubeadmin` user exists by default after an OpenShift Container Platform installation. You can log in to your cluster as the `kubeadmin` user by using the OpenShift Container Platform web console.

Prerequisites

- You have access to the installation host.
- You completed a cluster installation and all cluster Operators are available.

Procedure

1. Obtain the password for the `kubeadmin` user from the `kubeadmin-password` file on the installation host:

   ```
   $ cat <installation_directory>/auth/kubeadmin-password
   ```

   **NOTE**

   Alternatively, you can obtain the `kubeadmin` password from the `<installation_directory>/openshift_install.log` log file on the installation host.

2. List the OpenShift Container Platform web console route:

   ```
   $ oc get routes -n openshift-console | grep 'console-openshift'
   ```

   **NOTE**

   Alternatively, you can obtain the OpenShift Container Platform route from the `<installation_directory>/openshift_install.log` log file on the installation host.

Example output

```
console console-openshift-console.apps.<cluster_name>.<base_domain> console
https  reencrypt/Redirect  None
```

3. Navigate to the route detailed in the output of the preceding command in a web browser and log in as the `kubeadmin` user.

Additional resources

- See [Accessing the web console](#) for more details about accessing and understanding the OpenShift Container Platform web console.
1.4.10. Telemetry access for OpenShift Container Platform

In OpenShift Container Platform 4.6, the Telemetry service, which runs by default to provide metrics about cluster health and the success of updates, requires Internet access. If your cluster is connected to the Internet, Telemetry runs automatically, and your cluster is registered to the Red Hat OpenShift Cluster Manager (OCM).

After you confirm that your Red Hat OpenShift Cluster Manager inventory is correct, either maintained automatically by Telemetry or manually by using OCM, use subscription watch to track your OpenShift Container Platform subscriptions at the account or multi-cluster level.

Additional resources

- See About remote health monitoring for more information about the Telemetry service.

1.4.11. Next steps

- Validating an installation.
- Customize your cluster.
- If necessary, you can opt out of remote health reporting.
- If necessary, you can remove cloud provider credentials.

1.5. INSTALLING A CLUSTER ON AWS WITH NETWORK CUSTOMIZATIONS

In OpenShift Container Platform version 4.6, you can install a cluster on Amazon Web Services (AWS) with customized network configuration options. By customizing your network configuration, your cluster can coexist with existing IP address allocations in your environment and integrate with existing MTU and VXLAN configurations.

You must set most of the network configuration parameters during installation, and you can modify only kubeProxy configuration parameters in a running cluster.

1.5.1. Prerequisites

- Review details about the OpenShift Container Platform installation and update processes.
- Configure an AWS account to host the cluster.

**IMPORTANT**

If you have an AWS profile stored on your computer, it must not use a temporary session token that you generated while using a multi-factor authentication device. The cluster continues to use your current AWS credentials to create AWS resources for the entire life of the cluster, so you must use key-based, long-lived credentials. To generate appropriate keys, see Managing Access Keys for IAM Users in the AWS documentation. You can supply the keys when you run the installation program.

- If you use a firewall, you must configure it to allow the sites that your cluster requires access to.
If you do not allow the system to manage identity and access management (IAM), then a cluster administrator can manually create and maintain IAM credentials. Manual mode can also be used in environments where the cloud IAM APIs are not reachable.

1.5.2. Internet access for OpenShift Container Platform

In OpenShift Container Platform 4.6, you require access to the Internet to install your cluster.

You must have Internet access to:

- Access the Red Hat OpenShift Cluster Manager page to download the installation program and perform subscription management. If the cluster has Internet access and you do not disable Telemetry, that service automatically entitles your cluster.
- Access Quay.io to obtain the packages that are required to install your cluster.
- Obtain the packages that are required to perform cluster updates.

**IMPORTANT**

If your cluster cannot have direct Internet access, you can perform a restricted network installation on some types of infrastructure that you provision. During that process, you download the content that is required and use it to populate a mirror registry with the packages that you need to install a cluster and generate the installation program. With some installation types, the environment that you install your cluster in will not require Internet access. Before you update the cluster, you update the content of the mirror registry.

1.5.3. Generating an SSH private key and adding it to the agent

If you want to perform installation debugging or disaster recovery on your cluster, you must provide an SSH key to both your ssh-agent and the installation program. You can use this key to access the bootstrap machine in a public cluster to troubleshoot installation issues.

**NOTE**

In a production environment, you require disaster recovery and debugging.

You can use this key to SSH into the master nodes as the user core. When you deploy the cluster, the key is added to the core user’s ~/.ssh/authorized_keys list.

**NOTE**

You must use a local key, not one that you configured with platform-specific approaches such as AWS key pairs.

**Procedure**

1. If you do not have an SSH key that is configured for password-less authentication on your computer, create one. For example, on a computer that uses a Linux operating system, run the following command:

   ```
   $ ssh-keygen -t ed25519 -N "" -f <path>/<file_name>
   ```
1. Specify the path and file name, such as `~/.ssh/id_rsa`, of the new SSH key. If you have an existing key pair, ensure your public key is in the your `~/.ssh` directory.

Running this command generates an SSH key that does not require a password in the location that you specified.

**NOTE**

If you plan to install an OpenShift Container Platform cluster that uses FIPS Validated / Modules in Process cryptographic libraries on the x86_64 architecture, do not create a key that uses the ed25519 algorithm. Instead, create a key that uses the rsa or ecdsa algorithm.

2. Start the ssh-agent process as a background task:

   `$ eval "$(ssh-agent -s)"

**Example output**

   Agent pid 31874

**NOTE**

If your cluster is in FIPS mode, only use FIPS-compliant algorithms to generate the SSH key. The key must be either RSA or ECDSA.

3. Add your SSH private key to the ssh-agent:

   `$ ssh-add <path>/<file_name>`

**Example output**

   Identity added: /home/<you>/<path>/<file_name> (<computer_name>)

   Specify the path and file name for your SSH private key, such as `~/.ssh/id_rsa`

**Next steps**

- When you install OpenShift Container Platform, provide the SSH public key to the installation program.

**1.5.4. Obtaining the installation program**

Before you install OpenShift Container Platform, download the installation file on a local computer.

**Prerequisites**

- You have a computer that runs Linux or macOS, with 500 MB of local disk space
Procedure

1. Access the Infrastructure Provider page on the Red Hat OpenShift Cluster Manager site. If you have a Red Hat account, log in with your credentials. If you do not, create an account.

2. Select your infrastructure provider.

3. Navigate to the page for your installation type, download the installation program for your operating system, and place the file in the directory where you will store the installation configuration files.

   **IMPORTANT**

   The installation program creates several files on the computer that you use to install your cluster. You must keep the installation program and the files that the installation program creates after you finish installing the cluster. Both files are required to delete the cluster.

   **IMPORTANT**

   Deleting the files created by the installation program does not remove your cluster, even if the cluster failed during installation. To remove your cluster, complete the OpenShift Container Platform uninstallation procedures for your specific cloud provider.

4. Extract the installation program. For example, on a computer that uses a Linux operating system, run the following command:

   ```
   $ tar xvf openshift-install-linux.tar.gz
   ```

5. From the Pull Secret page on the Red Hat OpenShift Cluster Manager site, download your installation pull secret as a .txt file. This pull secret allows you to authenticate with the services that are provided by the included authorities, including Quay.io, which serves the container images for OpenShift Container Platform components.

### 1.5.5. Network configuration phases

When specifying a cluster configuration prior to installation, there are several phases in the installation procedures when you can modify the network configuration:

**Phase 1**

After entering the `openshift-install create install-config` command. In the `install-config.yaml` file, you can customize the following network-related fields:

- `networking.networkType`
- `networking.clusterNetwork`
- `networking.serviceNetwork`
- `networking.machineNetwork`

For more information on these fields, refer to "Installation configuration parameters".
NOTE

Set the `networking.machineNetwork` to match the CIDR that the preferred NIC resides in.

Phase 2

After entering the `openshift-install create manifests` command. If you must specify advanced network configuration, during this phase you can define a customized Cluster Network Operator manifest with only the fields you want to modify.

You cannot override the values specified in phase 1 in the `install-config.yaml` file during phase 2. However, you can further customize the cluster network provider during phase 2.

1.5.6. Creating the installation configuration file

You can customize the OpenShift Container Platform cluster you install on Amazon Web Services (AWS).

Prerequisites

- Obtain the OpenShift Container Platform installation program and the pull secret for your cluster.

Procedure

1. Create the `install-config.yaml` file.
   a. Change to the directory that contains the installation program and run the following command:

   ```
   $ ./openshift-install create install-config --dir <installation_directory> ①
   
   ① For `<installation_directory>`, specify the directory name to store the files that the installation program creates.
   ```

   IMPORTANT

   Specify an empty directory. Some installation assets, like bootstrap X.509 certificates have short expiration intervals, so you must not reuse an installation directory. If you want to reuse individual files from another cluster installation, you can copy them into your directory. However, the file names for the installation assets might change between releases. Use caution when copying installation files from an earlier OpenShift Container Platform version.

   b. At the prompts, provide the configuration details for your cloud:

   i. Optional: Select an SSH key to use to access your cluster machines.
NOTE

For production OpenShift Container Platform clusters on which you want to perform installation debugging or disaster recovery, specify an SSH key that your ssh-agent process uses.

ii. Select **AWS** as the platform to target.

iii. If you do not have an Amazon Web Services (AWS) profile stored on your computer, enter the AWS access key ID and secret access key for the user that you configured to run the installation program.

iv. Select the AWS region to deploy the cluster to.

v. Select the base domain for the Route 53 service that you configured for your cluster.

vi. Enter a descriptive name for your cluster.

vii. Paste the pull secret that you obtained from the Pull Secret page on the Red Hat OpenShift Cluster Manager site.

2. Modify the **install-config.yaml** file. You can find more information about the available parameters in the **Installation configuration parameters** section.

3. Back up the **install-config.yaml** file so that you can use it to install multiple clusters.

**IMPORTANT**

The **install-config.yaml** file is consumed during the installation process. If you want to reuse the file, you must back it up now.

### 1.5.6.1. Installation configuration parameters

Before you deploy an OpenShift Container Platform cluster, you provide parameter values to describe your account on the cloud platform that hosts your cluster and optionally customize your cluster’s platform. When you create the **install-config.yaml** installation configuration file, you provide values for the required parameters through the command line. If you customize your cluster, you can modify the **install-config.yaml** file to provide more details about the platform.

**NOTE**

After installation, you cannot modify these parameters in the **install-config.yaml** file.

**IMPORTANT**

The **openshift-install** command does not validate field names for parameters. If an incorrect name is specified, the related file or object is not created, and no error is reported. Ensure that the field names for any parameters that are specified are correct.

### 1.5.6.1.1. Required configuration parameters

Required installation configuration parameters are described in the following table:

**Table 1.5. Required parameters**
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>apiVersion</td>
<td>The API version for the <code>install-config.yaml</code> content. The current version is v1. The installer may also support older API versions.</td>
<td>String</td>
</tr>
<tr>
<td>baseDomain</td>
<td>The base domain of your cloud provider. The base domain is used to create routes to your OpenShift Container Platform cluster components. The full DNS name for your cluster is a combination of the <code>baseDomain</code> and <code>metadata.name</code> parameter values that uses the <code>&lt;metadata.name&gt;</code>. <code>&lt;baseDomain&gt;</code> format.</td>
<td>A fully-qualified domain or subdomain name, such as example.com.</td>
</tr>
<tr>
<td>metadata</td>
<td>Kubernetes resource <code>ObjectMeta</code>, from which only the <code>name</code> parameter is consumed.</td>
<td>Object</td>
</tr>
<tr>
<td>metadata.name</td>
<td>The name of the cluster. DNS records for the cluster are all subdomains of <code>{{.metadata.name}}</code>. <code>{{.baseDomain}}</code>.</td>
<td>String of lowercase letters, hyphens (-), and periods (.), such as dev.</td>
</tr>
<tr>
<td>platform</td>
<td>The configuration for the specific platform upon which to perform the installation: <code>aws</code>, <code>baremetal</code>, <code>azure</code>, <code>openstack</code>, <code>ovirt</code>, <code>vSphere</code>. For additional information about <code>platform.&lt;platform&gt;</code> parameters, consult the following table for your specific platform.</td>
<td>Object</td>
</tr>
</tbody>
</table>
Get a pull secret from
https://console.redhat.com/openshift/install/pull-secret
to authenticate downloading container images for
OpenShift Container Platform components from services such as Quay.io.

```json
{
  
  "auths":{
    "cloud.openshift.com":{
      "auth":"b3Blb=",
      "email":"you@example.com"
    },
    "quay.io":{
      "auth":"b3Blb=",
      "email":"you@example.com"
    }
  }
}
```

1.5.6.1.2. Network configuration parameters

You can customize your installation configuration based on the requirements of your existing network infrastructure. For example, you can expand the IP address block for the cluster network or provide different IP address blocks than the defaults.

Only IPv4 addresses are supported.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>pullSecret</td>
<td>Get a pull secret from</td>
<td></td>
</tr>
<tr>
<td></td>
<td><a href="https://console.redhat.com/">https://console.redhat.com/</a></td>
<td></td>
</tr>
<tr>
<td></td>
<td>openshift/install/pull-secret</td>
<td></td>
</tr>
<tr>
<td></td>
<td>to authenticate downloading</td>
<td></td>
</tr>
<tr>
<td></td>
<td>container images for</td>
<td></td>
</tr>
<tr>
<td></td>
<td>OpenShift Container Platform</td>
<td></td>
</tr>
<tr>
<td></td>
<td>components from services</td>
<td></td>
</tr>
<tr>
<td></td>
<td>such as Quay.io.</td>
<td></td>
</tr>
</tbody>
</table>

### Table 1.6. Network parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>networking</td>
<td>The configuration for the</td>
<td></td>
</tr>
<tr>
<td></td>
<td>cluster network.</td>
<td>Object</td>
</tr>
<tr>
<td>networking.network</td>
<td>The cluster network provider</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Container Network Interface</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(CNI) plug-in to install.</td>
<td></td>
</tr>
<tr>
<td>networking.type</td>
<td>Either OpenShiftSDN or</td>
<td></td>
</tr>
<tr>
<td></td>
<td>OVNKubernetes.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>The default value is</td>
<td></td>
</tr>
<tr>
<td></td>
<td>OpenShiftSDN.</td>
<td></td>
</tr>
<tr>
<td>networking.clusterNetwork</td>
<td>The IP address blocks for</td>
<td></td>
</tr>
<tr>
<td></td>
<td>pods. The default value is</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10.128.0.0/14 with a host</td>
<td></td>
</tr>
<tr>
<td></td>
<td>prefix of /23.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>If you specify multiple IP</td>
<td></td>
</tr>
<tr>
<td></td>
<td>address blocks, the blocks</td>
<td></td>
</tr>
<tr>
<td></td>
<td>must not overlap.</td>
<td></td>
</tr>
<tr>
<td>Parameter</td>
<td>Description</td>
<td>Values</td>
</tr>
<tr>
<td>-----------</td>
<td>-------------</td>
<td>--------</td>
</tr>
<tr>
<td>networking.clusterNetwork.cidr</td>
<td>Required if you use networking.clusterNetwork. An IP address block. An IPv4 network.</td>
<td>An IP address block in Classless Inter-Domain Routing (CIDR) notation. The prefix length for an IPv4 block is between 0 and 32.</td>
</tr>
<tr>
<td>networking.clusterNetwork.hostPrefix</td>
<td>The subnet prefix length to assign to each individual node. For example, if hostPrefix is set to 23 then each node is assigned a /23 subnet out of the given cidr. A hostPrefix value of 23 provides 510 (2^{(32 - 23)} - 2) pod IP addresses.</td>
<td>A subnet prefix. The default value is 23.</td>
</tr>
<tr>
<td>networking.serviceNetwork</td>
<td>The IP address block for services. The default value is 172.30.0.0/16. The OpenShift SDN and OVN-Kubernetes network providers support only a single IP address block for the service network.</td>
<td>An array with an IP address block in CIDR format. For example: networking: serviceNetwork: - 172.30.0.0/16</td>
</tr>
<tr>
<td>networking.machineNetwork</td>
<td>The IP address blocks for machines. If you specify multiple IP address blocks, the blocks must not overlap.</td>
<td>An array of objects. For example: networking: machineNetwork: - cidr: 10.0.0.0/16</td>
</tr>
<tr>
<td>networking.machineNetwork.cidr</td>
<td>Required if you use networking.machineNetwork. An IP address block. The default value is 10.0.0.0/16 for all platforms other than libvirt. For libvirt, the default value is 192.168.126.0/24.</td>
<td>An IP network block in CIDR notation. For example, 10.0.0.0/16.</td>
</tr>
</tbody>
</table>

**NOTE**

Set the networking.machineNetwork to match the CIDR that the preferred NIC resides in.

### 1.5.6.1.3. Optional configuration parameters

Optional installation configuration parameters are described in the following table:

**Table 1.7. Optional parameters**
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>additionalTrustBundle</strong></td>
<td>A PEM-encoded X.509 certificate bundle that is added to the nodes' trusted certificate store. This trust bundle may also be used when a proxy has been configured.</td>
<td>String</td>
</tr>
<tr>
<td><strong>compute</strong></td>
<td>The configuration for the machines that comprise the compute nodes.</td>
<td>Array of machine-pool objects. For details, see the following &quot;Machine-pool&quot; table.</td>
</tr>
<tr>
<td><strong>compute.architecture</strong></td>
<td>Determines the instruction set architecture of the machines in the pool. Currently, heterogeneous clusters are not supported, so all pools must specify the same architecture. Valid values are <strong>amd64</strong> (the default).</td>
<td>String</td>
</tr>
<tr>
<td><strong>compute.hyperthreading</strong></td>
<td>Whether to enable or disable simultaneous multithreading, or <strong>hyperthreading</strong>, on compute machines. By default, simultaneous multithreading is enabled to increase the performance of your machines' cores.</td>
<td><strong>Enabled</strong> or <strong>Disabled</strong></td>
</tr>
<tr>
<td><strong>compute.name</strong></td>
<td>Required if you use <strong>compute</strong>. The name of the machine pool.</td>
<td><strong>worker</strong></td>
</tr>
<tr>
<td><strong>compute.platform</strong></td>
<td>Required if you use <strong>compute</strong>. Use this parameter to specify the cloud provider to host the worker machines. This parameter value must match the <strong>controlPlane.platform</strong> parameter value.</td>
<td><strong>aws</strong>, <strong>azure</strong>, <strong>gcp</strong>, <strong>openstack</strong>, <strong>ovirt</strong>, <strong>vsphere</strong>, or {}</td>
</tr>
<tr>
<td><strong>compute.replicas</strong></td>
<td>The number of compute machines, which are also known as worker machines, to provision.</td>
<td>A positive integer greater than or equal to 2. The default value is 3.</td>
</tr>
<tr>
<td>Parameter</td>
<td>Description</td>
<td>Values</td>
</tr>
<tr>
<td>------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>---------------------------------------------</td>
</tr>
<tr>
<td>controlPlane</td>
<td>The configuration for the machines that comprise the control plane.</td>
<td>Array of <strong>MachinePool</strong> objects. For details, see the following &quot;Machine-pool&quot; table.</td>
</tr>
<tr>
<td>controlPlane.architecture</td>
<td>Determines the instruction set architecture of the machines in the pool. Currently, heterogeneous clusters are not supported, so all pools must specify the same architecture. Valid values are <strong>amd64</strong> (the default).</td>
<td>String</td>
</tr>
<tr>
<td>controlPlane.hypertreading</td>
<td>Whether to enable or disable simultaneous multithreading, or <strong>hyperthreading</strong>, on control plane machines. By default, simultaneous multithreading is enabled to increase the performance of your machines' cores.</td>
<td><strong>Enabled</strong> or <strong>Disabled</strong></td>
</tr>
<tr>
<td>controlPlane.name</td>
<td>Required if you use controlPlane. The name of the machine pool.</td>
<td>master</td>
</tr>
<tr>
<td>controlPlane.platform</td>
<td>Required if you use controlPlane. Use this parameter to specify the cloud provider that hosts the control plane machines. This parameter value must match the <strong>compute.platform</strong> parameter value.</td>
<td><strong>aws</strong>, <strong>azure</strong>, <strong>gcp</strong>, <strong>openstack</strong>, <strong>ovirt</strong>, <strong>vsphere</strong>, or {}</td>
</tr>
<tr>
<td>controlPlane.replicas</td>
<td>The number of control plane machines to provision.</td>
<td>The only supported value is 3, which is the default value.</td>
</tr>
</tbody>
</table>
The Cloud Credential Operator (CCO) mode. If no mode is specified, the CCO dynamically tries to determine the capabilities of the provided credentials, with a preference for mint mode on the platforms where multiple modes are supported.

**NOTE**
Not all CCO modes are supported for all cloud providers. For more information on CCO modes, see the Cloud Credential Operator entry in the Red Hat Operators reference content.
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>fips</td>
<td>Enable or disable FIPS mode. The default is <strong>false</strong> (disabled). If FIPS mode is enabled, the Red Hat Enterprise Linux CoreOS (RHCOS) machines that OpenShift Container Platform runs on bypass the default Kubernetes cryptography suite and use the cryptography modules that are provided with RHCOS instead.</td>
<td><strong>false</strong> or <strong>true</strong></td>
</tr>
<tr>
<td><strong>IMPORTANT</strong></td>
<td>The use of FIPS Validated / Modules in Process cryptographic libraries is only supported on OpenShift Container Platform deployments on the <strong>x86_64</strong> architecture.</td>
<td></td>
</tr>
<tr>
<td><strong>NOTE</strong></td>
<td>If you are using Azure File storage, you cannot enable FIPS mode.</td>
<td></td>
</tr>
<tr>
<td>imageContentSources</td>
<td>Sources and repositories for the release-image content.</td>
<td>Array of objects. Includes a <strong>source</strong> and, optionally, <strong>mirrors</strong>, as described in the following rows of this table.</td>
</tr>
<tr>
<td>imageContentSources.source</td>
<td>Required if you use <strong>imageContentSources</strong>. Specify the repository that users refer to, for example, in image pull specifications.</td>
<td><strong>String</strong></td>
</tr>
<tr>
<td>imageContentSources.mirrors</td>
<td>Specify one or more repositories that may also contain the same images.</td>
<td>Array of strings</td>
</tr>
<tr>
<td>publish</td>
<td>How to publish or expose the user-facing endpoints of your cluster, such as the Kubernetes API, OpenShift routes.</td>
<td><strong>Internal</strong> or <strong>External</strong>. To deploy a private cluster, which cannot be accessed from the internet, set <strong>publish</strong> to <strong>Internal</strong>. The default value is <strong>External</strong>.</td>
</tr>
</tbody>
</table>
The SSH key or keys to authenticate access your cluster machines.

**NOTE**

For production OpenShift Container Platform clusters on which you want to perform installation debugging or disaster recovery, specify an SSH key that your `ssh-agent` process uses.

One or more keys. For example:

```
sshKey:
  <key1>
  <key2>
  <key3>
```

### 1.5.6.14. Optional AWS configuration parameters

Optional AWS configuration parameters are described in the following table:

**Table 1.8. Optional AWS parameters**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>compute.platform.aws.amiID</code></td>
<td>The AWS AMI used to boot compute machines for the cluster. This is required for regions that require a custom RHCOS AMI.</td>
<td>Any published or custom RHCOS AMI that belongs to the set AWS region.</td>
</tr>
<tr>
<td><code>compute.platform.aws.rootVolume.iops</code></td>
<td>The Input/Output Operations Per Second (IOPS) that is reserved for the root volume.</td>
<td>Integer, for example <strong>4000</strong>.</td>
</tr>
<tr>
<td><code>compute.platform.aws.rootVolume.size</code></td>
<td>The size in GiB of the root volume.</td>
<td>Integer, for example <strong>500</strong>.</td>
</tr>
<tr>
<td><code>compute.platform.aws.rootVolume.type</code></td>
<td>The type of the root volume.</td>
<td>Valid AWS EBS volume type, such as <strong>io1</strong>.</td>
</tr>
<tr>
<td><code>compute.platform.aws.type</code></td>
<td>The EC2 instance type for the compute machines.</td>
<td>Valid AWS instance type, such as <strong>c5.9xlarge</strong>.</td>
</tr>
<tr>
<td>Parameter</td>
<td>Description</td>
<td>Values</td>
</tr>
<tr>
<td>----------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------</td>
</tr>
<tr>
<td><code>compute.platform.aws.zones</code></td>
<td>The availability zones where the installation program creates machines for the compute machine pool. If you provide your own VPC, you must provide a subnet in that availability zone.</td>
<td>A list of valid AWS availability zones, such as <code>us-east-1c</code>, in a YAML sequence.</td>
</tr>
<tr>
<td><code>compute.aws.region</code></td>
<td>The AWS region that the installation program creates compute resources in.</td>
<td>Any valid AWS region, such as <code>us-east-1</code>.</td>
</tr>
<tr>
<td><code>controlPlane.platform.aws.amiID</code></td>
<td>The AWS AMI used to boot control plane machines for the cluster. This is required for regions that require a custom RHCOS AMI.</td>
<td>Any published or custom RHCOS AMI that belongs to the set AWS region.</td>
</tr>
<tr>
<td><code>controlPlane.platform.aws.type</code></td>
<td>The EC2 instance type for the control plane machines.</td>
<td>Valid AWS instance type, such as <code>c5.9xlarge</code>.</td>
</tr>
<tr>
<td><code>controlPlane.platform.aws.zones</code></td>
<td>The availability zones where the installation program creates machines for the control plane machine pool.</td>
<td>A list of valid AWS availability zones, such as <code>us-east-1c</code>, in a YAML sequence.</td>
</tr>
<tr>
<td><code>controlPlane.aws.region</code></td>
<td>The AWS region that the installation program creates control plane resources in.</td>
<td>Valid AWS region, such as <code>us-east-1</code>.</td>
</tr>
<tr>
<td><code>platform.aws.amiID</code></td>
<td>The AWS AMI used to boot all machines for the cluster. If set, the AMI must belong to the same region as the cluster. This is required for regions that require a custom RHCOS AMI.</td>
<td>Any published or custom RHCOS AMI that belongs to the set AWS region.</td>
</tr>
<tr>
<td><code>platform.aws.serviceEndpoints.name</code></td>
<td>The AWS service endpoint name. Custom endpoints are only required for cases where alternative AWS endpoints, like FIPS, must be used. Custom API endpoints can be specified for EC2, S3, IAM, Elastic Load Balancing, Tagging, Route 53, and STS AWS services.</td>
<td>Valid AWS service endpoint name.</td>
</tr>
</tbody>
</table>
### Parameter Description Values

**platform.aws.serviceEndpoints.url**  
The AWS service endpoint URL. The URL must use the https protocol and the host must trust the certificate.  
Valid AWS service endpoint URL.

**platform.aws.userTags**  
A map of keys and values that the installation program adds as tags to all resources that it creates.  
Any valid YAML map, such as key value pairs in the `<key>: <value>` format. For more information about AWS tags, see Tagging Your Amazon EC2 Resources in the AWS documentation.

**platform.aws.subnets**  
If you provide the VPC instead of allowing the installation program to create the VPC for you, specify the subnet for the cluster to use. The subnet must be part of the same `machineNetwork[].cidr` ranges that you specify. For a standard cluster, specify a public and a private subnet for each availability zone. For a private cluster, specify a private subnet for each availability zone.  
Valid subnet IDs.

### 1.5.6.2. Sample customized `install-config.yaml` file for AWS

You can customize the `install-config.yaml` file to specify more details about your OpenShift Container Platform cluster’s platform or modify the values of the required parameters.

**IMPORTANT**

This sample YAML file is provided for reference only. You must obtain your `install-config.yaml` file by using the installation program and modify it.

```yaml
apiVersion: v1
baseDomain: example.com  
credentialsMode: Mint  
controlPlane:  
  hyperthreading: Enabled  
  name: master  
  platform:  
    aws:  
      zones:  
        - us-west-2a  
        - us-west-2b  
      rootVolume:  
        iops: 4000
```
size: 500
  type: io1
  type: m5.xlarge
replicas: 3
compute: - hyperthreading: Enabled
  name: worker
platform:
  aws:
    rootVolume:
    iops: 2000
    size: 500
    type: io1
    type: c5.4xlarge
    zones:
      - us-west-2c
    replicas: 3
metadata:
  name: test-cluster
networking:
  clusterNetwork:
    - cidr: 10.128.0.0/14
    hostPrefix: 23
  machineNetwork:
    - cidr: 10.0.0.0/16
  networkType: OpenShiftSDN
  serviceNetwork:
    - 172.30.0.0/16
platform:
  aws:
    region: us-west-2
  userTags:
    adminContact: jdoe
    costCenter: 7536
  amiID: ami-96c6f8f7
  serviceEndpoints:
    - name: ec2
      url: https://vpce-id.ec2.us-west-2.vpce.amazonaws.com
  fips: false
  sshKey: ssh-ed25519 AAAA...
pullSecret: '{"auths": ...}'

1. Required. The installation program prompts you for this value.
2. Optional: Add this parameter to force the Cloud Credential Operator (CCO) to use the specified mode, instead of having the CCO dynamically try to determine the capabilities of the credentials. For details about CCO modes, see the Cloud Credential Operator entry in the Red Hat Operators reference content.
3. If you do not provide these parameters and values, the installation program provides the default value.
4. The controlPlane section is a single mapping, but the compute section is a sequence of mappings. To meet the requirements of the different data structures, the first line of the compute section must begin with a hyphen, -, and the first line of the controlPlane section must not. Although both
sections currently define a single machine pool, it is possible that future versions of OpenShift Container Platform will support defining multiple compute pools during installation. Only one control plane pool is used.

Whether to enable or disable simultaneous multithreading, or **hyperthreading**. By default, simultaneous multithreading is enabled to increase the performance of your machines’ cores. You can disable it by setting the parameter value to **Disabled**. If you disable simultaneous multithreading in some cluster machines, you must disable it in all cluster machines.

**IMPORTANT**

If you disable simultaneous multithreading, ensure that your capacity planning accounts for the dramatically decreased machine performance. Use larger instance types, such as **m4.2xlarge** or **m5.2xlarge**, for your machines if you disable simultaneous multithreading.

To configure faster storage for etcd, especially for larger clusters, set the storage type as **io1** and set **iops** to **2000**.

The ID of the AMI used to boot machines for the cluster. If set, the AMI must belong to the same region as the cluster.

The AWS service endpoints. Custom endpoints are required when installing to an unknown AWS region. The endpoint URL must use the **https** protocol and the host must trust the certificate.

Whether to enable or disable FIPS mode. By default, FIPS mode is not enabled. If FIPS mode is enabled, the Red Hat Enterprise Linux CoreOS (RHCOS) machines that OpenShift Container Platform runs on bypass the default Kubernetes cryptography suite and use the cryptography modules that are provided with RHCOS instead.

**IMPORTANT**

The use of FIPS Validated / Modules in Process cryptographic libraries is only supported on OpenShift Container Platform deployments on the **x86_64** architecture.

You can optionally provide the **sshKey** value that you use to access the machines in your cluster.

**NOTE**

For production OpenShift Container Platform clusters on which you want to perform installation debugging or disaster recovery, specify an SSH key that your **ssh-agent** process uses.

### 1.5.6.3. Configuring the cluster-wide proxy during installation

Production environments can deny direct access to the Internet and instead have an HTTP or HTTPS proxy available. You can configure a new OpenShift Container Platform cluster to use a proxy by configuring the proxy settings in the **install-config.yaml** file.

**Prerequisites**

- You have an existing **install-config.yaml** file.
• You reviewed the sites that your cluster requires access to and determined whether any of them need to bypass the proxy. By default, all cluster egress traffic is proxied, including calls to hosting cloud provider APIs. You added sites to the Proxy object’s spec.noProxy field to bypass the proxy if necessary.

NOTE

The Proxy object status.noProxy field is populated with the values of the networking.machineNetwork[].cidr, networking.clusterNetwork[].cidr, and networking.serviceNetwork[] fields from your installation configuration.

For installations on Amazon Web Services (AWS), Google Cloud Platform (GCP), Microsoft Azure, and Red Hat OpenStack Platform (RHOSP), the Proxy object status.noProxy field is also populated with the instance metadata endpoint (169.254.169.254).

• If your cluster is on AWS, you added the ec2.<region>.amazonaws.com, elasticloadbalancing.<region>.amazonaws.com, and s3.<region>.amazonaws.com endpoints to your VPC endpoint. These endpoints are required to complete requests from the nodes to the AWS EC2 API. Because the proxy works on the container level, not the node level, you must route these requests to the AWS EC2 API through the AWS private network. Adding the public IP address of the EC2 API to your allowlist in your proxy server is not sufficient.

Procedure

1. Edit your install-config.yaml file and add the proxy settings. For example:

```
apiVersion: v1
crds.config.openshift.io/v1

baseDomain: my.domain.com

proxy:
  httpProxy: http://<username>:<pswd>@<ip>:<port> 1
  httpsProxy: https://<username>:<pswd>@<ip>:<port> 2
  noProxy: example.com 3
  additionalTrustBundle: | 4
    -----BEGIN CERTIFICATE-----
    <MY_TRUSTED_CA_CERT>
    -----END CERTIFICATE-----
    ...
```

1 A proxy URL to use for creating HTTP connections outside the cluster. The URL scheme must be http.

2 A proxy URL to use for creating HTTPS connections outside the cluster.

3 A comma-separated list of destination domain names, IP addresses, or other network CIDRs to exclude from proxying. Preface a domain with . to match subdomains only. For example, .y.com matches x.y.com, but not y.com. Use * to bypass the proxy for all destinations.

4 If provided, the installation program generates a config map that is named user-ca-bundle in the openshift-config namespace to hold the additional CA certificates. If you provide additionalTrustBundle and at least one proxy setting, the Proxy object is configured to reference the user-ca-bundle config map in the trustedCA field. The Cluster Network Operator then creates a trusted-ca-bundle config map that merges the contents specified for the trustedCA parameter with the RHCOS trust bundle. The
**additionalTrustBundle** field is required unless the proxy’s identity certificate is signed by an authority from the RHCOS trust bundle.

**NOTE**

The installation program does not support the proxy **readinessEndpoints** field.

2. Save the file and reference it when installing OpenShift Container Platform.

The installation program creates a cluster-wide proxy that is named **cluster** that uses the proxy settings in the provided **install-config.yaml** file. If no proxy settings are provided, a **cluster Proxy** object is still created, but it will have a nil **spec**.

**NOTE**

Only the **Proxy** object named **cluster** is supported, and no additional proxies can be created.

### 1.5.7. Cluster Network Operator configuration

The configuration for the cluster network is specified as part of the Cluster Network Operator (CNO) configuration and stored in a custom resource (CR) object that is named **cluster**. The CR specifies the fields for the **Network** API in the **operator.openshift.io** API group.

The CNO configuration inherits the following fields during cluster installation from the **Network** API in the **Network.config.openshift.io** API group and these fields cannot be changed:

- **clusterNetwork**
  - IP address pools from which pod IP addresses are allocated.
- **serviceNetwork**
  - IP address pool for services.
- **defaultNetwork.type**
  - Cluster network provider, such as OpenShift SDN or OVN-Kubernetes.

You can specify the cluster network provider configuration for your cluster by setting the fields for the **defaultNetwork** object in the CNO object named **cluster**.

#### 1.5.7.1. Cluster Network Operator configuration object

The fields for the Cluster Network Operator (CNO) are described in the following table:

<table>
<thead>
<tr>
<th>Field</th>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>metadata.name</td>
<td>string</td>
<td>The name of the CNO object. This name is always <strong>cluster</strong>.</td>
</tr>
</tbody>
</table>
The `spec.clusterNetwork` field is an array of objects that specify the blocks of IP addresses from which pod IP addresses are allocated and the subnet prefix length assigned to each individual node in the cluster. For example:

```
spec:
  clusterNetwork:
    - cidr: 10.128.0.0/19
      hostPrefix: 23
    - cidr: 10.128.32.0/19
      hostPrefix: 23
```

This value is ready-only and specified in the `install-config.yaml` file.

The `spec.serviceNetwork` field is an array that specifies a block of IP addresses for services. The OpenShift SDN and OVN-Kubernetes Container Network Interface (CNI) network providers support only a single IP address block for the service network. For example:

```
spec:
  serviceNetwork:
    - 172.30.0.0/14
```

This value is ready-only and specified in the `install-config.yaml` file.

The `spec.defaultNetwork` field is an object that configures the Container Network Interface (CNI) cluster network provider for the cluster network.

The `spec.kubeProxyConfig` field is an object that specifies the kube-proxy configuration. If you are using the OVN-Kubernetes cluster network provider, the kube-proxy configuration has no effect.

### defaultNetwork object configuration

The values for the `defaultNetwork` object are defined in the following table:

<table>
<thead>
<tr>
<th>Field</th>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>spec.clusterNetwork</code> array</td>
<td></td>
<td>A list specifying the blocks of IP addresses from which pod IP addresses are allocated and the subnet prefix length assigned to each individual node in the cluster. For example:</td>
</tr>
<tr>
<td><code>spec.serviceNetwork</code> array</td>
<td></td>
<td>A block of IP addresses for services. The OpenShift SDN and OVN-Kubernetes Container Network Interface (CNI) network providers support only a single IP address block for the service network. For example:</td>
</tr>
<tr>
<td><code>spec.defaultNetwork</code> object</td>
<td></td>
<td>Configures the Container Network Interface (CNI) cluster network provider for the cluster network.</td>
</tr>
<tr>
<td><code>spec.kubeProxyConfig</code> object</td>
<td></td>
<td>The fields for this object specify the kube-proxy configuration. If you are using the OVN-Kubernetes cluster network provider, the kube-proxy configuration has no effect.</td>
</tr>
</tbody>
</table>
Either **OpenShiftSDN** or **OVNKubernetes**. The cluster network provider is selected during installation. This value cannot be changed after cluster installation.

**NOTE**

OpenShift Container Platform uses the OpenShift SDN Container Network Interface (CNI) cluster network provider by default.

<table>
<thead>
<tr>
<th>Field</th>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>type</strong></td>
<td><strong>string</strong></td>
<td>Either <strong>OpenShiftSDN</strong> or <strong>OVNKubernetes</strong>. The cluster network provider is selected during installation. This value cannot be changed after cluster installation.</td>
</tr>
</tbody>
</table>

**openshiftSDNConfig** object

This object is only valid for the OpenShift SDN cluster network provider.

**ovnKubernetesConfig** object

This object is only valid for the OVN-Kubernetes cluster network provider.

**Configuration for the OpenShift SDN CNI cluster network provider**

The following table describes the configuration fields for the OpenShift SDN Container Network Interface (CNI) cluster network provider.

**Table 1.11. openshiftSDNConfig object**

<table>
<thead>
<tr>
<th>Field</th>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>mode</strong></td>
<td><strong>string</strong></td>
<td>Configures the network isolation mode for OpenShift SDN. The default value is <strong>NetworkPolicy</strong>. The values <strong>Multitenant</strong> and <strong>Subnet</strong> are available for backwards compatibility with OpenShift Container Platform 3.x but are not recommended. This value cannot be changed after cluster installation.</td>
</tr>
</tbody>
</table>
The maximum transmission unit (MTU) for the VXLAN overlay network. This is detected automatically based on the MTU of the primary network interface. You do not normally need to override the detected MTU.

If the auto-detected value is not what you expected it to be, confirm that the MTU on the primary network interface on your nodes is correct. You cannot use this option to change the MTU value of the primary network interface on the nodes.

If your cluster requires different MTU values for different nodes, you must set this value to 50 less than the lowest MTU value in your cluster. For example, if some nodes in your cluster have an MTU of 9001, and some have an MTU of 1500, you must set this value to 1450.

This value cannot be changed after cluster installation.

The port to use for all VXLAN packets. The default value is 4789. This value cannot be changed after cluster installation.

If you are running in a virtualized environment with existing nodes that are part of another VXLAN network, then you might be required to change this. For example, when running an OpenShift SDN overlay on top of VMware NSX-T, you must select an alternate port for the VXLAN, because both SDNs use the same default VXLAN port number.

On Amazon Web Services (AWS), you can select an alternate port for the VXLAN between port 9000 and port 9999.

### Example OpenShift SDN configuration

```yaml
defaultNetwork:
  type: OpenShiftSDN
  openshiftSDNConfig:
    mode: NetworkPolicy
    mtu: 1450
    vxlanPort: 4789
```

### Configuration for the OVN-Kubernetes CNI cluster network provider

The following table describes the configuration fields for the OVN-Kubernetes CNI cluster network provider.

#### Table 1.12. ovnKubernetesConfig object

<table>
<thead>
<tr>
<th>Field</th>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>defaultNetwork</td>
<td>object</td>
<td></td>
</tr>
<tr>
<td>type</td>
<td>OpenShiftSDN</td>
<td></td>
</tr>
<tr>
<td>openshiftSDNConfig</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mode</td>
<td>NetworkPolicy</td>
<td></td>
</tr>
<tr>
<td>mtu</td>
<td>integer</td>
<td></td>
</tr>
<tr>
<td>vxlanPort</td>
<td>integer</td>
<td></td>
</tr>
</tbody>
</table>

---

79
The maximum transmission unit (MTU) for the Geneve (Generic Network Virtualization Encapsulation) overlay network. This is detected automatically based on the MTU of the primary network interface. You do not normally need to override the detected MTU.

If the auto-detected value is not what you expected it to be, confirm that the MTU on the primary network interface on your nodes is correct. You cannot use this option to change the MTU value of the primary network interface on the nodes.

If your cluster requires different MTU values for different nodes, you must set this value to 100 less than the lowest MTU value in your cluster. For example, if some nodes in your cluster have an MTU of 9001, and some have an MTU of 1500, you must set this value to 1400.

This value cannot be changed after cluster installation.

genevePort integer

The port to use for all Geneve packets. The default value is 6081. This value cannot be changed after cluster installation.

Example OVN-Kubernetes configuration

defaultNetwork:
  type: OVNKubernetes
  ovnKubernetesConfig:
    mtu: 1400
    genevePort: 6081

kubeProxyConfig object configuration
The values for the kubeProxyConfig object are defined in the following table:

<table>
<thead>
<tr>
<th>Field</th>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>defaultNetwork:</td>
<td>type: OVNKubernetes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ovnKubernetesConfig:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>mtu: 1400</td>
<td></td>
</tr>
<tr>
<td></td>
<td>genevePort: 6081</td>
<td></td>
</tr>
</tbody>
</table>
### Field | Type | Description
--- | --- | ---
**iptablesSyncPeriod** | string | The refresh period for *iptables* rules. The default value is `30s`. Valid suffixes include `s`, `m`, and `h` and are described in the *Go time* package documentation.

**proxyArguments.iptables-min-sync-period** | array | The minimum duration before refreshing *iptables* rules. This field ensures that the refresh does not happen too frequently. Valid suffixes include `s`, `m`, and `h` and are described in the *Go time* package. The default value is:

```yaml
kubeProxyConfig:
  proxyArguments:
    iptables-min-sync-period:
      - 0s
```

### 1.5.8. Specifying advanced network configuration

You can use advanced configuration customization to integrate your cluster into your existing network environment by specifying additional configuration for your cluster network provider. You can specify advanced network configuration only before you install the cluster.

**IMPORTANT**

Modifying the OpenShift Container Platform manifest files created by the installation program is not supported. Applying a manifest file that you create, as in the following procedure, is supported.

#### Prerequisites

- Create the `install-config.yaml` file and complete any modifications to it.

#### Procedure

1. Change to the directory that contains the installation program and create the manifests:

   ```bash
   $ ./openshift-install create manifests --dir <installation_directory>
   ```

   where:

   ```bash
   <installation_directory>
   ```
Specifies the name of the directory that contains the `install-config.yaml` file for your cluster.

2. Create a stub manifest file for the advanced network configuration that is named `cluster-network-03-config.yml` in the `<installation_directory>/manifests/` directory:

```bash
$ cat <<EOF > <installation_directory>/manifests/cluster-network-03-config.yml
apiVersion: operator.openshift.io/v1
kind: Network
metadata:
  name: cluster
spec:
EOF
```

where:

```
<installation_directory>
```

Specifies the directory name that contains the `manifests/` directory for your cluster.

3. Open the `cluster-network-03-config.yml` file in an editor and specify the advanced network configuration for your cluster, such as in the following example:

**Specify a different VXLAN port for the OpenShift SDN network provider**

```yaml
apiVersion: operator.openshift.io/v1
kind: Network
metadata:
  name: cluster
spec:
defaultNetwork:
  openshiftSDNConfig:
    vxlanPort: 4800
```

4. Save the `cluster-network-03-config.yml` file and quit the text editor.

5. Optional: Back up the `manifests/cluster-network-03-config.yml` file. The installation program deletes the `manifests/` directory when creating the cluster.

**NOTE**

For more information on using a Network Load Balancer (NLB) on AWS, see Configuring Ingress cluster traffic on AWS using a Network Load Balancer.

### 1.5.9. Configuring an Ingress Controller Network Load Balancer on a new AWS cluster

You can create an Ingress Controller backed by an AWS Network Load Balancer (NLB) on a new cluster.

**Prerequisites**

- Create the `install-config.yaml` file and complete any modifications to it.

**Procedure**

Create an Ingress Controller backed by an AWS NLB on a new cluster.
1. Change to the directory that contains the installation program and create the manifests:

   $ ./openshift-install create manifests --dir <installation_directory>

   For `<installation_directory>`, specify the name of the directory that contains the `install-config.yaml` file for your cluster.

2. Create a file that is named `cluster-ingress-default-ingresscontroller.yaml` in the `<installation_directory>/manifests/` directory:

   $ touch <installation_directory>/manifests/cluster-ingress-default-ingresscontroller.yaml

   For `<installation_directory>`, specify the directory name that contains the `manifests/` directory for your cluster.

   After creating the file, several network configuration files are in the `manifests/` directory, as shown:

   $ ls <installation_directory>/manifests/cluster-ingress-default-ingresscontroller.yaml

   **Example output**

   `cluster-ingress-default-ingresscontroller.yaml`

3. Open the `cluster-ingress-default-ingresscontroller.yaml` file in an editor and enter a CR that describes the Operator configuration you want:

   ```yaml
   apiVersion: operator.openshift.io/v1
   kind: IngressController
   metadata:
     creationTimestamp: null
     name: default
     namespace: openshift-ingress-operator
   spec:
     endpointPublishingStrategy:
       loadBalancer:
         scope: External
         providerParameters:
          type: AWS
          aws:
            type: NLB
            type: LoadBalancerService
   ```

4. Save the `cluster-ingress-default-ingresscontroller.yaml` file and quit the text editor.

5. Optional: Back up the `manifests/cluster-ingress-default-ingresscontroller.yaml` file. The installation program deletes the `manifests/` directory when creating the cluster.

1.5.10. Configuring hybrid networking with OVN-Kubernetes
You can configure your cluster to use hybrid networking with OVN-Kubernetes. This allows a hybrid cluster that supports different node networking configurations. For example, this is necessary to run both Linux and Windows nodes in a cluster.

**IMPORTANT**

You must configure hybrid networking with OVN-Kubernetes during the installation of your cluster. You cannot switch to hybrid networking after the installation process.

**Prerequisites**

- You defined `OVNKubernetes` for the `networking.networkType` parameter in the `install-config.yaml` file. See the installation documentation for configuring OpenShift Container Platform network customizations on your chosen cloud provider for more information.

**Procedure**

1. Change to the directory that contains the installation program and create the manifests:

   ```
   $ ./openshift-install create manifests --dir <installation_directory>
   ```

   where:

   `<installation_directory>`

   Specifies the name of the directory that contains the `install-config.yaml` file for your cluster.

2. Create a stub manifest file for the advanced network configuration that is named `cluster-network-03-config.yml` in the `<installation_directory>/manifests/` directory:

   ```
   $ cat <<EOF > <installation_directory>/manifests/cluster-network-03-config.yml
   apiVersion: operator.openshift.io/v1
   kind: Network
   metadata:
     name: cluster
   spec:
   EOF
   ```

   where:

   `<installation_directory>`

   Specifies the directory name that contains the `manifests/` directory for your cluster.

3. Open the `cluster-network-03-config.yml` file in an editor and configure OVN-Kubernetes with hybrid networking, such as in the following example:

   **Specify a hybrid networking configuration**

   ```
   apiVersion: operator.openshift.io/v1
   kind: Network
   metadata:
     name: cluster
   spec:
     defaultNetwork:
   ```
Specify the CIDR configuration used for nodes on the additional overlay network. The `hybridClusterNetwork` CIDR cannot overlap with the `clusterNetwork` CIDR.

Specify a custom VXLAN port for the additional overlay network. This is required for running Windows nodes in a cluster installed on vSphere, and must not be configured for any other cloud provider. The custom port can be any open port excluding the default 4789 port. For more information on this requirement, see the Microsoft documentation on Pod-to-pod connectivity between hosts is broken.

4. Save the `cluster-network-03-config.yml` file and quit the text editor.

5. Optional: Back up the `manifests/cluster-network-03-config.yml` file. The installation program deletes the `manifests/` directory when creating the cluster.

**NOTE**

For more information on using Linux and Windows nodes in the same cluster, see Understanding Windows container workloads.

### 1.5.11. Deploying the cluster

You can install OpenShift Container Platform on a compatible cloud platform.

**IMPORTANT**

You can run the `create cluster` command of the installation program only once, during initial installation.

**Prerequisites**

- Configure an account with the cloud platform that hosts your cluster.
- Obtain the OpenShift Container Platform installation program and the pull secret for your cluster.

**Procedure**

1. Change to the directory that contains the installation program and initialize the cluster deployment:

   ```bash
   $ ./openshift-install create cluster --dir <installation_directory> \  
   --log-level=info
   ```

   1. For `<installation_directory>`, specify the location of your customized `./install-config.yaml` file.
To view different installation details, specify `warn`, `debug`, or `error` instead of `info`.

**NOTE**

If the cloud provider account that you configured on your host does not have sufficient permissions to deploy the cluster, the installation process stops, and the missing permissions are displayed.

When the cluster deployment completes, directions for accessing your cluster, including a link to its web console and credentials for the `kubeadmin` user, display in your terminal.

**Example output**

...  
INFO Install complete!  
INFO To access the cluster as the system:admin user when using 'oc', run 'export KUBECONFIG=/home/myuser/install_dir/auth/kubeconfig'  
INFO Access the OpenShift web-console here: https://console-openshift-console.apps.mycluster.example.com  
INFO Login to the console with user: "kubeadmin", and password: "4vYBz-Ee6gm-yMBZj-Wt5AL"  
INFO Time elapsed: 36m22s

**NOTE**

The cluster access and credential information also outputs to `<installation_directory>/openshift_install.log` when an installation succeeds.

**IMPORTANT**

The Ignition config files that the installation program generates contain certificates that expire after 24 hours, which are then renewed at that time. If the cluster is shut down before renewing the certificates and the cluster is later restarted after the 24 hours have elapsed, the cluster automatically recovers the expired certificates. The exception is that you must manually approve the pending `node-bootstrapper` certificate signing requests (CSRs) to recover `kubelet` certificates. See the documentation for `Recovering from expired control plane certificates` for more information.

**IMPORTANT**

You must not delete the installation program or the files that the installation program creates. Both are required to delete the cluster.

2. Optional: Remove or disable the `AdministratorAccess` policy from the IAM account that you used to install the cluster.

**NOTE**

The elevated permissions provided by the `AdministratorAccess` policy are required only during installation.
1.5.12. Installing the OpenShift CLI by downloading the binary

You can install the OpenShift CLI (oc) in order to interact with OpenShift Container Platform from a command-line interface. You can install oc on Linux, Windows, or macOS.

**IMPORTANT**

If you installed an earlier version of oc, you cannot use it to complete all of the commands in OpenShift Container Platform 4.6. Download and install the new version of oc.

1.5.12.1. Installing the OpenShift CLI on Linux

You can install the OpenShift CLI (oc) binary on Linux by using the following procedure.

**Procedure**


2. Select the appropriate version in the Version drop-down menu.

3. Click Download Now next to the OpenShift v4.6 Linux Client entry and save the file.

4. Unpack the archive:

   ```$ tar xvzf <file>```

5. Place the oc binary in a directory that is on your PATH.
   To check your PATH, execute the following command:

   ```$ echo $PATH```

After you install the OpenShift CLI, it is available using the oc command:

```$ oc <command>```

1.5.12.2. Installing the OpenShift CLI on Windows

You can install the OpenShift CLI (oc) binary on Windows by using the following procedure.

**Procedure**


2. Select the appropriate version in the Version drop-down menu.

3. Click Download Now next to the OpenShift v4.6 Windows Client entry and save the file.

4. Unzip the archive with a ZIP program.

5. Move the oc binary to a directory that is on your PATH.
   To check your PATH, open the command prompt and execute the following command:

   ```$ echo $PATH```

   After you install the OpenShift CLI, it is available using the oc command:

   ```$ oc <command>```
After you install the OpenShift CLI, it is available using the `oc` command:

```bash
C:\> oc <command>
```

### 1.5.12.3. Installing the OpenShift CLI on macOS

You can install the OpenShift CLI (`oc`) binary on macOS by using the following procedure.

**Procedure**

2. Select the appropriate version in the **Version** drop-down menu.
3. Click **Download Now** next to the **OpenShift v4.6 MacOSX Client** entry and save the file.
4. Unpack and unzip the archive.
5. Move the `oc` binary to a directory on your PATH.
   To check your **PATH**, open a terminal and execute the following command:

```bash
$ echo $PATH
```

After you install the OpenShift CLI, it is available using the `oc` command:

```bash
$ oc <command>
```

### 1.5.13. Logging in to the cluster by using the CLI

You can log in to your cluster as a default system user by exporting the cluster `kubeconfig` file. The `kubeconfig` file contains information about the cluster that is used by the CLI to connect a client to the correct cluster and API server. The file is specific to a cluster and is created during OpenShift Container Platform installation.

**Prerequisites**

- You deployed an OpenShift Container Platform cluster.
- You installed the `oc` CLI.

**Procedure**

1. Export the `kubeadmin` credentials:

```bash
$ export KUBECONFIG=<installation_directory>/auth/kubeconfig
```

   ![1] For **<installation_directory>**, specify the path to the directory that you stored the installation files in.
2. Verify you can run `oc` commands successfully using the exported configuration:

```
$ oc whoami
```

**Example output**

```
system:admin
```

### 1.5.14. Logging in to the cluster by using the web console

The `kubeadmin` user exists by default after an OpenShift Container Platform installation. You can log in to your cluster as the `kubeadmin` user by using the OpenShift Container Platform web console.

**Prerequisites**

- You have access to the installation host.
- You completed a cluster installation and all cluster Operators are available.

**Procedure**

1. Obtain the password for the `kubeadmin` user from the `kubeadmin-password` file on the installation host:

```
$ cat <installation_directory>/auth/kubeadmin-password
```

**NOTE**

Alternatively, you can obtain the `kubeadmin` password from the `<installation_directory>/openshift_install.log` log file on the installation host.

2. List the OpenShift Container Platform web console route:

```
$ oc get routes -n openshift-console | grep 'console-openshift'
```

**Example output**

```
console     console-openshift-console.apps.<cluster_name>.<base_domain>            console
https   reencrypt/Redirect   None
```

3. Navigate to the route detailed in the output of the preceding command in a web browser and log in as the `kubeadmin` user.

**Additional resources**

- See [Accessing the web console](#) for more details about accessing and understanding the output of `oc whoami`.

```bash
system:admin
```

$ cat <installation_directory>/auth/kubeadmin-password

```
console     console-openshift-console.apps.<cluster_name>.<base_domain>            console
https   reencrypt/Redirect   None
```
• See Accessing the web console for more details about accessing and understanding the OpenShift Container Platform web console.

1.5.15. Telemetry access for OpenShift Container Platform

In OpenShift Container Platform 4.6, the Telemetry service, which runs by default to provide metrics about cluster health and the success of updates, requires Internet access. If your cluster is connected to the Internet, Telemetry runs automatically, and your cluster is registered to the Red Hat OpenShift Cluster Manager (OCM).

After you confirm that your Red Hat OpenShift Cluster Manager inventory is correct, either maintained automatically by Telemetry or manually by using OCM, use subscription watch to track your OpenShift Container Platform subscriptions at the account or multi-cluster level.

Additional resources
• See About remote health monitoring for more information about the Telemetry service.

1.5.16. Next steps
• Validating an installation.
• Customize your cluster.
• If necessary, you can opt out of remote health reporting.
• If necessary, you can remove cloud provider credentials.

1.6. INSTALLING A CLUSTER ON AWS INTO AN EXISTING VPC

In OpenShift Container Platform version 4.6, you can install a cluster into an existing Amazon Virtual Private Cloud (VPC) on Amazon Web Services (AWS). The installation program provisions the rest of the required infrastructure, which you can further customize. To customize the installation, you modify parameters in the install-config.yaml file before you install the cluster.

1.6.1. Prerequisites
• Review details about the OpenShift Container Platform installation and update processes.
• Configure an AWS account to host the cluster.

IMPORTANT
If you have an AWS profile stored on your computer, it must not use a temporary session token that you generated while using a multi-factor authentication device. The cluster continues to use your current AWS credentials to create AWS resources for the entire life of the cluster, so you must use long-lived credentials. To generate appropriate keys, see Managing Access Keys for IAM Users in the AWS documentation. You can supply the keys when you run the installation program.

• If you use a firewall, you must configure it to allow the sites that your cluster requires access to.
• If you do not allow the system to manage identity and access management (IAM), then a cluster administrator can manually create and maintain IAM credentials. Manual mode can also be used in environments where the cloud IAM APIs are not reachable.

1.6.2. About using a custom VPC

In OpenShift Container Platform 4.6, you can deploy a cluster into existing subnets in an existing Amazon Virtual Private Cloud (VPC) in Amazon Web Services (AWS). By deploying OpenShift Container Platform into an existing AWS VPC, you might be able to avoid limit constraints in new accounts or more easily abide by the operational constraints that your company’s guidelines set. If you cannot obtain the infrastructure creation permissions that are required to create the VPC yourself, use this installation option.

Because the installation program cannot know what other components are also in your existing subnets, it cannot choose subnet CIDRs and so forth on your behalf. You must configure networking for the subnets that you install your cluster to yourself.

1.6.2.1. Requirements for using your VPC

The installation program no longer creates the following components:

- Internet gateways
- NAT gateways
- Subnets
- Route tables
- VPCs
- VPC DHCP options
- VPC endpoints

If you use a custom VPC, you must correctly configure it and its subnets for the installation program and the cluster to use. See Amazon VPC console wizard configurations and Work with VPCs and subnets in the AWS documentation for more information on creating and managing an AWS VPC.

The installation program cannot:

- Subdivide network ranges for the cluster to use.
- Set route tables for the subnets.
- Set VPC options like DHCP.

You must complete these tasks before you install the cluster. See VPC networking components and Route tables for your VPC for more information on configuring networking in an AWS VPC.

Your VPC must meet the following characteristics:

- Create a public and private subnet for each availability zone that your cluster uses. Each availability zone can contain no more than one public and one private subnet. For an example of this type of configuration, see VPC with public and private subnets (NAT) in the AWS documentation.
Record each subnet ID. Completing the installation requires that you enter these values in the platform section of the install-config.yaml file. See Finding a subnet ID in the AWS documentation.

- The VPC’s CIDR block must contain the Networking.MachineCIDR range, which is the IP address pool for cluster machines. The subnet CIDR blocks must belong to the machine CIDR that you specify.

- The VPC must have a public internet gateway attached to it. For each availability zone:
  - The public subnet requires a route to the internet gateway.
  - The public subnet requires a NAT gateway with an EIP address.
  - The private subnet requires a route to the NAT gateway in public subnet.

- The VPC must not use the kubernetes.io/cluster/.*: owned tag. The installation program modifies your subnets to add the kubernetes.io/cluster/.*: shared tag, so your subnets must have at least one free tag slot available for it. See Tag Restrictions in the AWS documentation to confirm that the installation program can add a tag to each subnet that you specify.

- You must enable the enableDnsSupport and enableDnsHostnames attributes in your VPC, so that the cluster can use the Route 53 zones that are attached to the VPC to resolve cluster’s internal DNS records. See DNS Support in Your VPC in the AWS documentation. If you prefer to use your own Route 53 hosted private zone, you must associate the existing hosted zone with your VPC prior to installing a cluster. You can define your hosted zone using the platform.aws.hostedZone field in the install-config.yaml file.

If you are working in a disconnected environment, you are unable to reach the public IP addresses for EC2 and ELB endpoints. To resolve this, you must create a VPC endpoint and attach it to the subnet that the clusters are using. The endpoints should be named as follows:

- ec2.<region>.amazonaws.com
- elasticloadbalancing.<region>.amazonaws.com
- s3.<region>.amazonaws.com

Required VPC components

You must provide a suitable VPC and subnets that allow communication to your machines.

<table>
<thead>
<tr>
<th>Component</th>
<th>AWS type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>VPC</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>AWS::EC2::VPC</td>
<td>You must provide a public VPC for the cluster to use. The VPC uses an endpoint that references the route tables for each subnet to improve communication with the registry that is hosted in S3.</td>
</tr>
<tr>
<td></td>
<td>AWS::EC2::VPCEndpoint</td>
<td></td>
</tr>
</tbody>
</table>

If you are working in a disconnected environment, you are unable to reach the public IP addresses for EC2 and ELB endpoints. To resolve this, you must create a VPC endpoint and attach it to the subnet that the clusters are using. The endpoints should be named as follows:

- ec2.<region>.amazonaws.com
- elasticloadbalancing.<region>.amazonaws.com
- s3.<region>.amazonaws.com

Required VPC components

You must provide a suitable VPC and subnets that allow communication to your machines.
<table>
<thead>
<tr>
<th>Component</th>
<th>AWS type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Public subnets</td>
<td>● AWS::EC2::Subnet</td>
<td>Your VPC must have public subnets for between 1 and 3 availability zones and associate them with appropriate Ingress rules.</td>
</tr>
<tr>
<td></td>
<td>● AWS::EC2::SubnetNetworkAclAss</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ociation</td>
<td></td>
</tr>
<tr>
<td>Internet gateway</td>
<td>● AWS::EC2::InternetGateway</td>
<td>You must have a public Internet gateway, with public routes, attached to the VPC. In the provided templates, each public subnet has a NAT gateway with an EIP address. These NAT gateways allow cluster resources, like private subnet instances, to reach the Internet and are not required for some restricted network or proxy scenarios.</td>
</tr>
<tr>
<td></td>
<td>● AWS::EC2::VPCGatewayAttachment</td>
<td></td>
</tr>
<tr>
<td></td>
<td>● AWS::EC2::RouteTable</td>
<td></td>
</tr>
<tr>
<td></td>
<td>● AWS::EC2::Route</td>
<td></td>
</tr>
<tr>
<td></td>
<td>● AWS::EC2::SubnetRouteTableAss</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ociation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>● AWS::EC2::NatGateway</td>
<td></td>
</tr>
<tr>
<td></td>
<td>● AWS::EC2::EIP</td>
<td></td>
</tr>
<tr>
<td>Network access</td>
<td>● AWS::EC2::NetworkAcl</td>
<td>You must allow the VPC to access the following ports:</td>
</tr>
<tr>
<td>control</td>
<td>● AWS::EC2::NetworkAclEntry</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Port</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>80</td>
</tr>
<tr>
<td></td>
<td></td>
<td>443</td>
</tr>
<tr>
<td></td>
<td></td>
<td>22</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1024 - 65535</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0 - 65535</td>
</tr>
<tr>
<td>Private subnets</td>
<td>● AWS::EC2::Subnet</td>
<td>Your VPC can have private subnets. The provided CloudFormation templates can create private subnets for between 1 and 3 availability zones. If you use private subnets, you must provide appropriate routes and tables for them.</td>
</tr>
<tr>
<td></td>
<td>● AWS::EC2::RouteTable</td>
<td></td>
</tr>
<tr>
<td></td>
<td>● AWS::EC2::SubnetRouteTableAss</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ociation</td>
<td></td>
</tr>
</tbody>
</table>
1.6.2.2. VPC validation

To ensure that the subnets that you provide are suitable, the installation program confirms the following data:

- All the subnets that you specify exist.
- You provide private subnets.
- The subnet CIDRs belong to the machine CIDR that you specified.
- You provide subnets for each availability zone. Each availability zone contains no more than one public and one private subnet. If you use a private cluster, provide only a private subnet for each availability zone. Otherwise, provide exactly one public and private subnet for each availability zone.
- You provide a public subnet for each private subnet availability zone. Machines are not provisioned in availability zones that you do not provide private subnets for.

If you destroy a cluster that uses an existing VPC, the VPC is not deleted. When you remove the OpenShift Container Platform cluster from a VPC, the `kubernetes.io/cluster/*/shared` tag is removed from the subnets that it used.

1.6.2.3. Division of permissions

Starting with OpenShift Container Platform 4.3, you do not need all of the permissions that are required for an installation program-provisioned infrastructure cluster to deploy a cluster. This change mimics the division of permissions that you might have at your company: some individuals can create different resource in your clouds than others. For example, you might be able to create application-specific items, like instances, buckets, and load balancers, but not networking-related components such as VPCs, subnets, or ingress rules.

The AWS credentials that you use when you create your cluster do not need the networking permissions that are required to make VPCs and core networking components within the VPC, such as subnets, routing tables, Internet gateways, NAT, and VPN. You still need permission to make the application resources that the machines within the cluster require, such as ELBs, security groups, S3 buckets, and nodes.

1.6.2.4. Isolation between clusters

If you deploy OpenShift Container Platform to an existing network, the isolation of cluster services is reduced in the following ways:

- You can install multiple OpenShift Container Platform clusters in the same VPC.
- ICMP ingress is allowed from the entire network.
- TCP 22 ingress (SSH) is allowed to the entire network.
- Control plane TCP 6443 ingress (Kubernetes API) is allowed to the entire network.
- Control plane TCP 22623 ingress (MCS) is allowed to the entire network.

1.6.3. Internet access for OpenShift Container Platform

In OpenShift Container Platform 4.6, you require access to the Internet to install your cluster.
You must have Internet access to:

- Access the Red Hat OpenShift Cluster Manager page to download the installation program and perform subscription management. If the cluster has Internet access and you do not disable Telemetry, that service automatically entitles your cluster.

- Access Quay.io to obtain the packages that are required to install your cluster.

- Obtain the packages that are required to perform cluster updates.

**IMPORTANT**

If your cluster cannot have direct Internet access, you can perform a restricted network installation on some types of infrastructure that you provision. During that process, you download the content that is required and use it to populate a mirror registry with the packages that you need to install a cluster and generate the installation program. With some installation types, the environment that you install your cluster in will not require Internet access. Before you update the cluster, you update the content of the mirror registry.

1.6.4. Generating an SSH private key and adding it to the agent

If you want to perform installation debugging or disaster recovery on your cluster, you must provide an SSH key to both your `ssh-agent` and the installation program. You can use this key to access the bootstrap machine in a public cluster to troubleshoot installation issues.

**NOTE**

In a production environment, you require disaster recovery and debugging.

You can use this key to SSH into the master nodes as the user `core`. When you deploy the cluster, the key is added to the `core` user’s `~/.ssh/authorized_keys` list.

**NOTE**

You must use a local key, not one that you configured with platform-specific approaches such as AWS key pairs.

**Procedure**

1. If you do not have an SSH key that is configured for password-less authentication on your computer, create one. For example, on a computer that uses a Linux operating system, run the following command:

```
$ ssh-keygen -t ed25519 -N "" \
-f <path>/<file_name> 1
```

1 Specify the path and file name, such as `~/.ssh/id_rsa`, of the new SSH key. If you have an existing key pair, ensure your public key is in the your `~/.ssh` directory.

Running this command generates an SSH key that does not require a password in the location that you specified.
If you plan to install an OpenShift Container Platform cluster that uses FIPS Validated / Modules in Process cryptographic libraries on the x86_64 architecture, do not create a key that uses the ed25519 algorithm. Instead, create a key that uses the rsa or ecdsa algorithm.

2. Start the ssh-agent process as a background task:

   $ eval "$(ssh-agent -s)"

Example output

   Agent pid 31874

   NOTE

   If your cluster is in FIPS mode, only use FIPS-compliant algorithms to generate the SSH key. The key must be either RSA or ECDSA.

3. Add your SSH private key to the ssh-agent:

   $ ssh-add <path>/<file_name>

Example output

   Identity added: /home/<you>/<path>/<file_name> (<computer_name>)

   Specify the path and file name for your SSH private key, such as ~/.ssh/id_rsa

Next steps

- When you install OpenShift Container Platform, provide the SSH public key to the installation program.

1.6.5. Obtaining the installation program

Before you install OpenShift Container Platform, download the installation file on a local computer.

Prerequisites

- You have a computer that runs Linux or macOS, with 500 MB of local disk space

Procedure

1. Access the Infrastructure Provider page on the Red Hat OpenShift Cluster Manager site. If you have a Red Hat account, log in with your credentials. If you do not, create an account.

2. Select your infrastructure provider.
3. Navigate to the page for your installation type, download the installation program for your operating system, and place the file in the directory where you will store the installation configuration files.

**IMPORTANT**

The installation program creates several files on the computer that you use to install your cluster. You must keep the installation program and the files that the installation program creates after you finish installing the cluster. Both files are required to delete the cluster.

**IMPORTANT**

Deleting the files created by the installation program does not remove your cluster, even if the cluster failed during installation. To remove your cluster, complete the OpenShift Container Platform uninstallation procedures for your specific cloud provider.

4. Extract the installation program. For example, on a computer that uses a Linux operating system, run the following command:

```
$ tar xvf openshift-install-linux.tar.gz
```

5. From the Pull Secret page on the Red Hat OpenShift Cluster Manager site, download your installation pull secret as a `.txt` file. This pull secret allows you to authenticate with the services that are provided by the included authorities, including Quay.io, which serves the container images for OpenShift Container Platform components.

1.6.6. Creating the installation configuration file

You can customize the OpenShift Container Platform cluster you install on Amazon Web Services (AWS).

**Prerequisites**

- Obtain the OpenShift Container Platform installation program and the pull secret for your cluster.

**Procedure**

1. Create the `install-config.yaml` file.

   a. Change to the directory that contains the installation program and run the following command:

   ```
   $ ./openshift-install create install-config --dir <installation_directory>
   ```

   **1** For `<installation_directory>`, specify the directory name to store the files that the installation program creates.
IMPORTANT

Specify an empty directory. Some installation assets, like bootstrap X.509 certificates have short expiration intervals, so you must not reuse an installation directory. If you want to reuse individual files from another cluster installation, you can copy them into your directory. However, the file names for the installation assets might change between releases. Use caution when copying installation files from an earlier OpenShift Container Platform version.

b. At the prompts, provide the configuration details for your cloud:
   
i. Optional: Select an SSH key to use to access your cluster machines.

   **NOTE**
   
   For production OpenShift Container Platform clusters on which you want to perform installation debugging or disaster recovery, specify an SSH key that your `ssh-agent` process uses.

   ii. Select **AWS** as the platform to target.

   iii. If you do not have an Amazon Web Services (AWS) profile stored on your computer, enter the AWS access key ID and secret access key for the user that you configured to run the installation program.

   iv. Select the AWS region to deploy the cluster to.

   v. Select the base domain for the Route 53 service that you configured for your cluster.

   vi. Enter a descriptive name for your cluster.

   vii. Paste the pull secret that you obtained from the Pull Secret page on the Red Hat OpenShift Cluster Manager site.

2. Modify the `install-config.yaml` file. You can find more information about the available parameters in the Installation configuration parameters section.

3. Back up the `install-config.yaml` file so that you can use it to install multiple clusters.

   **IMPORTANT**
   
   The `install-config.yaml` file is consumed during the installation process. If you want to reuse the file, you must back it up now.

1.6.6.1. Installation configuration parameters

Before you deploy an OpenShift Container Platform cluster, you provide parameter values to describe your account on the cloud platform that hosts your cluster and optionally customize your cluster’s platform. When you create the `install-config.yaml` installation configuration file, you provide values for the required parameters through the command line. If you customize your cluster, you can modify the `install-config.yaml` file to provide more details about the platform.
NOTE
After installation, you cannot modify these parameters in the `install-config.yaml` file.

IMPORTANT
The `openshift-install` command does not validate field names for parameters. If an incorrect name is specified, the related file or object is not created, and no error is reported. Ensure that the field names for any parameters that are specified are correct.

1.6.6.1.1. Required configuration parameters

Required installation configuration parameters are described in the following table:

Table 1.14. Required parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>apiVersion</td>
<td>The API version for the <code>install-config.yaml</code> content. The current version is <code>v1</code>. The installer may also support older API versions.</td>
<td>String</td>
</tr>
<tr>
<td>baseDomain</td>
<td>The base domain of your cloud provider. The base domain is used to create routes to your OpenShift Container Platform cluster components. The full DNS name for your cluster is a combination of the <code>baseDomain</code> and <code>metadata.name</code> parameter values that uses the <code>&lt;metadata.name&gt;.&lt;baseDomain&gt;</code> format.</td>
<td>A fully-qualified domain or subdomain name, such as example.com.</td>
</tr>
<tr>
<td>metadata</td>
<td>Kubernetes resource <code>ObjectMeta</code>, from which only the <code>name</code> parameter is consumed.</td>
<td>Object</td>
</tr>
<tr>
<td>metadata.name</td>
<td>The name of the cluster. DNS records for the cluster are all subdomains of <code>{{.metadata.name}}.{{.baseDomain}}</code>.</td>
<td>String of lowercase letters, hyphens (-), and periods (.), such as dev.</td>
</tr>
</tbody>
</table>
The configuration for the specific platform upon which to perform the installation: **aws, baremetal, azure, openstack, ovirt, vsphere.** For additional information about platform.<platform> parameters, consult the following table for your specific platform.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>platform</td>
<td>The configuration for the specific platform upon which to perform the installation: <strong>aws, baremetal, azure, openstack, ovirt, vsphere.</strong> For additional information about platform.&lt;platform&gt; parameters, consult the following table for your specific platform.</td>
<td>Object</td>
</tr>
</tbody>
</table>

Get a pull secret from [https://console.redhat.com/openshift/install/pull-secret](https://console.redhat.com/openshift/install/pull-secret) to authenticate downloading container images for OpenShift Container Platform components from services such as Quay.io.

```json
{
  "auths":{
    "cloud.openshift.com":{
      "auth":"b3Blb=",
      "email":"you@example.com"
    },
    "quay.io":{
      "auth":"b3Blb=",
      "email":"you@example.com"
    }
  }
}
```

1.6.6.1.2. Network configuration parameters

You can customize your installation configuration based on the requirements of your existing network infrastructure. For example, you can expand the IP address block for the cluster network or provide different IP address blocks than the defaults.

Only IPv4 addresses are supported.

### Table 1.15. Network parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>networking</td>
<td>The configuration for the cluster network.</td>
<td>Object</td>
</tr>
</tbody>
</table>

**NOTE**

You cannot modify parameters specified by the networking object after installation.
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>networking.network Type</td>
<td>The cluster network provider Container Network Interface (CNI) plug-in to install.</td>
<td>Either OpenShiftSDN or OVNKubernetes. The default value is OpenShiftSDN.</td>
</tr>
<tr>
<td>networking.clusterNetwork</td>
<td>The IP address blocks for pods.</td>
<td>An array of objects. For example:</td>
</tr>
<tr>
<td></td>
<td>The default value is 10.128.0.0/14 with a host prefix of /23.</td>
<td>networking:</td>
</tr>
<tr>
<td></td>
<td>If you specify multiple IP address blocks, the blocks must not overlap.</td>
<td>clusterNetwork:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- cidr: 10.128.0.0/14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>hostPrefix: 23</td>
</tr>
<tr>
<td>networking.clusterNetwork.cidr</td>
<td>Required if you use networking.clusterNetwork. An IP address block.</td>
<td>An IP address block in Classless Inter-Domain Routing (CIDR) notation. The prefix length for an IPv4 block is between 0 and 32.</td>
</tr>
<tr>
<td></td>
<td>An IPv4 network.</td>
<td></td>
</tr>
<tr>
<td>networking.clusterNetwork.hostPrefix</td>
<td>The subnet prefix length to assign to each individual node. For example, if hostPrefix is set to 23 then each node is assigned a /23 subnet out of the given cidr. A hostPrefix value of 23 provides 510 (2^(32 - 23) - 2) pod IP addresses.</td>
<td>A subnet prefix.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The default value is 23.</td>
</tr>
<tr>
<td>networking.serviceNetwork</td>
<td>The IP address block for services. The default value is 172.30.0.0/16.</td>
<td>An array with an IP address block in CIDR format. For example:</td>
</tr>
<tr>
<td></td>
<td>The OpenShift SDN and OVN-Kubernetes network providers support only a single IP address block for the service network.</td>
<td>networking:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>serviceNetwork:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- 172.30.0.0/16</td>
</tr>
<tr>
<td>networking.machineNetwork</td>
<td>The IP address blocks for machines.</td>
<td>An array of objects. For example:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>networking:</td>
</tr>
<tr>
<td></td>
<td>If you specify multiple IP address blocks, the blocks must not overlap.</td>
<td>machineNetwork:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- cidr: 10.0.0.0/16</td>
</tr>
</tbody>
</table>
1.6.6.3. Optional configuration parameters

Optional installation configuration parameters are described in the following table:

Table 1.16. Optional parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Values</th>
</tr>
</thead>
</table>
| networking.machin
  Network.cidr           | Required if you use networking.machin
  NetworkNetwork. An IP address block. The default value is 10.0.0.0/16 for all platforms other than libvirt. For libvirt, the default value is 192.168.126.0/24. | An IP network block in CIDR notation. For example, 10.0.0.0/16. |

**NOTE**

Set the networking.machin
eNetworkNetwork to match the CIDR that the preferred NIC resides in.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Values</th>
</tr>
</thead>
</table>
| additionalTrustBund
  lele                     | A PEM-encoded X.509 certificate bundle that is added to the nodes’ trusted certificate store. This trust bundle may also be used when a proxy has been configured. | String                                           |
| compute                 | The configuration for the machines that comprise the compute nodes.         | Array of machine-pool objects. For details, see the following "Machine-pool" table. |
| compute.architectur
  e                     | Determines the instruction set architecture of the machines in the pool. Currently, heterogeneous clusters are not supported, so all pools must specify the same architecture. Valid values are amd64 (the default). | String                                           |
Whether to enable or disable simultaneous multithreading, or **hyperthreading**, on compute machines. By default, simultaneous multithreading is enabled to increase the performance of your machines’ cores.

**IMPORTANT**

If you disable simultaneous multithreading, ensure that your capacity planning accounts for the dramatically decreased machine performance.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>compute.hyperthreading</strong></td>
<td>Whether to enable or disable simultaneous multithreading, or <strong>hyperthreading</strong>, on compute machines. By default, simultaneous multithreading is enabled to increase the performance of your machines’ cores.</td>
<td><strong>Enabled</strong> or <strong>Disabled</strong></td>
</tr>
<tr>
<td><strong>compute.name</strong></td>
<td>Required if you use <strong>compute</strong>. The name of the machine pool.</td>
<td><strong>worker</strong></td>
</tr>
<tr>
<td><strong>compute.platform</strong></td>
<td>Required if you use <strong>compute</strong>. Use this parameter to specify the cloud provider to host the worker machines. This parameter value must match the <strong>controlPlane.platform</strong> parameter value.</td>
<td><strong>aws</strong>, <strong>azure</strong>, <strong>gcp</strong>, <strong>openstack</strong>, <strong>ovirt</strong>, <strong>vsphere</strong>, or <strong>{}</strong></td>
</tr>
<tr>
<td><strong>compute.replicas</strong></td>
<td>The number of compute machines, which are also known as worker machines, to provision.</td>
<td>A positive integer greater than or equal to 2. The default value is 3.</td>
</tr>
<tr>
<td><strong>controlPlane</strong></td>
<td>The configuration for the machines that comprise the control plane.</td>
<td>Array of <strong>MachinePool</strong> objects. For details, see the following &quot;Machine-pool&quot; table.</td>
</tr>
<tr>
<td><strong>controlPlane.architecture</strong></td>
<td>Determines the instruction set architecture of the machines in the pool. Currently, heterogeneous clusters are not supported, so all pools must specify the same architecture. Valid values are <strong>amd64</strong> (the default).</td>
<td><strong>String</strong></td>
</tr>
<tr>
<td>Parameter</td>
<td>Description</td>
<td>Values</td>
</tr>
<tr>
<td>----------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>-----------------------------</td>
</tr>
<tr>
<td>controlPlane.hypertreading</td>
<td>Whether to enable or disable simultaneous multithreading, or hyperthreading, on control plane machines. By default, simultaneous multithreading is enabled to increase the performance of your machines’ cores.</td>
<td>Enabled or Disabled</td>
</tr>
<tr>
<td></td>
<td><strong>IMPORTANT</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>If you disable simultaneous multithreading, ensure that your capacity planning accounts for the dramatically decreased machine performance.</td>
<td></td>
</tr>
<tr>
<td>controlPlane.name</td>
<td>Required if you use controlPlane. The name of the machine pool.</td>
<td>master</td>
</tr>
<tr>
<td>controlPlane.platform</td>
<td>Required if you use controlPlane. Use this parameter to specify the cloud provider that hosts the control plane machines. This parameter value must match the compute.platform parameter value.</td>
<td>aws, azure, gcp, openstack, ovirt, vsphere, or {}</td>
</tr>
<tr>
<td>controlPlane.replicas</td>
<td>The number of control plane machines to provision.</td>
<td>The only supported value is 3, which is the default value.</td>
</tr>
<tr>
<td>Parameter</td>
<td>Description</td>
<td>Values</td>
</tr>
<tr>
<td>-----------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>credentialsMode</td>
<td>The Cloud Credential Operator (CCO) mode. If no mode is specified, the CCO dynamically tries to determine the capabilities of the provided credentials, with a preference for mint mode on the platforms where multiple modes are supported.</td>
<td><em>Mint, Passthrough, Manual</em>, or an empty string (&quot;&quot;&quot;).</td>
</tr>
</tbody>
</table>

**NOTE**

Not all CCO modes are supported for all cloud providers. For more information on CCO modes, see the Cloud Credential Operator entry in the Red Hat Operators reference content.
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>fips</td>
<td>Enable or disable FIPS mode. The default is <strong>false</strong> (disabled). If FIPS mode is enabled, the Red Hat Enterprise Linux CoreOS (RHCOS) machines that OpenShift Container Platform runs on bypass the default Kubernetes cryptography suite and use the cryptography modules that are provided with RHCOS instead.</td>
<td><strong>false</strong> or <strong>true</strong></td>
</tr>
<tr>
<td></td>
<td><strong>IMPORTANT</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>The use of FIPS Validated / Modules in Process cryptographic libraries is only supported on OpenShift Container Platform deployments on the <strong>x86_64</strong> architecture.</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>NOTE</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>If you are using Azure File storage, you cannot enable FIPS mode.</td>
<td></td>
</tr>
<tr>
<td>imageContentSources</td>
<td>Sources and repositories for the release-image content.</td>
<td>Array of objects. Includes a <strong>source</strong> and, optionally, <strong>mirrors</strong>, as described in the following rows of this table.</td>
</tr>
<tr>
<td>imageContentSources.source</td>
<td>Required if you use <strong>imageContentSources</strong>. Specify the repository that users refer to, for example, in image pull specifications.</td>
<td><strong>String</strong></td>
</tr>
<tr>
<td>imageContentSources.mirrors</td>
<td>Specify one or more repositories that may also contain the same images.</td>
<td>Array of strings</td>
</tr>
<tr>
<td>publish</td>
<td>How to publish or expose the user-facing endpoints of your cluster, such as the Kubernetes API, OpenShift routes.</td>
<td><strong>Internal</strong> or <strong>External</strong>. To deploy a private cluster, which cannot be accessed from the internet, set <strong>publish</strong> to <strong>Internal</strong>. The default value is <strong>External</strong>.</td>
</tr>
</tbody>
</table>
The SSH key or keys to authenticate access your cluster machines.

**NOTE**
For production OpenShift Container Platform clusters on which you want to perform installation debugging or disaster recovery, specify an SSH key that your `ssh-agent` process uses.

One or more keys. For example:

```
sshKey:
  <key1>
  <key2>
  <key3>
```

1.6.6.1.4. Optional AWS configuration parameters

Optional AWS configuration parameters are described in the following table:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>compute.platform.aws.amiID</code></td>
<td>The AWS AMI used to boot compute machines for the cluster. This is required for regions that require a custom RHCOS AMI.</td>
<td>Any published or custom RHCOS AMI that belongs to the set AWS region.</td>
</tr>
<tr>
<td><code>compute.platform.aws.rootVolume.size</code></td>
<td>The size in GiB of the root volume.</td>
<td>Integer, for example <strong>500</strong>.</td>
</tr>
<tr>
<td><code>compute.platform.aws.rootVolume.type</code></td>
<td>The type of the root volume.</td>
<td>Valid AWS EBS volume type, such as <strong>io1</strong>.</td>
</tr>
<tr>
<td><code>compute.platform.aws.type</code></td>
<td>The EC2 instance type for the compute machines.</td>
<td>Valid AWS instance type, such as <strong>c5.9xlarge</strong>.</td>
</tr>
<tr>
<td>Parameter</td>
<td>Description</td>
<td>Values</td>
</tr>
<tr>
<td>-----------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>compute.platform.aws.zones</td>
<td>The availability zones where the installation program creates machines for</td>
<td>A list of valid AWS availability zones, such as <strong>us-east-1c</strong>, in a YAML sequence.</td>
</tr>
<tr>
<td></td>
<td>the compute machine pool. If you provide your own VPC, you must provide a</td>
<td></td>
</tr>
<tr>
<td></td>
<td>subnet in that availability zone.</td>
<td></td>
</tr>
<tr>
<td>compute.aws.region</td>
<td>The AWS region that the installation program creates compute resources in.</td>
<td>Any valid <strong>AWS region</strong>, such as <strong>us-east-1</strong>.</td>
</tr>
<tr>
<td>controlPlane.platform.aws.amiID</td>
<td>The AWS AMI used to boot control plane machines for the cluster. This is</td>
<td>Any published or custom RHCOS AMI that belongs to the set AWS region.</td>
</tr>
<tr>
<td></td>
<td>required for regions that require a custom RHCOS AMI.</td>
<td></td>
</tr>
<tr>
<td>controlPlane.platform.aws.type</td>
<td>The EC2 instance type for the control plane machines.</td>
<td>Valid <strong>AWS instance type</strong>, such as <strong>c5.9xlarge</strong>.</td>
</tr>
<tr>
<td>controlPlane.platform.aws.zones</td>
<td>The availability zones where the installation program creates machines for</td>
<td>A list of valid AWS availability zones, such as <strong>us-east-1c</strong>, in a YAML sequence.</td>
</tr>
<tr>
<td></td>
<td>the control plane machine pool.</td>
<td></td>
</tr>
<tr>
<td>controlPlane.aws.region</td>
<td>The AWS region that the installation program creates control plane resources in.</td>
<td>Valid <strong>AWS region</strong>, such as <strong>us-east-1</strong>.</td>
</tr>
<tr>
<td>platform.aws.amiID</td>
<td>The AWS AMI used to boot all machines for the cluster. If set, the AMI must</td>
<td>Any published or custom RHCOS AMI that belongs to the set AWS region.</td>
</tr>
<tr>
<td></td>
<td>belong to the same region as the cluster. This is required for regions that</td>
<td></td>
</tr>
<tr>
<td></td>
<td>require a custom RHCOS AMI.</td>
<td></td>
</tr>
<tr>
<td>platform.aws.serviceEndpoints.name</td>
<td>The AWS service endpoint name. Custom endpoints are only required for cases</td>
<td>Valid <strong>AWS service endpoint name</strong>.</td>
</tr>
<tr>
<td></td>
<td>where alternative AWS endpoints, like FIPS, must be used. Custom API</td>
<td></td>
</tr>
<tr>
<td></td>
<td>endpoints can be specified for EC2, S3, IAM, Elastic Load Balancing,</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tagging, Route 53, and STS AWS services.</td>
<td></td>
</tr>
</tbody>
</table>
## Parameter | Description | Values
--- | --- | ---
platform.aws.serviceEndpoints.url | The AWS service endpoint URL. The URL must use the https protocol and the host must trust the certificate. | Valid AWS service endpoint URL.
platform.aws.userTags | A map of keys and values that the installation program adds as tags to all resources that it creates. | Any valid YAML map, such as key value pairs in the `<key>: <value>` format. For more information about AWS tags, see Tagging Your Amazon EC2 Resources in the AWS documentation.
platform.aws.subnets | If you provide the VPC instead of allowing the installation program to create the VPC for you, specify the subnet for the cluster to use. The subnet must be part of the same `machineNetwork[].cidr` ranges that you specify. For a standard cluster, specify a public and a private subnet for each availability zone. For a private cluster, specify a private subnet for each availability zone. | Valid subnet IDs.

### 1.6.6.2. Sample customized `install-config.yaml` file for AWS

You can customize the `install-config.yaml` file to specify more details about your OpenShift Container Platform cluster’s platform or modify the values of the required parameters.

**IMPORTANT**

This sample YAML file is provided for reference only. You must obtain your `install-config.yaml` file by using the installation program and modify it.

```yaml
apiVersion: v1
baseDomain: example.com
credentialsMode: Mint
controlPlane:
  hyperthreading: Enabled
  name: master
platform:
  aws:
    zones:
    - us-west-2a
    - us-west-2b
    rootVolume:
      iops: 4000
      size: 500
```
Required. The installation program prompts you for this value.

Optional: Add this parameter to force the Cloud Credential Operator (CCO) to use the specified mode, instead of having the CCO dynamically try to determine the capabilities of the credentials. For details about CCO modes, see the Cloud Credential Operator entry in the Red Hat Operators reference content.

If you do not provide these parameters and values, the installation program provides the default
The controlPlane section is a single mapping, but the compute section is a sequence of mappings. To meet the requirements of the different data structures, the first line of the compute section

Whether to enable or disable simultaneous multithreading, or hyperthreading. By default, simultaneous multithreading is enabled to increase the performance of your machines’ cores. You can disable it by setting the parameter value to Disabled. If you disable simultaneous multithreading in some cluster machines, you must disable it in all cluster machines.

IMPORTANT
If you disable simultaneous multithreading, ensure that your capacity planning accounts for the dramatically decreased machine performance. Use larger instance types, such as m4.2xlarge or m5.2xlarge, for your machines if you disable simultaneous multithreading.

To configure faster storage for etcd, especially for larger clusters, set the storage type as io1 and set iops to 2000.

If you provide your own VPC, specify subnets for each availability zone that your cluster uses.

The ID of the AMI used to boot machines for the cluster. If set, the AMI must belong to the same region as the cluster.

The AWS service endpoints. Custom endpoints are required when installing to an unknown AWS region. The endpoint URL must use the https protocol and the host must trust the certificate.

The ID of your existing Route 53 private hosted zone. Providing an existing hosted zone requires that you supply your own VPC and the hosted zone is already associated with the VPC prior to installing your cluster. If undefined, the installation program creates a new hosted zone.

Whether to enable or disable FIPS mode. By default, FIPS mode is not enabled. If FIPS mode is enabled, the Red Hat Enterprise Linux CoreOS (RHCOS) machines that OpenShift Container Platform runs on bypass the default Kubernetes cryptography suite and use the cryptography modules that are provided with RHCOS instead.

IMPORTANT
The use of FIPS Validated / Modules in Process cryptographic libraries is only supported on OpenShift Container Platform deployments on the x86_64 architecture.

You can optionally provide the sshKey value that you use to access the machines in your cluster.

NOTE
For production OpenShift Container Platform clusters on which you want to perform installation debugging or disaster recovery, specify an SSH key that your ssh-agent process uses.

1.6.6.3. Configuring the cluster-wide proxy during installation
Production environments can deny direct access to the Internet and instead have an HTTP or HTTPS proxy available. You can configure a new OpenShift Container Platform cluster to use a proxy by configuring the proxy settings in the `install-config.yaml` file.

**Prerequisites**

- You have an existing `install-config.yaml` file.
- You reviewed the sites that your cluster requires access to and determined whether any of them need to bypass the proxy. By default, all cluster egress traffic is proxied, including calls to hosting cloud provider APIs. You added sites to the `Proxy` object’s `spec.noProxy` field to bypass the proxy if necessary.

**NOTE**

The `Proxy` object `status.noProxy` field is populated with the values of the `networking.machineNetwork[].cidr`, `networking.clusterNetwork[].cidr`, and `networking.serviceNetwork[]` fields from your installation configuration.

For installations on Amazon Web Services (AWS), Google Cloud Platform (GCP), Microsoft Azure, and Red Hat OpenStack Platform (RHOSP), the `Proxy` object `status.noProxy` field is also populated with the instance metadata endpoint (169.254.169.254).

- If your cluster is on AWS, you added the `ec2.<region>.amazonaws.com`, `elasticloadbalancing.<region>.amazonaws.com`, and `s3.<region>.amazonaws.com` endpoints to your VPC endpoint. These endpoints are required to complete requests from the nodes to the AWS EC2 API. Because the proxy works on the container level, not the node level, you must route these requests to the AWS EC2 API through the AWS private network. Adding the public IP address of the EC2 API to your allowlist in your proxy server is not sufficient.

**Procedure**

1. Edit your `install-config.yaml` file and add the proxy settings. For example:

   ```yaml
   apiVersion: v1
   baseDomain: my.domain.com
   proxy:
     httpProxy: http://<username>:<pswd>@<ip>:<port> ¹
     httpsProxy: https://<username>:<pswd>@<ip>:<port> ²
     noProxy: example.com ³
   additionalTrustBundle: |
     -----BEGIN CERTIFICATE-----
     <MY_TRUSTED_CA_CERT>
     -----END CERTIFICATE-----
   ... ⁴
   ``

   - ¹ A proxy URL to use for creating HTTP connections outside the cluster. The URL scheme must be `http`.
   - ² A proxy URL to use for creating HTTPS connections outside the cluster.
   - ³ A comma-separated list of destination domain names, IP addresses, or other network
     apiVersion: v1
     baseDomain: my.domain.com
     proxy:
     httpProxy: http://<username>:<pswd>@<ip>:<port> ¹
     httpsProxy: https://<username>:<pswd>@<ip>:<port> ²
     noProxy: example.com ³
     additionalTrustBundle: |
     -----BEGIN CERTIFICATE-----
     <MY_TRUSTED_CA_CERT>
     -----END CERTIFICATE-----
   ...

   - ¹ A proxy URL to use for creating HTTP connections outside the cluster. The URL scheme must be `http`.
   - ² A proxy URL to use for creating HTTPS connections outside the cluster.
   - ³ A comma-separated list of destination domain names, IP addresses, or other network
A comma-separated list of destination domain names, IP addresses, or other network CIDRs to exclude from proxying. Preface a domain with . to match subdomains only. For

If provided, the installation program generates a config map that is named user-ca-bundle in the openshift-config namespace to hold the additional CA certificates. If you provide additionalTrustBundle and at least one proxy setting, the Proxy object is configured to reference the user-ca-bundle config map in the trustedCA field. The Cluster Network Operator then creates a trusted-ca-bundle config map that merges the contents specified for the trustedCA parameter with the RHCOS trust bundle. The additionalTrustBundle field is required unless the proxy’s identity certificate is signed by an authority from the RHCOS trust bundle.

NOTE

The installation program does not support the proxy readinessEndpoints field.

2. Save the file and reference it when installing OpenShift Container Platform.

The installation program creates a cluster-wide proxy that is named cluster that uses the proxy settings in the provided install-config.yaml file. If no proxy settings are provided, a cluster Proxy object is still created, but it will have a nil spec.

NOTE

Only the Proxy object named cluster is supported, and no additional proxies can be created.

1.6.7. Deploying the cluster

You can install OpenShift Container Platform on a compatible cloud platform.

IMPORTANT

You can run the create cluster command of the installation program only once, during initial installation.

Prerequisites

- Configure an account with the cloud platform that hosts your cluster.
- Obtain the OpenShift Container Platform installation program and the pull secret for your cluster.

Procedure

1. Change to the directory that contains the installation program and initialize the cluster deployment:

   ```
   $ ./openshift-install create cluster --dir <installation_directory> \ 
      --log-level=info
   ```

   For <installation_directory>, specify the location of your customized ./install-config.yaml file.
To view different installation details, specify **warn**, **debug**, or **error** instead of **info**.

**NOTE**

If the cloud provider account that you configured on your host does not have sufficient permissions to deploy the cluster, the installation process stops, and the missing permissions are displayed.

When the cluster deployment completes, directions for accessing your cluster, including a link to its web console and credentials for the **kubeadmin** user, display in your terminal.

**Example output**

...  
INFO Install complete!  
INFO To access the cluster as the system:admin user when using 'oc', run 'export KUBECONFIG=/home/myuser/install_dir/auth/kubeconfig'  
INFO Access the OpenShift web-console here: https://console-openshift-console.apps.mycluster.example.com  
INFO Login to the console with user: "kubeadmin", and password: "4vYBz-Ee6gm-ymBZj-Wi5AL"  
INFO Time elapsed: 36m22s

**NOTE**

The cluster access and credential information also outputs to `<installation_directory>/openshift_install.log` when an installation succeeds.

**IMPORTANT**

The Ignition config files that the installation program generates contain certificates that expire after 24 hours, which are then renewed at that time. If the cluster is shut down before renewing the certificates and the cluster is later restarted after the 24 hours have elapsed, the cluster automatically recovers the expired certificates. The exception is that you must manually approve the pending **node-bootstrapper** certificate signing requests (CSRs) to recover kubelet certificates. See the documentation for **Recovering from expired control plane certificates** for more information.

**IMPORTANT**

You must not delete the installation program or the files that the installation program creates. Both are required to delete the cluster.

2. Optional: Remove or disable the **AdministratorAccess** policy from the IAM account that you used to install the cluster.

**NOTE**

The elevated permissions provided by the **AdministratorAccess** policy are required only during installation.
1.6.8. Installing the OpenShift CLI by downloading the binary

You can install the OpenShift CLI (oc) in order to interact with OpenShift Container Platform from a command-line interface. You can install oc on Linux, Windows, or macOS.

**IMPORTANT**

If you installed an earlier version of oc, you cannot use it to complete all of the commands in OpenShift Container Platform 4.6. Download and install the new version of oc.

1.6.8.1. Installing the OpenShift CLI on Linux

You can install the OpenShift CLI (oc) binary on Linux by using the following procedure.

**Procedure**

2. Select the appropriate version in the Version drop-down menu.
3. Click Download Now next to the OpenShift v4.6 Linux Client entry and save the file.
4. Unpack the archive:
   
   ```
   $ tar xvzf <file>
   ```
5. Place the oc binary in a directory that is on your PATH.
   To check your PATH, execute the following command:
   
   ```
   $ echo $PATH
   ```

After you install the OpenShift CLI, it is available using the oc command:

```
$ oc <command>
```
1.6.8.3. Installing the OpenShift CLI on macOS

You can install the OpenShift CLI (oc) binary on macOS by using the following procedure.

**Procedure**

2. Select the appropriate version in the Version drop-down menu.
3. Click Download Now next to the OpenShift v4.6 MacOSX Client entry and save the file.
4. Unpack and unzip the archive.
5. Move the oc binary to a directory on your PATH.
   To check your PATH, open a terminal and execute the following command:

   ```
   $ echo $PATH
   ```

After you install the OpenShift CLI, it is available using the oc command:

```
C:\> oc <command>
```

1.6.9. Logging in to the cluster by using the CLI

You can log in to your cluster as a default system user by exporting the cluster kubeconfig file. The kubeconfig file contains information about the cluster that is used by the CLI to connect a client to the correct cluster and API server. The file is specific to a cluster and is created during OpenShift Container Platform installation.

**Prerequisites**

- You deployed an OpenShift Container Platform cluster.
- You installed the oc CLI.

**Procedure**

1. Export the kubeadmin credentials:

   ```
   $ export KUBECONFIG=<installation_directory>/auth/kubeconfig
   ```

   For `<installation_directory>`, specify the path to the directory that you stored the installation files in.

```
$ oc <command>
```
2. Verify you can run `oc` commands successfully using the exported configuration:

```
$ oc whoami
```

**Example output**

```
system:admin
```

### 1.6.10. Logging in to the cluster by using the web console

The `kubeadmin` user exists by default after an OpenShift Container Platform installation. You can log in to your cluster as the `kubeadmin` user by using the OpenShift Container Platform web console.

**Prerequisites**

- You have access to the installation host.
- You completed a cluster installation and all cluster Operators are available.

**Procedure**

1. Obtain the password for the `kubeadmin` user from the `kubeadmin-password` file on the installation host:

```
$ cat <installation_directory>/auth/kubeadmin-password
```

**NOTE**

Alternatively, you can obtain the `kubeadmin` password from the `<installation_directory>/openshift_install.log` log file on the installation host.

2. List the OpenShift Container Platform web console route:

```
$ oc get routes -n openshift-console | grep 'console-openshift'
```

**NOTE**

Alternatively, you can obtain the OpenShift Container Platform route from the `<installation_directory>/openshift_install.log` log file on the installation host.

**Example output**

```
console  console-openshift-console.apps.<cluster_name>.<base_domain>  console
https   reencrypt/Redirect   None
```

3. Navigate to the route detailed in the output of the preceding command in a web browser and log in as the `kubeadmin` user.

**Additional resources**

- See [Accessing the web console](#) for more details about accessing and understanding the `$ oc whoami` system:admin
- cat <installation_directory>/auth/kubeadmin-password
- oc get routes -n openshift-console | grep 'console-openshift'
- console  console-openshift-console.apps.<cluster_name>.<base_domain>  console
- https   reencrypt/Redirect   None
- Navigate to the route detailed in the output of the preceding command in a web browser and log in as the `kubeadmin` user.
1.6.11. Telemetry access for OpenShift Container Platform

In OpenShift Container Platform 4.6, the Telemetry service, which runs by default to provide metrics about cluster health and the success of updates, requires Internet access. If your cluster is connected to the Internet, Telemetry runs automatically, and your cluster is registered to the Red Hat OpenShift Cluster Manager (OCM).

After you confirm that your Red Hat OpenShift Cluster Manager inventory is correct, either maintained automatically by Telemetry or manually by using OCM, use subscription watch to track your OpenShift Container Platform subscriptions at the account or multi-cluster level.

Additional resources

- See About remote health monitoring for more information about the Telemetry service.

1.6.12. Next steps

- Validating an installation.
- Customize your cluster.
- If necessary, you can opt out of remote health reporting.
- If necessary, you can remove cloud provider credentials.

1.7. INSTALLING A PRIVATE CLUSTER ON AWS

In OpenShift Container Platform version 4.6, you can install a private cluster into an existing VPC on Amazon Web Services (AWS). The installation program provisions the rest of the required infrastructure, which you can further customize. To customize the installation, you modify parameters in the install-config.yaml file before you install the cluster.

1.7.1. Prerequisites

- Review details about the OpenShift Container Platform installation and update processes.
- Configure an AWS account to host the cluster.

IMPORTANT

If you have an AWS profile stored on your computer, it must not use a temporary session token that you generated while using a multi-factor authentication device. The cluster continues to use your current AWS credentials to create AWS resources for the entire life of the cluster, so you must use long-lived credentials. To generate appropriate keys, see Managing Access Keys for IAM Users in the AWS documentation. You can supply the keys when you run the installation program.

- If you use a firewall, you must configure it to allow the sites that your cluster requires access to.
If you do not allow the system to manage identity and access management (IAM), then a cluster administrator can manually create and maintain IAM credentials. Manual mode can also be used in environments where the cloud IAM APIs are not reachable.

1.7.2. Private clusters

You can deploy a private OpenShift Container Platform cluster that does not expose external endpoints. Private clusters are accessible from only an internal network and are not visible to the Internet.

By default, OpenShift Container Platform is provisioned to use publicly-accessible DNS and endpoints. A private cluster sets the DNS, Ingress Controller, and API server to private when you deploy your cluster. This means that the cluster resources are only accessible from your internal network and are not visible to the internet.

To deploy a private cluster, you must use existing networking that meets your requirements. Your cluster resources might be shared between other clusters on the network.

Additionally, you must deploy a private cluster from a machine that has access the API services for the cloud you provision to, the hosts on the network that you provision, and to the internet to obtain installation media. You can use any machine that meets these access requirements and follows your company's guidelines. For example, this machine can be a bastion host on your cloud network or a machine that has access to the network through a VPN.

1.7.2.1. Private clusters in AWS

To create a private cluster on Amazon Web Services (AWS), you must provide an existing private VPC and subnets to host the cluster. The installation program must also be able to resolve the DNS records that the cluster requires. The installation program configures the Ingress Operator and API server for access from only the private network.

The cluster still requires access to Internet to access the AWS APIs.

The following items are not required or created when you install a private cluster:

- Public subnets
- Public load balancers, which support public ingress
- A public Route 53 zone that matches the baseDomain for the cluster

The installation program does use the baseDomain that you specify to create a private Route 53 zone and the required records for the cluster. The cluster is configured so that the Operators do not create public records for the cluster and all cluster machines are placed in the private subnets that you specify.

1.7.2.1.1. Limitations

The ability to add public functionality to a private cluster is limited.

- You cannot make the Kubernetes API endpoints public after installation without taking additional actions, including creating public subnets in the VPC for each availability zone in use, creating a public load balancer, and configuring the control plane security groups to allow traffic from Internet on 6443 (Kubernetes API port).
If you use a public Service type load balancer, you must tag a public subnet in each availability zone with `kubernetes.io/cluster/<cluster-infra-id>: shared` so that AWS can use them to create public load balancers.

1.7.3. About using a custom VPC

In OpenShift Container Platform 4.6, you can deploy a cluster into existing subnets in an existing Amazon Virtual Private Cloud (VPC) in Amazon Web Services (AWS). By deploying OpenShift Container Platform into an existing AWS VPC, you might be able to avoid limit constraints in new accounts or more easily abide by the operational constraints that your company’s guidelines set. If you cannot obtain the infrastructure creation permissions that are required to create the VPC yourself, use this installation option.

Because the installation program cannot know what other components are also in your existing subnets, it cannot choose subnet CIDRs and so forth on your behalf. You must configure networking for the subnets that you install your cluster to yourself.

1.7.3.1. Requirements for using your VPC

The installation program no longer creates the following components:

- Internet gateways
- NAT gateways
- Subnets
- Route tables
- VPCs
- VPC DHCP options
- VPC endpoints

If you use a custom VPC, you must correctly configure it and its subnets for the installation program and the cluster to use. See Amazon VPC console wizard configurations and Work with VPCs and subnets in the AWS documentation for more information on creating and managing an AWS VPC.

The installation program cannot:

- Subdivide network ranges for the cluster to use.
- Set route tables for the subnets.
- Set VPC options like DHCP.

You must complete these tasks before you install the cluster. See VPC networking components and Route tables for your VPC for more information on configuring networking in an AWS VPC.

Your VPC must meet the following characteristics:

- The VPC must not use the `kubernetes.io/cluster/.*: owned` tag.
  The installation program modifies your subnets to add the `kubernetes.io/cluster/.*: shared` tag, so your subnets must have at least one free tag slot available for it. See Tag Restrictions in the AWS documentation to confirm that the installation program can add a tag to each subnet that you specify.
You must enable the `enableDnsSupport` and `enableDnsHostnames` attributes in your VPC, so that the cluster can use the Route 53 zones that are attached to the VPC to resolve cluster’s internal DNS records. See DNS Support in Your VPC in the AWS documentation. If you prefer to use your own Route 53 hosted private zone, you must associate the existing hosted zone with your VPC prior to installing a cluster. You can define your hosted zone using the `platform.aws.hostedZone` field in the `install-config.yaml` file.

- If you use a cluster with public access, you must create a public and a private subnet for each availability zone that your cluster uses. Each availability zone can contain no more than one public and one private subnet.

If you are working in a disconnected environment, you are unable to reach the public IP addresses for EC2 and ELB endpoints. To resolve this, you must create a VPC endpoint and attach it to the subnet that the clusters are using. The endpoints should be named as follows:

- `ec2.<region>.amazonaws.com`
- `elasticloadbalancing.<region>.amazonaws.com`
- `s3.<region>.amazonaws.com`

**Required VPC components**

You must provide a suitable VPC and subnets that allow communication to your machines.

<table>
<thead>
<tr>
<th>Component</th>
<th>AWS type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>VPC</td>
<td>- AWS::EC2::VPC&lt;br&gt;- AWS::EC2::VPCEndpoint</td>
<td>You must provide a public VPC for the cluster to use. The VPC uses an endpoint that references the route tables for each subnet to improve communication with the registry that is hosted in S3.</td>
</tr>
<tr>
<td>Public subnets</td>
<td>- AWS::EC2::Subnet&lt;br&gt;- AWS::EC2::SubnetNetworkAclAssociation</td>
<td>Your VPC must have public subnets for between 1 and 3 availability zones and associate them with appropriate Ingress rules.</td>
</tr>
<tr>
<td>Internet gateway</td>
<td>- AWS::EC2::InternetGateway&lt;br&gt;- AWS::EC2::VPCGatewayAttachment&lt;br&gt;- AWS::EC2::RouteTable&lt;br&gt;- AWS::EC2::Route&lt;br&gt;- AWS::EC2::SubnetRouteTableAssociation&lt;br&gt;- AWS::EC2::NatGateway&lt;br&gt;- AWS::EC2::EIP</td>
<td>You must have a public Internet gateway, with public routes, attached to the VPC. In the provided templates, each public subnet has a NAT gateway with an EIP address. These NAT gateways allow cluster resources, like private subnet instances, to reach the Internet and are not required for some restricted network or proxy scenarios.</td>
</tr>
</tbody>
</table>
Network access control

- **AWS::EC2::NetworkAcl**
- **AWS::EC2::NetworkAclEntry**

You must allow the VPC to access the following ports:

<table>
<thead>
<tr>
<th>Port</th>
<th>Reason</th>
</tr>
</thead>
<tbody>
<tr>
<td>80</td>
<td>Inbound HTTP traffic</td>
</tr>
<tr>
<td>443</td>
<td>Inbound HTTPS traffic</td>
</tr>
<tr>
<td>22</td>
<td>Inbound SSH traffic</td>
</tr>
<tr>
<td>1024 - 65535</td>
<td>Inbound ephemeral traffic</td>
</tr>
<tr>
<td>0 - 65535</td>
<td>Outbound ephemeral traffic</td>
</tr>
</tbody>
</table>

Private subnets

- **AWS::EC2::Subnet**
- **AWS::EC2::RouteTable**
- **AWS::EC2::SubnetRouteTableAssociation**

Your VPC can have private subnets. The provided CloudFormation templates can create private subnets for between 1 and 3 availability zones. If you use private subnets, you must provide appropriate routes and tables for them.

1.7.3.2. VPC validation

To ensure that the subnets that you provide are suitable, the installation program confirms the following data:

- All the subnets that you specify exist.
- You provide private subnets.
- The subnet CIDRs belong to the machine CIDR that you specified.
- You provide subnets for each availability zone. Each availability zone contains no more than one public and one private subnet. If you use a private cluster, provide only a private subnet for each availability zone. Otherwise, provide exactly one public and private subnet for each availability zone.
- You provide a public subnet for each private subnet availability zone. Machines are not provisioned in availability zones that you do not provide private subnets for.
If you destroy a cluster that uses an existing VPC, the VPC is not deleted. When you remove the OpenShift Container Platform cluster from a VPC, the `kubernetes.io/cluster/*: shared` tag is removed from the subnets that it used.

### 1.7.3.3. Division of permissions

Starting with OpenShift Container Platform 4.3, you do not need all of the permissions that are required for an installation program-provisioned infrastructure cluster to deploy a cluster. This change mimics the division of permissions that you might have at your company: some individuals can create different resource in your clouds than others. For example, you might be able to create application-specific items, like instances, buckets, and load balancers, but not networking-related components such as VPCs, subnets, or ingress rules.

The AWS credentials that you use when you create your cluster do not need the networking permissions that are required to make VPCs and core networking components within the VPC, such as subnets, routing tables, Internet gateways, NAT, and VPN. You still need permission to make the application resources that the machines within the cluster require, such as ELBs, security groups, S3 buckets, and nodes.

### 1.7.3.4. Isolation between clusters

If you deploy OpenShift Container Platform to an existing network, the isolation of cluster services is reduced in the following ways:

- You can install multiple OpenShift Container Platform clusters in the same VPC.
- ICMP ingress is allowed from the entire network.
- TCP 22 ingress (SSH) is allowed to the entire network.
- Control plane TCP 6443 ingress (Kubernetes API) is allowed to the entire network.
- Control plane TCP 22623 ingress (MCS) is allowed to the entire network.

### 1.7.4. Internet access for OpenShift Container Platform

In OpenShift Container Platform 4.6, you require access to the Internet to install your cluster.

You must have Internet access to:

- Access the [Red Hat OpenShift Cluster Manager](https://cluster-managed.openshift.redhat.com) page to download the installation program and perform subscription management. If the cluster has Internet access and you do not disable Telemetry, that service automatically entitles your cluster.
- Access [Quay.io](https://quay.io) to obtain the packages that are required to install your cluster.
- Obtain the packages that are required to perform cluster updates.
1.7.5. Generating an SSH private key and adding it to the agent

If you want to perform installation debugging or disaster recovery on your cluster, you must provide an SSH key to both your ssh-agent and the installation program. You can use this key to access the bootstrap machine in a public cluster to troubleshoot installation issues.

**NOTE**

In a production environment, you require disaster recovery and debugging.

You can use this key to SSH into the master nodes as the user core. When you deploy the cluster, the key is added to the core user’s ~/.ssh/authorized_keys list.

**NOTE**

You must use a local key, not one that you configured with platform-specific approaches such as AWS key pairs.

**Procedure**

1. If you do not have an SSH key that is configured for password-less authentication on your computer, create one. For example, on a computer that uses a Linux operating system, run the following command:

```
$ ssh-keygen -t ed25519 -N "" \
-f <path>/<file_name>
```

Specify the path and file name, such as ~/.ssh/id_rsa, of the new SSH key. If you have an existing key pair, ensure your public key is in the your ~/.ssh directory.

Running this command generates an SSH key that does not require a password in the location that you specified.

**NOTE**

If you plan to install an OpenShift Container Platform cluster that uses FIPS Validated / Modules in Process cryptographic libraries on the x86_64 architecture, do not create a key that uses the ed25519 algorithm. Instead, create a key that uses the rsa or ecdsa algorithm.

2. Start the ssh-agent process as a background task:
NOTE
If your cluster is in FIPS mode, only use FIPS-compliant algorithms to generate the SSH key. The key must be either RSA or ECDSA.

3. Add your SSH private key to the `ssh-agent`:

```bash
$ eval "$(ssh-agent -s)"
```

Example output

```
Agent pid 31874
```

Example output

```
Identity added: /home/<you>/<path>/<file_name> (<computer_name>)
```

Specify the path and file name for your SSH private key, such as `~/.ssh/id_rsa`

Next steps

- When you install OpenShift Container Platform, provide the SSH public key to the installation program.

1.7.6. Obtaining the installation program

Before you install OpenShift Container Platform, download the installation file on a local computer.

Prerequisites

- You have a computer that runs Linux or macOS, with 500 MB of local disk space

Procedure

1. Access the Infrastructure Provider page on the Red Hat OpenShift Cluster Manager site. If you have a Red Hat account, log in with your credentials. If you do not, create an account.

2. Select your infrastructure provider.

3. Navigate to the page for your installation type, download the installation program for your operating system, and place the file in the directory where you will store the installation configuration files.

**IMPORTANT**

The installation program creates several files on the computer that you use to install your cluster. You must keep the installation program and the files that the installation program creates after you finish installing the cluster. Both files are required to delete the cluster.
Deleting the files created by the installation program does not remove your cluster, even if the cluster failed during installation. To remove your cluster, complete the OpenShift Container Platform uninstallation procedures for your specific cloud provider.

4. Extract the installation program. For example, on a computer that uses a Linux operating system, run the following command:

   $ tar xvf openshift-install-linux.tar.gz

5. From the Pull Secret page on the Red Hat OpenShift Cluster Manager site, download your installation pull secret as a .txt file. This pull secret allows you to authenticate with the services that are provided by the included authorities, including Quay.io, which serves the container images for OpenShift Container Platform components.

1.7.7. Manually creating the installation configuration file

For installations of a private OpenShift Container Platform cluster that are only accessible from an internal network and are not visible to the Internet, you must manually generate your installation configuration file.

Prerequisites

- Obtain the OpenShift Container Platform installation program and the access token for your cluster.

Procedure

1. Create an installation directory to store your required installation assets in:

   $ mkdir <installation_directory>

   IMPORTANT

   You must create a directory. Some installation assets, like bootstrap X.509 certificates have short expiration intervals, so you must not reuse an installation directory. If you want to reuse individual files from another cluster installation, you can copy them into your directory. However, the file names for the installation assets might change between releases. Use caution when copying installation files from an earlier OpenShift Container Platform version.

2. Customize the following install-config.yaml file template and save it in the <installation_directory>.

   NOTE

   You must name this configuration file install-config.yaml.

3. Back up the install-config.yaml file so that you can use it to install multiple clusters.
1.7.7.1. Installation configuration parameters

Before you deploy an OpenShift Container Platform cluster, you provide parameter values to describe your account on the cloud platform that hosts your cluster and optionally customize your cluster’s platform. When you create the `install-config.yaml` installation configuration file, you provide values for the required parameters through the command line. If you customize your cluster, you can modify the `install-config.yaml` file to provide more details about the platform.

**NOTE**

After installation, you cannot modify these parameters in the `install-config.yaml` file.

**IMPORTANT**

The `openshift-install` command does not validate field names for parameters. If an incorrect name is specified, the related file or object is not created, and no error is reported. Ensure that the field names for any parameters that are specified are correct.

1.7.7.1.1. Required configuration parameters

Required installation configuration parameters are described in the following table:

Table 1.18. Required parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>apiVersion</td>
<td>The API version for the <code>install-config.yaml</code> content. The current version is v1. The installer may also support older API versions.</td>
<td>String</td>
</tr>
<tr>
<td>baseDomain</td>
<td>The base domain of your cloud provider. The base domain is used to create routes to your OpenShift Container Platform cluster components. The full DNS name for your cluster is a combination of the <code>baseDomain</code> and <code>metadata.name</code> parameter values that uses the <code>&lt;metadata.name&gt;</code>.&lt;baseDomain&gt; format.</td>
<td>A fully-qualified domain or subdomain name, such as example.com.</td>
</tr>
</tbody>
</table>
**Parameter** | **Description** | **Values**
--- | --- | ---
**metadata** | Kubernetes resource `ObjectMeta`, from which only the `name` parameter is consumed. | Object

**metadata.name** | The name of the cluster. DNS records for the cluster are all subdomains of `{{.metadata.name}}.{{.baseDomain}}`. | String of lowercase letters, hyphens (-), and periods (.), such as `dev`.

**platform** | The configuration for the specific platform upon which to perform the installation: `aws`, `baremetal`, `azure`, `openstack`, `ovirt`, `vsphere`. For additional information about `platform.<platform>` parameters, consult the following table for your specific platform. | Object

**pullSecret** | Get a pull secret from [https://console.redhat.com/openshift/install/pull-secret](https://console.redhat.com/openshift/install/pull-secret) to authenticate downloading container images for OpenShift Container Platform components from services such as Quay.io. | `{  "auths":{  "cloud.openshift.com":{  "auth":"b3Blb=",  "email":"you@example.com"  },  "quay.io":{  "auth":"b3Blb=",  "email":"you@example.com"  }  }  }

1.7.7.1.2. Network configuration parameters

You can customize your installation configuration based on the requirements of your existing network infrastructure. For example, you can expand the IP address block for the cluster network or provide different IP address blocks than the defaults.

Only IPv4 addresses are supported.

**Table 1.19. Network parameters**
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>networking</td>
<td>The configuration for the cluster network.</td>
<td>Object</td>
</tr>
<tr>
<td></td>
<td><strong>NOTE</strong></td>
<td>You cannot modify parameters specified by the <code>networking</code> object after installation.</td>
</tr>
<tr>
<td>networking.network.Type</td>
<td>The cluster network provider Container Network Interface (CNI) plug-in to install.</td>
<td>Either OpenShiftSDN or OVNKubernetes. The default value is OpenShiftSDN.</td>
</tr>
<tr>
<td>networking.clusterNetwork</td>
<td>The IP address blocks for pods.</td>
<td>An array of objects. For example:</td>
</tr>
<tr>
<td></td>
<td>The default value is <strong>10.128.0.0/14</strong> with a host prefix of <strong>/23</strong>.</td>
<td>networking:</td>
</tr>
<tr>
<td></td>
<td>If you specify multiple IP address blocks, the blocks must not overlap.</td>
<td>clusterNetwork:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- cidr: 10.128.0.0/14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>hostPrefix: 23</td>
</tr>
<tr>
<td>networking.clusterNetwork.cidr</td>
<td>Required if you use <code>networking.clusterNetwork</code>. An IP address block.</td>
<td>An IP address block in Classless Inter-Domain Routing (CIDR) notation. The prefix length for an IPv4 block is between <strong>0</strong> and <strong>32</strong>.</td>
</tr>
<tr>
<td></td>
<td>An IPv4 network.</td>
<td></td>
</tr>
<tr>
<td>networking.clusterNetwork.hostPrefix</td>
<td>The subnet prefix length to assign to each individual node. For example, if <code>hostPrefix</code> is set to <strong>23</strong> then each node is assigned a /23 subnet out of the given cidr. A <code>hostPrefix</code> value of <strong>23</strong> provides 510 (2^(32 - 23) - 2) pod IP addresses.</td>
<td>A subnet prefix.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The default value is <strong>23</strong>.</td>
</tr>
<tr>
<td>networking.serviceNetwork</td>
<td>The IP address block for services. The default value is <strong>172.30.0.0/16</strong>.</td>
<td>An array with an IP address block in CIDR format. For example:</td>
</tr>
<tr>
<td></td>
<td>The OpenShift SDN and OVN-Kubernetes network providers support only a single IP address block for the service network.</td>
<td>networking:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>serviceNetwork:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- 172.30.0.0/16</td>
</tr>
<tr>
<td>networking.machineNetwork</td>
<td>The IP address blocks for machines.</td>
<td>An array of objects. For example:</td>
</tr>
<tr>
<td></td>
<td>If you specify multiple IP address blocks, the blocks must not overlap.</td>
<td>networking:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>machineNetwork:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- cidr: 10.0.0.0/16</td>
</tr>
</tbody>
</table>
**networking.machineNetwork.cidr**

Required if you use `networking.machineNetwork`. An IP address block. The default value is **10.0.0.0/16** for all platforms other than libvirt. For libvirt, the default value is **192.168.126.0/24**.

An IP network block in CIDR notation.

For example, **10.0.0.0/16**.

**NOTE**

Set the `networking.machineNetwork` to match the CIDR that the preferred NIC resides in.

---

1.7.7.1.3. Optional configuration parameters

Optional installation configuration parameters are described in the following table:

**Table 1.20. Optional parameters**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>additionalTrustBundle</td>
<td>A PEM-encoded X.509 certificate bundle that is added to the nodes’ trusted certificate store. This trust bundle may also be used when a proxy has been configured.</td>
<td>String</td>
</tr>
<tr>
<td>compute</td>
<td>The configuration for the machines that comprise the compute nodes.</td>
<td>Array of machine-pool objects. For details, see the following &quot;Machine-pool&quot; table.</td>
</tr>
<tr>
<td>compute.architecture</td>
<td>Determines the instruction set architecture of the machines in the pool. Currently, heterogeneous clusters are not supported, so all pools must specify the same architecture. Valid values are <strong>amd64</strong> (the default).</td>
<td>String</td>
</tr>
</tbody>
</table>
Whether to enable or disable simultaneous multithreading, or hyperthreading, on compute machines. By default, simultaneous multithreading is enabled to increase the performance of your machines' cores.

**IMPORTANT**

If you disable simultaneous multithreading, ensure that your capacity planning accounts for the dramatically decreased machine performance.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>compute.hyperthreading</td>
<td>Whether to enable or disable simultaneous multithreading, or hyperthreading, on compute machines. By default, simultaneous multithreading is enabled to increase the performance of your machines' cores.</td>
<td>Enabled or Disabled</td>
</tr>
<tr>
<td>compute.name</td>
<td>Required if you use compute. The name of the machine pool.</td>
<td>worker</td>
</tr>
<tr>
<td>compute.platform</td>
<td>Required if you use compute. Use this parameter to specify the cloud provider to host the worker machines. This parameter value must match the controlPlane.platform parameter value.</td>
<td>aws, azure, gcp, openstack, ovirt, vsphere, or {}</td>
</tr>
<tr>
<td>compute.replicas</td>
<td>The number of compute machines, which are also known as worker machines, to provision.</td>
<td>A positive integer greater than or equal to 2. The default value is 3.</td>
</tr>
<tr>
<td>controlPlane</td>
<td>The configuration for the machines that comprise the control plane.</td>
<td>Array of MachinePool objects. For details, see the following &quot;Machine-pool&quot; table.</td>
</tr>
<tr>
<td>controlPlane.architecture</td>
<td>Determines the instruction set architecture of the machines in the pool. Currently, heterogeneous clusters are not supported, so all pools must specify the same architecture. Valid values are amd64 (the default).</td>
<td>String</td>
</tr>
<tr>
<td>Parameter</td>
<td>Description</td>
<td>Values</td>
</tr>
<tr>
<td>-----------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>-----------------------------</td>
</tr>
<tr>
<td>controlPlane.hypertreading</td>
<td>Whether to enable or disable simultaneous multithreading, or <strong>hypertreading</strong>, on control plane machines. By default, simultaneous multithreading is enabled to increase the performance of your machines’ cores.</td>
<td>Enabled or Disabled</td>
</tr>
<tr>
<td></td>
<td><strong>IMPORTANT</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>If you disable simultaneous multithreading, ensure that your capacity planning accounts for the dramatically decreased machine performance.</td>
<td></td>
</tr>
<tr>
<td>controlPlane.name</td>
<td>Required if you use <code>controlPlane</code>. The name of the machine pool.</td>
<td>master</td>
</tr>
<tr>
<td>controlPlane.platform</td>
<td>Required if you use <code>controlPlane</code>. Use this parameter to specify the cloud provider that hosts the control plane machines. This parameter value must match the <code>compute.platform</code> parameter value.</td>
<td><code>aws, azure, gcp, openstack, ovirt, vsphere, or {}</code></td>
</tr>
<tr>
<td>controlPlane.replicas</td>
<td>The number of control plane machines to provision.</td>
<td>The only supported value is 3, which is the default value.</td>
</tr>
<tr>
<td>credentialsMode</td>
<td>The Cloud Credential Operator (CCO) mode. If no mode is specified, the CCO dynamically tries to determine the capabilities of the provided credentials, with a preference for mint mode on the platforms where multiple modes are supported.</td>
<td><code>Mint, Passthrough, Manual</code>, or an empty string (<code>&quot;&quot;</code>).</td>
</tr>
<tr>
<td></td>
<td><strong>NOTE</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Not all CCO modes are supported for all cloud providers. For more information on CCO modes, see the Cloud Credential Operator entry in the Red Hat Operators reference content.</td>
<td></td>
</tr>
</tbody>
</table>
**fips**
Enable or disable FIPS mode. The default is `false` (disabled). If FIPS mode is enabled, the Red Hat Enterprise Linux CoreOS (RHCOS) machines that OpenShift Container Platform runs on bypass the default Kubernetes cryptography suite and use the cryptography modules that are provided with RHCOS instead.

**IMPORTANT**
The use of FIPS Validated / Modules in Process cryptographic libraries is only supported on OpenShift Container Platform deployments on the `x86_64` architecture.

**NOTE**
If you are using Azure File storage, you cannot enable FIPS mode.

---

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>imageContentSources</td>
<td>Sources and repositories for the release-image content.</td>
<td>Array of objects. Includes a <code>source</code> and, optionally, <code>mirrors</code>, as described in the following rows of this table.</td>
</tr>
<tr>
<td>imageContentSources.source</td>
<td>Required if you use <code>imageContentSources</code>. Specify the repository that users refer to, for example, in image pull specifications.</td>
<td>String</td>
</tr>
<tr>
<td>imageContentSources.mirrors</td>
<td>Specify one or more repositories that may also contain the same images.</td>
<td>Array of strings</td>
</tr>
<tr>
<td>publish</td>
<td>How to publish or expose the user-facing endpoints of your cluster, such as the Kubernetes API, OpenShift routes.</td>
<td><code>Internal</code> or <code>External</code>. To deploy a private cluster, which cannot be accessed from the internet, set <code>publish</code> to <code>Internal</code>. The default value is <code>External</code>.</td>
</tr>
</tbody>
</table>
**sshKey**

The SSH key or keys to authenticate access your cluster machines.

**NOTE**

For production OpenShift Container Platform clusters on which you want to perform installation debugging or disaster recovery, specify an SSH key that your `ssh-agent` process uses.

One or more keys. For example:

```
sshKey:
<key1>
<key2>
<key3>
```

### 1.7.7.4. Optional AWS configuration parameters

Optional AWS configuration parameters are described in the following table:

**Table 1.21. Optional AWS parameters**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>compute.platform.aws.amiID</code></td>
<td>The AWS AMI used to boot compute machines for the cluster. This is required for regions that require a custom RHCOS AMI.</td>
<td>Any published or custom RHCOS AMI that belongs to the set AWS region.</td>
</tr>
<tr>
<td><code>compute.platform.aws.rootVolume.iops</code></td>
<td>The Input/Output Operations Per Second (IOPS) that is reserved for the root volume.</td>
<td>Integer, for example <strong>4000</strong>.</td>
</tr>
<tr>
<td><code>compute.platform.aws.rootVolume.size</code></td>
<td>The size in GiB of the root volume.</td>
<td>Integer, for example <strong>500</strong>.</td>
</tr>
<tr>
<td><code>compute.platform.aws.rootVolume.type</code></td>
<td>The type of the root volume.</td>
<td>Valid AWS EBS volume type, such as <strong>io1</strong>.</td>
</tr>
<tr>
<td><code>compute.platform.aws.type</code></td>
<td>The EC2 instance type for the compute machines.</td>
<td>Valid AWS instance type, such as <strong>c5.9xlarge</strong>.</td>
</tr>
<tr>
<td>Parameter</td>
<td>Description</td>
<td>Values</td>
</tr>
<tr>
<td>-----------------------------------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td><code>compute.platform.aws.zones</code></td>
<td>The availability zones where the installation program creates machines for the compute machine pool. If you provide your own VPC, you must provide a subnet in that availability zone.</td>
<td>A list of valid AWS availability zones, such as <strong>us-east-1c</strong>, in a <strong>YAML sequence</strong>.</td>
</tr>
<tr>
<td><code>compute.aws.region</code></td>
<td>The AWS region that the installation program creates compute resources in.</td>
<td>Any valid <strong>AWS region</strong>, such as <strong>us-east-1</strong>.</td>
</tr>
<tr>
<td><code>controlPlane.platform.aws.amiID</code></td>
<td>The AWS AMI used to boot control plane machines for the cluster. This is required for regions that require a custom RHCOS AMI.</td>
<td>Any published or custom RHCOS AMI that belongs to the set AWS region.</td>
</tr>
<tr>
<td><code>controlPlane.platform.aws.type</code></td>
<td>The EC2 instance type for the control plane machines.</td>
<td>Valid <strong>AWS instance type</strong>, such as <strong>c5.9xlarge</strong>.</td>
</tr>
<tr>
<td><code>controlPlane.platform.aws.zones</code></td>
<td>The availability zones where the installation program creates machines for the control plane machine pool.</td>
<td>A list of valid AWS availability zones, such as <strong>us-east-1c</strong>, in a <strong>YAML sequence</strong>.</td>
</tr>
<tr>
<td><code>controlPlane.aws.region</code></td>
<td>The AWS region that the installation program creates control plane resources in.</td>
<td>Valid <strong>AWS region</strong>, such as <strong>us-east-1</strong>.</td>
</tr>
<tr>
<td><code>platform.aws.amiID</code></td>
<td>The AWS AMI used to boot all machines for the cluster. If set, the AMI must belong to the same region as the cluster. This is required for regions that require a custom RHCOS AMI.</td>
<td>Any published or custom RHCOS AMI that belongs to the set AWS region.</td>
</tr>
<tr>
<td><code>platform.aws.serviceEndpoints.name</code></td>
<td>The AWS service endpoint name. Custom endpoints are only required for cases where alternative AWS endpoints, like FIPS, must be used. Custom API endpoints can be specified for EC2, S3, IAM, Elastic Load Balancing, Tagging, Route 53, and STS AWS services.</td>
<td>Valid <strong>AWS service endpoint</strong> name.</td>
</tr>
</tbody>
</table>
### Parameter Description Values

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>platform.aws.serviceEndpoints.url</td>
<td>The AWS service endpoint URL. The URL must use the https protocol and the host must trust the certificate.</td>
<td>Valid AWS service endpoint URL.</td>
</tr>
<tr>
<td>platform.aws.userTags</td>
<td>A map of keys and values that the installation program adds as tags to all resources that it creates.</td>
<td>Any valid YAML map, such as key value pairs in the &lt;key&gt;: &lt;value&gt; format. For more information about AWS tags, see Tagging Your Amazon EC2 Resources in the AWS documentation.</td>
</tr>
<tr>
<td>platform.aws.subnets</td>
<td>If you provide the VPC instead of allowing the installation program to create the VPC for you, specify the subnet for the cluster to use. The subnet must be part of the same machineNetwork[].cidr ranges that you specify. For a standard cluster, specify a public and a private subnet for each availability zone. For a private cluster, specify a private subnet for each availability zone.</td>
<td>Valid subnet IDs.</td>
</tr>
</tbody>
</table>

#### 1.7.7.2. Sample customized install-config.yaml file for AWS

You can customize the install-config.yaml file to specify more details about your OpenShift Container Platform cluster’s platform or modify the values of the required parameters.

**IMPORTANT**

This sample YAML file is provided for reference only. You must obtain your install-config.yaml file by using the installation program and modify it.

```yaml
apiVersion: v1
baseDomain: example.com
credentialsMode: Mint
controlPlane: 
  name: master
  platform:
    aws:
      zones:
      - us-west-2a
      - us-west-2b
      rootVolume:
        iops: 4000
        size: 500
```
<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>type: io1</td>
<td>type: m5.xlarge</td>
<td>replicas: 3</td>
<td>compute:</td>
<td>- hyperthreading: Enabled</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>size: 500</td>
<td>type: io1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>type: c5.4xlarge</td>
<td>zones:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- us-west-2c</td>
<td>replicas: 3</td>
</tr>
<tr>
<td>metadata:</td>
<td>networking:</td>
<td>clusterNetwork:</td>
<td>- cidr: 10.128.0.0/14</td>
<td>hostPrefix: 23</td>
</tr>
<tr>
<td>name: test-cluster</td>
<td>platform:</td>
<td>aws:</td>
<td>region: us-west-2</td>
<td>userTags:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>adminContact: jdoe</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>costCenter: 7536</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>subnets:</td>
<td>subnet-1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>subnet-2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>subnet-3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>amiID: ami-96c6f8f7</td>
<td>serviceEndpoints:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>url: <a href="https://vpce-id.ec2.us-west-2.vpce.amazonaws.com">https://vpce-id.ec2.us-west-2.vpce.amazonaws.com</a></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>hostedZone: Z3URY6TWQ91KVV</td>
</tr>
<tr>
<td></td>
<td></td>
<td>fips: false</td>
<td>sshKey: ssh-ed25519 AAAA...</td>
<td>publish: Internal</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>pullSecret: '{&quot;auths&quot;: ...}'</td>
</tr>
</tbody>
</table>

1. Required. The installation program prompts you for this value.

2. Optional: Add this parameter to force the Cloud Credential Operator (CCO) to use the specified mode, instead of having the CCO dynamically try to determine the capabilities of the credentials. For details about CCO modes, see the Cloud Credential Operator entry in the Red Hat Operators reference content.
If you do not provide these parameters and values, the installation program provides the default value.

The controlPlane section is a single mapping, but the compute section is a sequence of mappings. To meet the requirements of the different data structures, the first line of the compute section must begin with a hyphen, -, and the first line of the controlPlane section must not. Although both sections currently define a single machine pool, it is possible that future versions of OpenShift Container Platform will support defining multiple compute pools during installation. Only one control plane pool is used.

Whether to enable or disable simultaneous multithreading, or hyperthreading. By default, simultaneous multithreading is enabled to increase the performance of your machines’ cores. You can disable it by setting the parameter value to Disabled. If you disable simultaneous multithreading in some cluster machines, you must disable it in all cluster machines.

IMPORTANT

If you disable simultaneous multithreading, ensure that your capacity planning accounts for the dramatically decreased machine performance. Use larger instance types, such as m4.2xlarge or m5.2xlarge, for your machines if you disable simultaneous multithreading.

To configure faster storage for etcd, especially for larger clusters, set the storage type as io1 and set iops to 2000.

If you provide your own VPC, specify subnets for each availability zone that your cluster uses.

The ID of the AMI used to boot machines for the cluster. If set, the AMI must belong to the same region as the cluster.

The AWS service endpoints. Custom endpoints are required when installing to an unknown AWS region. The endpoint URL must use the https protocol and the host must trust the certificate.

The ID of your existing Route 53 private hosted zone. Providing an existing hosted zone requires that you supply your own VPC and the hosted zone is already associated with the VPC prior to installing your cluster. If undefined, the installation program creates a new hosted zone.

Whether to enable or disable FIPS mode. By default, FIPS mode is not enabled. If FIPS mode is enabled, the Red Hat Enterprise Linux CoreOS (RHCOS) machines that OpenShift Container Platform runs on bypass the default Kubernetes cryptography suite and use the cryptography modules that are provided with RHCOS instead.

IMPORTANT

The use of FIPS Validated / Modules in Process cryptographic libraries is only supported on OpenShift Container Platform deployments on the x86_64 architecture.

You can optionally provide the sshKey value that you use to access the machines in your cluster.

NOTE

For production OpenShift Container Platform clusters on which you want to perform installation debugging or disaster recovery, specify an SSH key that your ssh-agent process uses.
1.7.7.3. Configuring the cluster-wide proxy during installation

Production environments can deny direct access to the Internet and instead have an HTTP or HTTPS proxy available. You can configure a new OpenShift Container Platform cluster to use a proxy by configuring the proxy settings in the `install-config.yaml` file.

Prerequisites

- You have an existing `install-config.yaml` file.
- You reviewed the sites that your cluster requires access to and determined whether any of them need to bypass the proxy. By default, all cluster egress traffic is proxied, including calls to hosting cloud provider APIs. You added sites to the `Proxy` object’s `spec.noProxy` field to bypass the proxy if necessary.

NOTE

The `Proxy` object `status.noProxy` field is populated with the values of the `networking.machineNetwork[].cidr`, `networking.clusterNetwork[].cidr`, and `networking.serviceNetwork[]` fields from your installation configuration.

For installations on Amazon Web Services (AWS), Google Cloud Platform (GCP), Microsoft Azure, and Red Hat OpenStack Platform (RHOSP), the `Proxy` object `status.noProxy` field is also populated with the instance metadata endpoint (169.254.169.254).

- If your cluster is on AWS, you added the `ec2.<region>.amazonaws.com`, `elasticloadbalancing.<region>.amazonaws.com`, and `s3.<region>.amazonaws.com` endpoints to your VPC endpoint. These endpoints are required to complete requests from the nodes to the AWS EC2 API. Because the proxy works on the container level, not the node level, you must route these requests to the AWS EC2 API through the AWS private network. Adding the public IP address of the EC2 API to your allowlist in your proxy server is not sufficient.

Procedure

1. Edit your `install-config.yaml` file and add the proxy settings. For example:

```yaml
apiVersion: v1
baseDomain: my.domain.com
proxy:
  httpProxy: http://<username>:<pswd>@<ip>:<port> ①
  httpsProxy: https://<username>:<pswd>@<ip>:<port> ②
  noProxy: example.com ③
additionalTrustBundle: | ④
    -----BEGIN CERTIFICATE-----
    <MY_TRUSTED_CA_CERT>
    -----END CERTIFICATE-----
```

...
1. A proxy URL to use for creating HTTP connections outside the cluster. The URL scheme must be http.

2. A proxy URL to use for creating HTTPS connections outside the cluster.

3. A comma-separated list of destination domain names, IP addresses, or other network CIDRs to exclude from proxying. Preface a domain with . to match subdomains only. For example, .y.com matches x.y.com, but not y.com. Use * to bypass the proxy for all destinations.

4. If provided, the installation program generates a config map that is named user-ca-bundle in the openshift-config namespace to hold the additional CA certificates. If you provide additionalTrustBundle and at least one proxy setting, the Proxy object is configured to reference the user-ca-bundle config map in the trustedCA field. The Cluster Network Operator then creates a trusted-ca-bundle config map that merges the contents specified for the trustedCA parameter with the RHCOS trust bundle. The additionalTrustBundle field is required unless the proxy’s identity certificate is signed by an authority from the RHCOS trust bundle.

**NOTE**
The installation program does not support the proxy readinessEndpoints field.

2. Save the file and reference it when installing OpenShift Container Platform.

The installation program creates a cluster-wide proxy that is named cluster that uses the proxy settings in the provided install-config.yaml file. If no proxy settings are provided, a cluster Proxy object is still created, but it will have a nil spec.

**NOTE**
Only the Proxy object named cluster is supported, and no additional proxies can be created.

1.7.8. Deploying the cluster

You can install OpenShift Container Platform on a compatible cloud platform.

**IMPORTANT**
You can run the create cluster command of the installation program only once, during initial installation.

**Prerequisites**
- Configure an account with the cloud platform that hosts your cluster.
- Obtain the OpenShift Container Platform installation program and the pull secret for your cluster.

**Procedure**
1. Change to the directory that contains the installation program and initialize the cluster deployment:

```
$ ./openshift-install create cluster --dir <installation_directory> \ 
   --log-level=info
```

1. For `<installation_directory>`, specify the

2. To view different installation details, specify `warn`, `debug`, or `error` instead of `info`.

**NOTE**

If the cloud provider account that you configured on your host does not have sufficient permissions to deploy the cluster, the installation process stops, and the missing permissions are displayed.

When the cluster deployment completes, directions for accessing your cluster, including a link to its web console and credentials for the `kubeadmin` user, display in your terminal.

**Example output**

```
INFO Install complete!
INFO To access the cluster as the system:admin user when using 'oc', run 'export
KUBECONFIG=/home/myuser/install_dir/auth/kubeconfig'
INFO Access the OpenShift web-console here: https://console-openshift-console.apps.mycluster.example.com
INFO Login to the console with user: "kubeadmin", and password: "4vYBz-Ee6gm-ymBZj-Wt5AL"
INFO Time elapsed: 36m22s
```

**NOTE**

The cluster access and credential information also outputs to `<installation_directory>/openshift_install.log` when an installation succeeds.

**IMPORTANT**

The Ignition config files that the installation program generates contain certificates that expire after 24 hours, which are then renewed at that time. If the cluster is shut down before renewing the certificates and the cluster is later restarted after the 24 hours have elapsed, the cluster automatically recovers the expired certificates. The exception is that you must manually approve the pending `node-bootstrapper` certificate signing requests (CSRs) to recover kubelet certificates. See the documentation for *Recovering from expired control plane certificates* for more information.

**IMPORTANT**

You must not delete the installation program or the files that the installation program creates. Both are required to delete the cluster.
1.7.9. Installing the OpenShift CLI by downloading the binary

You can install the OpenShift CLI (oc) in order to interact with OpenShift Container Platform from a command-line interface. You can install oc on Linux, Windows, or macOS.

**IMPORTANT**

If you installed an earlier version of oc, you cannot use it to complete all of the commands in OpenShift Container Platform 4.6. Download and install the new version of oc.

1.7.9.1. Installing the OpenShift CLI on Linux

You can install the OpenShift CLI (oc) binary on Linux by using the following procedure.

**Procedure**


2. Select the appropriate version in the **Version** drop-down menu.

3. Click **Download Now** next to the **OpenShift v4.6 Linux Client** entry and save the file.

4. Unpack the archive:

   $ tar xvzf <file>

5. Place the oc binary in a directory that is on your PATH.

   To check your PATH, execute the following command:

   $ echo $PATH

After you install the OpenShift CLI, it is available using the oc command:

$ oc <command>

1.7.9.2. Installing the OpenShift CLI on Windows

You can install the OpenShift CLI (oc) binary on Windows by using the following procedure.

**Procedure**


2. Select the appropriate version in the **Version** drop-down menu.

3. Click **Download Now** next to the **OpenShift v4.6 Windows Client** entry and save the file.

4. Unzip the archive with a ZIP program.

5. Move the oc binary to a directory that is on your PATH.

   To check your PATH, open the command prompt and execute the following command:

   $ echo $PATH
After you install the OpenShift CLI, it is available using the `oc` command:

```bash
C:\> oc <command>
```

### 1.7.9.3. Installing the OpenShift CLI on macOS

You can install the OpenShift CLI (`oc`) binary on macOS by using the following procedure.

**Procedure**

2. Select the appropriate version in the **Version** drop-down menu.
3. Click **Download Now** next to the **OpenShift v4.6 MacOSX Client** entry and save the file.
4. Unpack and unzip the archive.
5. Move the `oc` binary to a directory on your PATH.
   
   To check your **PATH**, open a terminal and execute the following command:

   ```bash
   $ echo $PATH
   $ oc <command>
   ```

After you install the OpenShift CLI, it is available using the `oc` command:

```bash
$ oc <command>
```

### 1.7.10. Logging in to the cluster by using the CLI

You can log in to your cluster as a default system user by exporting the cluster `kubeconfig` file. The `kubeconfig` file contains information about the cluster that is used by the CLI to connect a client to the correct cluster and API server. The file is specific to a cluster and is created during OpenShift Container Platform installation.

**Prerequisites**

- You deployed an OpenShift Container Platform cluster.
- You installed the `oc` CLI.

**Procedure**

1. Export the `kubeadmin` credentials:

   ```bash
   $ export KUBECONFIG=<installation_directory>/auth/kubeconfig
   ```

   1 For `<installation_directory>`, specify the path to the directory that you stored the installation files in.
2. Verify you can run `oc` commands successfully using the exported configuration:

```bash
$ oc whoami
```

**Example output**

```
system:admin
```

### 1.7.11. Logging in to the cluster by using the web console

The `kubeadmin` user exists by default after an OpenShift Container Platform installation. You can log in to your cluster as the `kubeadmin` user by using the OpenShift Container Platform web console.

**Prerequisites**

- You have access to the installation host.
- You completed a cluster installation and all cluster Operators are available.

**Procedure**

1. Obtain the password for the `kubeadmin` user from the `kubeadmin-password` file on the installation host:

```bash
$ cat <installation_directory>/auth/kubeadmin-password
```

**NOTE**

Alternatively, you can obtain the `kubeadmin` password from the `<installation_directory>/openshift_install.log` log file on the installation host.

2. List the OpenShift Container Platform web console route:

```bash
$ oc get routes -n openshift-console | grep 'console-openshift'
```

**NOTE**

Alternatively, you can obtain the OpenShift Container Platform route from the `<installation_directory>/openshift_install.log` log file on the installation host.

**Example output**

```
console   console-openshift-console.apps.<cluster_name>.<base_domain>   console
https   reencrypt/Redirect   None
```

3. Navigate to the route detailed in the output of the preceding command in a web browser and log in as the `kubeadmin` user.

**Additional resources**

- See [Accessing the web console](#) for more details about accessing and understanding the $ `oc whoami`

  ```bash
  system:admin
  ```

OpenShift Container Platform 4.6 Installing on AWS

144
• See Accessing the web console for more details about accessing and understanding the OpenShift Container Platform web console.

1.7.12. Telemetry access for OpenShift Container Platform

In OpenShift Container Platform 4.6, the Telemetry service, which runs by default to provide metrics about cluster health and the success of updates, requires Internet access. If your cluster is connected to the Internet, Telemetry runs automatically, and your cluster is registered to the Red Hat OpenShift Cluster Manager (OCM).

After you confirm that your Red Hat OpenShift Cluster Manager inventory is correct, either maintained automatically by Telemetry or manually by using OCM, use subscription watch to track your OpenShift Container Platform subscriptions at the account or multi-cluster level.

Additional resources
• See About remote health monitoring for more information about the Telemetry service.

1.7.13. Next steps

• Validating an installation.
• Customize your cluster.
• If necessary, you can opt out of remote health reporting.
• If necessary, you can remove cloud provider credentials.

1.8. INSTALLING A CLUSTER ON AWS INTO A GOVERNMENT REGION

In OpenShift Container Platform version 4.6, you can install a cluster on Amazon Web Services (AWS) into a government region. To configure the government region, modify parameters in the install-config.yaml file before you install the cluster.

1.8.1. Prerequisites

• Review details about the OpenShift Container Platform installation and update processes.
• Configure an AWS account to host the cluster.

IMPORTANT

If you have an AWS profile stored on your computer, it must not use a temporary session token that you generated while using a multi-factor authentication device. The cluster continues to use your current AWS credentials to create AWS resources for the entire life of the cluster, so you must use long-lived credentials. To generate appropriate keys, see Managing Access Keys for IAM Users in the AWS documentation. You can supply the keys when you run the installation program.

• If you use a firewall, you must configure it to allow the sites that your cluster requires access to.
• If you do not allow the system to manage identity and access management (IAM), then a cluster administrator can manually create and maintain IAM credentials. Manual mode can also be used in environments where the cloud IAM APIs are not reachable.

1.8.2. AWS government regions

OpenShift Container Platform supports deploying a cluster to AWS GovCloud (US) regions. AWS GovCloud is specifically designed for US government agencies at the federal, state, and local level, as well as contractors, educational institutions, and other US customers that must run sensitive workloads in the cloud.

These regions do not have published Red Hat Enterprise Linux CoreOS (RHCOS) Amazon Machine Images (AMI) to select, so you must upload a custom AMI that belongs to that region.

The following AWS GovCloud partitions are supported:

• us-gov-west-1
• us-gov-east-1

The AWS GovCloud region and custom AMI must be manually configured in the install-config.yaml file since RHCOS AMIs are not provided by Red Hat for those regions.

1.8.3. Private clusters

You can deploy a private OpenShift Container Platform cluster that does not expose external endpoints. Private clusters are accessible from only an internal network and are not visible to the Internet.

**NOTE**

Public zones are not supported in Route 53 in AWS GovCloud. Therefore, clusters must be private if they are deployed to an AWS government region.

By default, OpenShift Container Platform is provisioned to use publicly-accessible DNS and endpoints. A private cluster sets the DNS, Ingress Controller, and API server to private when you deploy your cluster. This means that the cluster resources are only accessible from your internal network and are not visible to the internet.

To deploy a private cluster, you must use existing networking that meets your requirements. Your cluster resources might be shared between other clusters on the network.

Additionally, you must deploy a private cluster from a machine that has access the API services for the cloud you provision to, the hosts on the network that you provision, and to the internet to obtain installation media. You can use any machine that meets these access requirements and follows your company’s guidelines. For example, this machine can be a bastion host on your cloud network or a machine that has access to the network through a VPN.

1.8.3.1. Private clusters in AWS

To create a private cluster on Amazon Web Services (AWS), you must provide an existing private VPC and subnets to host the cluster. The installation program must also be able to resolve the DNS records that the cluster requires. The installation program configures the Ingress Operator and API server for access from only the private network.
The cluster still requires access to Internet to access the AWS APIs.

The following items are not required or created when you install a private cluster:

- Public subnets
- Public load balancers, which support public ingress
- A public Route 53 zone that matches the baseDomain for the cluster

The installation program does use the baseDomain that you specify to create a private Route 53 zone and the required records for the cluster. The cluster is configured so that the Operators do not create public records for the cluster and all cluster machines are placed in the private subnets that you specify.

### 1.8.3.1.1. Limitations

The ability to add public functionality to a private cluster is limited.

- You cannot make the Kubernetes API endpoints public after installation without taking additional actions, including creating public subnets in the VPC for each availability zone in use, creating a public load balancer, and configuring the control plane security groups to allow traffic from Internet on 6443 (Kubernetes API port).

- If you use a public Service type load balancer, you must tag a public subnet in each availability zone with kubernetes.io/cluster/<cluster-infra-id>: shared so that AWS can use them to create public load balancers.

### 1.8.4. About using a custom VPC

In OpenShift Container Platform 4.6, you can deploy a cluster into existing subnets in an existing Amazon Virtual Private Cloud (VPC) in Amazon Web Services (AWS). By deploying OpenShift Container Platform into an existing AWS VPC, you might be able to avoid limit constraints in new accounts or more easily abide by the operational constraints that your company's guidelines set. If you cannot obtain the infrastructure creation permissions that are required to create the VPC yourself, use this installation option.

Because the installation program cannot know what other components are also in your existing subnets, it cannot choose subnet CIDRs and so forth on your behalf. You must configure networking for the subnets that you install your cluster to yourself.

### 1.8.4.1. Requirements for using your VPC

The installation program no longer creates the following components:

- Internet gateways
- NAT gateways
- Subnets
- Route tables
- VPCs
- VPC DHCP options
VPC endpoints

If you use a custom VPC, you must correctly configure it and its subnets for the installation program and the cluster to use. See Amazon VPC console wizard configurations and Work with VPCs and subnets in the AWS documentation for more information on creating and managing an AWS VPC.

The installation program cannot:

- Subdivide network ranges for the cluster to use.
- Set route tables for the subnets.
- Set VPC options like DHCP.

You must complete these tasks before you install the cluster. See VPC networking components and Route tables for your VPC for more information on configuring networking in an AWS VPC.

Your VPC must meet the following characteristics:

- The VPC must not use the kubernetes.io/cluster/.*: owned tag. The installation program modifies your subnets to add the kubernetes.io/cluster/.*: shared tag, so your subnets must have at least one free tag slot available for it. See Tag Restrictions in the AWS documentation to confirm that the installation program can add a tag to each subnet that you specify.

- You must enable the enableDnsSupport and enableDnsHostnames attributes in your VPC, so that the cluster can use the Route 53 zones that are attached to the VPC to resolve cluster’s internal DNS records. See DNS Support in Your VPC in the AWS documentation.
  If you prefer to use your own Route 53 hosted private zone, you must associate the existing hosted zone with your VPC prior to installing a cluster. You can define your hosted zone using the platform.aws.hostedZone field in the install-config.yaml file.

- If you use a cluster with public access, you must create a public and a private subnet for each availability zone that your cluster uses. Each availability zone can contain no more than one public and one private subnet.

If you are working in a disconnected environment, you are unable to reach the public IP addresses for EC2 and ELB endpoints. To resolve this, you must create a VPC endpoint and attach it to the subnet that the clusters are using. The endpoints should be named as follows:

- ec2.<region>.amazonaws.com
- elasticloadbalancing.<region>.amazonaws.com
- s3.<region>.amazonaws.com

Required VPC components

You must provide a suitable VPC and subnets that allow communication to your machines.

<table>
<thead>
<tr>
<th>Component</th>
<th>AWS type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Component</td>
<td>AWS type</td>
<td>Description</td>
</tr>
<tr>
<td>--------------------</td>
<td>-----------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>VPC</td>
<td>• AWS::EC2::VPC</td>
<td>You must provide a public VPC for the cluster to use. The VPC uses an endpoint that references the route tables for each subnet to improve communication with the registry that is hosted in S3.</td>
</tr>
<tr>
<td></td>
<td>• AWS::EC2::VPCEndpoint</td>
<td></td>
</tr>
<tr>
<td>Public subnets</td>
<td>• AWS::EC2::Subnet</td>
<td>Your VPC must have public subnets for between 1 and 3 availability zones and associate them with appropriate Ingress rules.</td>
</tr>
<tr>
<td></td>
<td>• AWS::EC2::SubnetNetworkAclAssociation</td>
<td></td>
</tr>
<tr>
<td>Internet gateway</td>
<td>• AWS::EC2::InternetGateway</td>
<td>You must have a public Internet gateway, with public routes, attached to the VPC. In the provided templates, each public subnet has a NAT gateway with an EIP address. These NAT gateways allow cluster resources, like private subnet instances, to reach the Internet and are not required for some restricted network or proxy scenarios.</td>
</tr>
<tr>
<td></td>
<td>• AWS::EC2::VPCGatewayAttachment</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• AWS::EC2::RouteTable</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• AWS::EC2::Route</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• AWS::EC2::SubnetRouteTableAssociation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• AWS::EC2::NatGateway</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• AWS::EC2::EIP</td>
<td></td>
</tr>
<tr>
<td>Network access</td>
<td>• AWS::EC2::NetworkAcl</td>
<td>You must allow the VPC to access the following ports:</td>
</tr>
<tr>
<td>control</td>
<td>• AWS::EC2::NetworkAclEntry</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Port</td>
<td>Reason</td>
</tr>
<tr>
<td></td>
<td>80</td>
<td>Inbound HTTP traffic</td>
</tr>
<tr>
<td></td>
<td>443</td>
<td>Inbound HTTPS traffic</td>
</tr>
<tr>
<td></td>
<td>22</td>
<td>Inbound SSH traffic</td>
</tr>
<tr>
<td></td>
<td>1024 - 65535</td>
<td>Inbound ephemeral traffic</td>
</tr>
<tr>
<td></td>
<td>0 - 65535</td>
<td>Outbound ephemeral traffic</td>
</tr>
</tbody>
</table>
Your VPC can have private subnets. The provided CloudFormation templates can create private subnets for between 1 and 3 availability zones. If you use private subnets, you must provide appropriate routes and tables for them.

### 1.8.4.2. VPC validation

To ensure that the subnets that you provide are suitable, the installation program confirms the following data:

- All the subnets that you specify exist.
- You provide private subnets.
- The subnet CIDRs belong to the machine CIDR that you specified.
- You provide subnets for each availability zone. Each availability zone contains no more than one public and one private subnet. If you use a private cluster, provide only a private subnet for each availability zone. Otherwise, provide exactly one public and private subnet for each availability zone.
- You provide a public subnet for each private subnet availability zone. Machines are not provisioned in availability zones that you do not provide private subnets for.

If you destroy a cluster that uses an existing VPC, the VPC is not deleted. When you remove the OpenShift Container Platform cluster from a VPC, the `kubernetes.io/cluster/.*: shared` tag is removed from the subnets that it used.

### 1.8.4.3. Division of permissions

Starting with OpenShift Container Platform 4.3, you do not need all of the permissions that are required for an installation program-provisioned infrastructure cluster to deploy a cluster. This change mimics the division of permissions that you might have at your company: some individuals can create different resource in your clouds than others. For example, you might be able to create application-specific items, like instances, buckets, and load balancers, but not networking-related components such as VPCs, subnets, or ingress rules.

The AWS credentials that you use when you create your cluster do not need the networking permissions that are required to make VPCs and core networking components within the VPC, such as subnets, routing tables, Internet gateways, NAT, and VPN. You still need permission to make the application resources that the machines within the cluster require, such as ELBs, security groups, S3 buckets, and nodes.

### 1.8.4.4. Isolation between clusters

If you deploy OpenShift Container Platform to an existing network, the isolation of cluster services is reduced in the following ways:
- You can install multiple OpenShift Container Platform clusters in the same VPC.
- ICMP ingress is allowed from the entire network.
- TCP 22 ingress (SSH) is allowed to the entire network.
- Control plane TCP 6443 ingress (Kubernetes API) is allowed to the entire network.
- Control plane TCP 22623 ingress (MCS) is allowed to the entire network.

1.8.5. Internet access for OpenShift Container Platform

In OpenShift Container Platform 4.6, you require access to the Internet to install your cluster.

You must have Internet access to:

- Access the Red Hat OpenShift Cluster Manager page to download the installation program and perform subscription management. If the cluster has Internet access and you do not disable Telemetry, that service automatically entitles your cluster.
- Access Quay.io to obtain the packages that are required to install your cluster.
- Obtain the packages that are required to perform cluster updates.

**IMPORTANT**

If your cluster cannot have direct Internet access, you can perform a restricted network installation on some types of infrastructure that you provision. During that process, you download the content that is required and use it to populate a mirror registry with the packages that you need to install a cluster and generate the installation program. With some installation types, the environment that you install your cluster in will not require Internet access. Before you update the cluster, you update the content of the mirror registry.

1.8.6. Generating an SSH private key and adding it to the agent

If you want to perform installation debugging or disaster recovery on your cluster, you must provide an SSH key to both your ssh-agent and the installation program. You can use this key to access the bootstrap machine in a public cluster to troubleshoot installation issues.

**NOTE**

In a production environment, you require disaster recovery and debugging.

You can use this key to SSH into the master nodes as the user core. When you deploy the cluster, the key is added to the core user’s ~/.ssh/authorized_keys list.

**NOTE**

You must use a local key, not one that you configured with platform-specific approaches such as AWS key pairs.

**Procedure**
1. If you do not have an SSH key that is configured for password-less authentication on your computer, create one. For example, on a computer that uses a Linux operating system, run the following command:

```
$ ssh-keygen -t ed25519 -N " \n   -f <path>/<file_name> 1
```

Specify the path and file name, such as `~/.ssh/id_rsa`, of the new SSH key. If you have an existing key pair, ensure your public key is in the your `~/.ssh` directory.

Running this command generates an SSH key that does not require a password in the location that you specified.

**NOTE**

If you plan to install an OpenShift Container Platform cluster that uses FIPS Validated / Modules in Process cryptographic libraries on the `x86_64` architecture, do not create a key that uses the `ed25519` algorithm. Instead, create a key that uses the `rsa` or `ecdsa` algorithm.

2. Start the `ssh-agent` process as a background task:

```
$ eval "$(ssh-agent -s)"
```

**Example output**

```
Agent pid 31874
```

**NOTE**

If your cluster is in FIPS mode, only use FIPS-compliant algorithms to generate the SSH key. The key must be either RSA or ECDSA.

3. Add your SSH private key to the `ssh-agent`:

```
$ ssh-add <path>/<file_name> 1
```

**Example output**

```
Identity added: /home/<you>/<path>/<file_name> (<computer_name>)
```

Specify the path and file name for your SSH private key, such as `~/.ssh/id_rsa`

**Next steps**

- When you install OpenShift Container Platform, provide the SSH public key to the installation program.

**1.8.7. Obtaining the installation program**
Before you install OpenShift Container Platform, download the installation file on a local computer.

**Prerequisites**

- You have a computer that runs Linux or macOS, with 500 MB of local disk space

**Procedure**

1. Access the Infrastructure Provider page on the Red Hat OpenShift Cluster Manager site. If you have a Red Hat account, log in with your credentials. If you do not, create an account.

2. Select your infrastructure provider.

3. Navigate to the page for your installation type, download the installation program for your operating system, and place the file in the directory where you will store the installation configuration files.

   **IMPORTANT**

   The installation program creates several files on the computer that you use to install your cluster. You must keep the installation program and the files that the installation program creates after you finish installing the cluster. Both files are required to delete the cluster.

   **IMPORTANT**

   Deleting the files created by the installation program does not remove your cluster, even if the cluster failed during installation. To remove your cluster, complete the OpenShift Container Platform uninstallation procedures for your specific cloud provider.

4. Extract the installation program. For example, on a computer that uses a Linux operating system, run the following command:

   ```
   $ tar xvf openshift-install-linux.tar.gz
   ```

5. From the Pull Secret page on the Red Hat OpenShift Cluster Manager site, download your installation pull secret as a `.txt` file. This pull secret allows you to authenticate with the services that are provided by the included authorities, including Quay.io, which serves the container images for OpenShift Container Platform components.

### 1.8.8. Manually creating the installation configuration file

When installing OpenShift Container Platform on Amazon Web Services (AWS) into a region requiring a custom Red Hat Enterprise Linux CoreOS (RHCOS) AMI, you must manually generate your installation configuration file.

**Prerequisites**

- Obtain the OpenShift Container Platform installation program and the access token for your cluster.

**Procedure**
1. Create an installation directory to store your required installation assets in:

```
$ mkdir <installation_directory>
```

**IMPORTANT**

You must create a directory. Some installation assets, like bootstrap X.509 certificates have short expiration intervals, so you must not reuse an installation directory. If you want to reuse individual files from another cluster installation, you can copy them into your directory. However, the file names for the installation assets might change between releases. Use caution when copying installation files from an earlier OpenShift Container Platform version.

2. Customize the following `install-config.yaml` file template and save it in the `<installation_directory>`.

**NOTE**

You must name this configuration file `install-config.yaml`.

3. Back up the `install-config.yaml` file so that you can use it to install multiple clusters.

**IMPORTANT**

The `install-config.yaml` file is consumed during the next step of the installation process. You must back it up now.

### 1.8.8.1. Installation configuration parameters

Before you deploy an OpenShift Container Platform cluster, you provide parameter values to describe your account on the cloud platform that hosts your cluster and optionally customize your cluster’s platform. When you create the `install-config.yaml` installation configuration file, you provide values for the required parameters through the command line. If you customize your cluster, you can modify the `install-config.yaml` file to provide more details about the platform.

**NOTE**

After installation, you cannot modify these parameters in the `install-config.yaml` file.

**IMPORTANT**

The `openshift-install` command does not validate field names for parameters. If an incorrect name is specified, the related file or object is not created, and no error is reported. Ensure that the field names for any parameters that are specified are correct.

#### 1.8.8.1.1. Required configuration parameters

Required installation configuration parameters are described in the following table:

Table 1.22. Required parameters
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>apiVersion</td>
<td>The API version for the <code>install-config.yaml</code> content. The current version is <strong>v1</strong>. The installer may also support older API versions.</td>
<td>String</td>
</tr>
<tr>
<td>baseDomain</td>
<td>The base domain of your cloud provider. The base domain is used to create routes to your OpenShift Container Platform cluster components. The full DNS name for your cluster is a combination of the <code>baseDomain</code> and <code>metadata.name</code> parameter values that uses the <code>&lt;metadata.name&gt;.[baseDomain&gt;</code> format.</td>
<td>A fully-qualified domain or subdomain name, such as <code>example.com</code>.</td>
</tr>
<tr>
<td>metadata</td>
<td>Kubernetes resource <code>ObjectMeta</code>, from which only the <code>name</code> parameter is consumed.</td>
<td>Object</td>
</tr>
<tr>
<td>metadata.name</td>
<td>The name of the cluster. DNS records for the cluster are all subdomains of <code>{{.metadata.name}}.{{.baseDomain}}</code>.</td>
<td>String of lowercase letters, hyphens (-), and periods (.), such as <code>dev</code>.</td>
</tr>
<tr>
<td>platform</td>
<td>The configuration for the specific platform upon which to perform the installation: <code>aws</code>, <code>baremetal</code>, <code>azure</code>, <code>openstack</code>, <code>ovirt</code>, <code>vsphere</code>. For additional information about <code>platform.&lt;platform&gt;</code> parameters, consult the following table for your specific platform.</td>
<td>Object</td>
</tr>
</tbody>
</table>
1.8.8.1.2. Network configuration parameters

You can customize your installation configuration based on the requirements of your existing network infrastructure. For example, you can expand the IP address block for the cluster network or provide different IP address blocks than the defaults.

Only IPv4 addresses are supported.

Table 1.23. Network parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>networking</td>
<td>The configuration for the cluster network.</td>
<td>Object</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>networking.network</td>
<td>The cluster network provider Container Network Interface (CNI) plug-in to install.</td>
<td>Either OpenShiftSDN or OVNKubernetes. The default value is OpenShiftSDN.</td>
</tr>
<tr>
<td>networking.clusterNetwork</td>
<td>The IP address blocks for pods.</td>
<td>An array of objects. For example: networking: clusterNetwork: - cidr: 10.128.0.0/14</td>
</tr>
</tbody>
</table>

### pullSecret

Get a pull secret from https://console.redhat.com/openshift/install/pull-secret to authenticate downloading container images for OpenShift Container Platform components from services such as Quay.io.

```json
{
  "auths":{
    "cloud.openshift.com":{
      "auth":"b3Blb=",
      "email":"you@example.com"
    },
    "quay.io":{
      "auth":"b3Blb=",
      "email":"you@example.com"
    }
  }
}
```
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>networking.clusterNetwork.cidr</td>
<td>Required if you use networking.clusterNetwork. An IP address block. An IPv4 network.</td>
<td>An IP address block in Classless Inter-Domain Routing (CIDR) notation. The prefix length for an IPv4 block is between 0 and 32.</td>
</tr>
<tr>
<td>networking.clusterNetwork.hostPrefix</td>
<td>The subnet prefix length to assign to each individual node. For example, if hostPrefix is set to 23 then each node is assigned a /23 subnet out of the given cidr. A hostPrefix value of 23 provides 510 (2^{(32 - 23)} - 2) pod IP addresses.</td>
<td>A subnet prefix. The default value is 23.</td>
</tr>
<tr>
<td>networking.serviceNetwork</td>
<td>The IP address block for services. The default value is (172.30.0.0/16). The OpenShift SDN and OVN-Kubernetes network providers support only a single IP address block for the service network.</td>
<td>An array with an IP address block in CIDR format. For example:</td>
</tr>
<tr>
<td>networking.machineNetwork</td>
<td>The IP address blocks for machines. If you specify multiple IP address blocks, the blocks must not overlap.</td>
<td>An array of objects. For example:</td>
</tr>
<tr>
<td>networking.machineNetwork.cidr</td>
<td>Required if you use networking.machineNetwork. An IP address block. The default value is (10.0.0.0/16) for all platforms other than libvirt. For libvirt, the default value is (192.168.126.0/24).</td>
<td>An IP network block in CIDR notation. For example, (10.0.0.0/16).</td>
</tr>
</tbody>
</table>

1.8.8.1.3. Optional configuration parameters

Optional installation configuration parameters are described in the following table:

Table 1.24. Optional parameters
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>additionalTrustBundle</code></td>
<td>A PEM-encoded X.509 certificate bundle that is added to the nodes' trusted certificate store. This trust bundle may also be used when a proxy has been configured.</td>
<td>String</td>
</tr>
<tr>
<td><code>compute</code></td>
<td>The configuration for the machines that comprise the compute nodes.</td>
<td>Array of machine-pool objects. For details, see the following &quot;Machine-pool&quot; table.</td>
</tr>
<tr>
<td><code>compute.architecture</code></td>
<td>Determines the instruction set architecture of the machines in the pool. Currently, heterogeneous clusters are not supported, so all pools must specify the same architecture. Valid values are <code>amd64</code> (the default).</td>
<td>String</td>
</tr>
<tr>
<td><code>compute.hyperthreading</code></td>
<td>Whether to enable or disable simultaneous multithreading, or hyperthreading, on compute machines. By default, simultaneous multithreading is enabled to increase the performance of your machines' cores.</td>
<td>Enabled or Disabled</td>
</tr>
<tr>
<td></td>
<td>IMPORTANT</td>
<td></td>
</tr>
<tr>
<td></td>
<td>If you disable simultaneous multithreading, ensure that your capacity planning accounts for the dramatically decreased machine performance.</td>
<td></td>
</tr>
<tr>
<td><code>compute.name</code></td>
<td>Required if you use <code>compute</code>. The name of the machine pool.</td>
<td><code>worker</code></td>
</tr>
<tr>
<td><code>compute.platform</code></td>
<td>Required if you use <code>compute</code>. Use this parameter to specify the cloud provider to host the worker machines. This parameter value must match the <code>controlPlane.platform</code> parameter value.</td>
<td><code>aws, azure, gcp, openstack, ovirt, vsphere, or {}</code></td>
</tr>
<tr>
<td><code>compute.replicas</code></td>
<td>The number of compute machines, which are also known as worker machines, to provision.</td>
<td>A positive integer greater than or equal to 2. The default value is 3.</td>
</tr>
<tr>
<td>Parameter</td>
<td>Description</td>
<td>Values</td>
</tr>
<tr>
<td>---------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------</td>
<td>-----------------------------</td>
</tr>
<tr>
<td>controlPlane</td>
<td>The configuration for the machines that comprise the control plane.</td>
<td>Array of <strong>MachinePool</strong> objects. For details, see the following &quot;Machine-pool&quot; table.</td>
</tr>
<tr>
<td>controlPlane.architecture</td>
<td>Determines the instruction set architecture of the machines in the pool. Currently, heterogeneous clusters are not supported, so all pools must specify the same architecture. Valid values are <strong>amd64</strong> (the default).</td>
<td>String</td>
</tr>
<tr>
<td>controlPlane.hypertreading</td>
<td>Whether to enable or disable simultaneous multithreading, or <strong>hyperthreading</strong>, on control plane machines. By default, simultaneous multithreading is enabled to increase the performance of your machines' cores.</td>
<td><strong>Enabled</strong> or <strong>Disabled</strong></td>
</tr>
<tr>
<td>controlPlane.name</td>
<td>Required if you use controlPlane. The name of the machine pool.</td>
<td>master</td>
</tr>
<tr>
<td>controlPlane.platform</td>
<td>Required if you use controlPlane. Use this parameter to specify the cloud provider that hosts the control plane machines. This parameter value must match the <strong>compute.platform</strong> parameter value.</td>
<td>aws, azure, gcp, openstack, ovirt, vsphere, or {}</td>
</tr>
<tr>
<td>controlPlane.replicas</td>
<td>The number of control plane machines to provision.</td>
<td>The only supported value is 3, which is the default value.</td>
</tr>
<tr>
<td>Parameter</td>
<td>Description</td>
<td>Values</td>
</tr>
<tr>
<td>---------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>credentialsMode</td>
<td>The Cloud Credential Operator (CCO) mode. If no mode is specified, the CCO dynamically tries to determine the capabilities of the provided credentials, with a preference for mint mode on the platforms where multiple modes are supported.</td>
<td>Mint, Passthrough, Manual, or an empty string (&quot;&quot;).</td>
</tr>
</tbody>
</table>

**NOTE**

Not all CCO modes are supported for all cloud providers. For more information on CCO modes, see the Cloud Credential Operator entry in the Red Hat Operators reference content.
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>fips</td>
<td>Enable or disable FIPS mode. The default is false (disabled). If FIPS mode is enabled, the Red Hat Enterprise Linux CoreOS (RHCOS) machines that OpenShift Container Platform runs on bypass the default Kubernetes cryptography suite and use the cryptography modules that are provided with RHCOS instead.</td>
<td>false or true</td>
</tr>
<tr>
<td>IMPORTANT</td>
<td>The use of FIPS Validated / Modules in Process cryptographic libraries is only supported on OpenShift Container Platform deployments on the x86_64 architecture.</td>
<td></td>
</tr>
<tr>
<td>NOTE</td>
<td>If you are using Azure File storage, you cannot enable FIPS mode.</td>
<td></td>
</tr>
<tr>
<td>imageContentSources</td>
<td>Sources and repositories for the release-image content.</td>
<td>Array of objects. Includes a source and, optionally, mirrors, as described in the following rows of this table.</td>
</tr>
<tr>
<td>imageContentSources.source</td>
<td>Required if you use imageContentSources. Specify the repository that users refer to, for example, in image pull specifications.</td>
<td>String</td>
</tr>
<tr>
<td>imageContentSources.mirrors</td>
<td>Specify one or more repositories that may also contain the same images.</td>
<td>Array of strings</td>
</tr>
<tr>
<td>publish</td>
<td>How to publish or expose the user-facing endpoints of your cluster, such as the Kubernetes API, OpenShift routes.</td>
<td>Internal or External. To deploy a private cluster, which cannot be accessed from the internet, set publish to Internal. The default value is External.</td>
</tr>
</tbody>
</table>

**Parameter**

- **fips**: Enable or disable FIPS mode. The default is false (disabled). If FIPS mode is enabled, the Red Hat Enterprise Linux CoreOS (RHCOS) machines that OpenShift Container Platform runs on bypass the default Kubernetes cryptography suite and use the cryptography modules that are provided with RHCOS instead.

**Important Note**

The use of FIPS Validated / Modules in Process cryptographic libraries is only supported on OpenShift Container Platform deployments on the x86_64 architecture.

**Note**

If you are using Azure File storage, you cannot enable FIPS mode.

**imageContentSources**

- **imageContentSources.source**: Required if you use imageContentSources. Specify the repository that users refer to, for example, in image pull specifications.

**imageContentSources.mirrors**: Specify one or more repositories that may also contain the same images.

**publish**

- **Internal** or **External**. To deploy a private cluster, which cannot be accessed from the internet, set publish to Internal. The default value is External.
The SSH key or keys to authenticate access your cluster machines.

**NOTE**
For production OpenShift Container Platform clusters on which you want to perform installation debugging or disaster recovery, specify an SSH key that your `ssh-agent` process uses.

One or more keys. For example:

```
sshKey:
  <key1>
  <key2>
  <key3>
```

### 1.8.8.1.4. Optional AWS configuration parameters

Optional AWS configuration parameters are described in the following table:

**Table 1.25. Optional AWS parameters**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>compute.platform.aws.amiID</code></td>
<td>The AWS AMI used to boot compute machines for the cluster. This is required for regions that require a custom RHCOS AMI.</td>
<td>Any published or custom RHCOS AMI that belongs to the set AWS region.</td>
</tr>
<tr>
<td><code>compute.platform.aws.rootVolume.iops</code></td>
<td>The Input/Output Operations Per Second (IOPS) that is reserved for the root volume.</td>
<td>Integer, for example <strong>4000</strong>.</td>
</tr>
<tr>
<td><code>compute.platform.aws.rootVolume.size</code></td>
<td>The size in GiB of the root volume.</td>
<td>Integer, for example <strong>500</strong>.</td>
</tr>
<tr>
<td><code>compute.platform.aws.rootVolume.type</code></td>
<td>The type of the root volume.</td>
<td>Valid <strong>AWS EBS volume type</strong>, such as <strong>io1</strong>.</td>
</tr>
<tr>
<td><code>compute.platform.aws.type</code></td>
<td>The EC2 instance type for the compute machines.</td>
<td>Valid <strong>AWS instance type</strong>, such as <strong>c5.9xlarge</strong>.</td>
</tr>
<tr>
<td>Parameter</td>
<td>Description</td>
<td>Values</td>
</tr>
<tr>
<td>------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td><code>compute.platform.aws.zones</code></td>
<td>The availability zones where the installation program creates machines for the compute machine pool. If you provide your own VPC, you must provide a subnet in that availability zone.</td>
<td>A list of valid AWS availability zones, such as <code>us-east-1c</code>, in a YAML sequence.</td>
</tr>
<tr>
<td><code>compute.aws.region</code></td>
<td>The AWS region that the installation program creates compute resources in.</td>
<td>Any valid AWS region, such as <code>us-east-1</code>.</td>
</tr>
<tr>
<td><code>controlPlane.platform.aws.amiID</code></td>
<td>The AWS AMI used to boot control plane machines for the cluster. This is required for regions that require a custom RH COS AMI.</td>
<td>Any published or custom RH COS AMI that belongs to the set AWS region.</td>
</tr>
<tr>
<td><code>controlPlane.platform.aws.type</code></td>
<td>The EC2 instance type for the control plane machines.</td>
<td>Valid AWS instance type, such as <code>c5.9xlarge</code>.</td>
</tr>
<tr>
<td><code>controlPlane.platform.aws.zones</code></td>
<td>The availability zones where the installation program creates machines for the control plane machine pool.</td>
<td>A list of valid AWS availability zones, such as <code>us-east-1c</code>, in a YAML sequence.</td>
</tr>
<tr>
<td><code>controlPlane.aws.region</code></td>
<td>The AWS region that the installation program creates control plane resources in.</td>
<td>Valid AWS region, such as <code>us-east-1</code>.</td>
</tr>
<tr>
<td><code>platform.aws.amiID</code></td>
<td>The AWS AMI used to boot all machines for the cluster. If set, the AMI must belong to the same region as the cluster. This is required for regions that require a custom RH COS AMI.</td>
<td>Any published or custom RH COS AMI that belongs to the set AWS region.</td>
</tr>
<tr>
<td><code>platform.aws.serviceEndpoints.name</code></td>
<td>The AWS service endpoint name. Custom endpoints are only required for cases where alternative AWS endpoints, like FIPS, must be used. Custom API endpoints can be specified for EC2, S3, IAM, Elastic Load Balancing, Tagging, Route 53, and STS AWS services.</td>
<td>Valid AWS service endpoint name.</td>
</tr>
</tbody>
</table>
The AWS service endpoint URL. The URL must use the https protocol and the host must trust the certificate.

A map of keys and values that the installation program adds as tags to all resources that it creates.

If you provide the VPC instead of allowing the installation program to create the VPC for you, specify the subnet for the cluster to use. The subnet must be part of the same machineNetwork[].cidr ranges that you specify. For a standard cluster, specify a public and a private subnet for each availability zone. For a private cluster, specify a private subnet for each availability zone.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>platform.aws.serviceEndpoints.url</td>
<td>The AWS service endpoint URL. The URL must use the https protocol and the host must trust the certificate.</td>
<td>Valid AWS service endpoint URL.</td>
</tr>
<tr>
<td>platform.aws.userTags</td>
<td>A map of keys and values that the installation program adds as tags to all resources that it creates.</td>
<td>Any valid YAML map, such as key value pairs in the &lt;key&gt;: &lt;value&gt; format. For more information about AWS tags, see Tagging Your Amazon EC2 Resources in the AWS documentation.</td>
</tr>
<tr>
<td>platform.aws.subnets</td>
<td>If you provide the VPC instead of allowing the installation program to create the VPC for you, specify the subnet for the cluster to use. The subnet must be part of the same machineNetwork[].cidr ranges that you specify. For a standard cluster, specify a public and a private subnet for each availability zone. For a private cluster, specify a private subnet for each availability zone.</td>
<td>Valid subnet IDs.</td>
</tr>
</tbody>
</table>

1.8.8.2. Sample customized install-config.yaml file for AWS

You can customize the install-config.yaml file to specify more details about your OpenShift Container Platform cluster’s platform or modify the values of the required parameters.

IMPORTANT

This sample YAML file is provided for reference only. You must obtain your install-config.yaml file by using the installation program and modify it.

```yaml
apiVersion: v1
baseDomain: example.com
credentialsMode: Mint
controlPlane:
  - hyperthreading: Enabled
  - name: master
platform:
  aws:
    zones:
    - us-gov-west-1a
    - us-gov-west-1b
    rootVolume:
      iops: 4000
      size: 500
```
CHAPTER 1. INSTALLING ON AWS

Required.

Optional: Add this parameter to force the Cloud Credential Operator (CCO) to use the specified
3 If you do not provide these parameters and values, the installation program provides the default value.

4 The `controlPlane` section is a single mapping, but the compute section is a sequence of mappings. To meet the requirements of the different data structures, the first line of the `compute` section must begin with a hyphen, `-`, and the first line of the `controlPlane` section must not. Although both sections currently define a single machine pool, it is possible that future versions of OpenShift Container Platform will support defining multiple compute pools during installation. Only one control plane pool is used.

5 Whether to enable or disable simultaneous multithreading, or hyperthreading. By default, simultaneous multithreading is enabled to increase the performance of your machines’ cores. You can disable it by setting the parameter value to `Disabled`. If you disable simultaneous multithreading in some cluster machines, you must disable it in all cluster machines.

   **IMPORTANT**

   If you disable simultaneous multithreading, ensure that your capacity planning accounts for the dramatically decreased machine performance. Use larger instance types, such as `m4.2xlarge` or `m5.2xlarge`, for your machines if you disable simultaneous multithreading.

6 To configure faster storage for etcd, especially for larger clusters, set the storage type as `io1` and set `iops` to `2000`.

7 If you provide your own VPC, specify subnets for each availability zone that your cluster uses.

8 The ID of the AMI used to boot machines for the cluster. If set, the AMI must belong to the same region as the cluster.

9 The AWS service endpoints. Custom endpoints are required when installing to an unknown AWS region. The endpoint URL must use the `https` protocol and the host must trust the certificate.

10 The ID of your existing Route 53 private hosted zone. Providing an existing hosted zone requires that you supply your own VPC and the hosted zone is already associated with the VPC prior to installing your cluster. If undefined, the installation program creates a new hosted zone.

11 Whether to enable or disable FIPS mode. By default, FIPS mode is not enabled. If FIPS mode is enabled, the Red Hat Enterprise Linux CoreOS (RHCOS) machines that OpenShift Container Platform runs on bypass the default Kubernetes cryptography suite and use the cryptography modules that are provided with RHCOS instead.

   **IMPORTANT**

   The use of FIPS Validated / Modules in Process cryptographic libraries is only supported on OpenShift Container Platform deployments on the `x86_64` architecture.

12 You can optionally provide the `sshKey` value that you use to access the machines in your cluster.

   **NOTE**

   For production OpenShift Container Platform clusters on which you want to perform installation debugging or disaster recovery, specify an SSH key that your `ssh-agent` process uses.
How to publish the user-facing endpoints of your cluster. Set `publish` to `Internal` to deploy a private cluster, which cannot be accessed from the Internet. The default value is `External`.

The custom CA certificate. This is required when deploying to the AWS C2S Secret Region because the AWS API requires a custom CA trust bundle.

### 1.8.8.3. AWS regions without a published RHCOS AMI

You can deploy an OpenShift Container Platform cluster to Amazon Web Services (AWS) regions without native support for a Red Hat Enterprise Linux CoreOS (RHCOS) Amazon Machine Image (AMI) or the AWS software development kit (SDK). If a published AMI is not available for an AWS region, you can upload a custom AMI prior to installing the cluster. This is required if you are deploying your cluster to an AWS government region.

If you are deploying to a non-government region that does not have a published RHCOS AMI, and you do not specify a custom AMI, the installation program copies the `us-east-1` AMI to the user account automatically. Then the installation program creates the control plane machines with encrypted EBS volumes using the default or user-specified Key Management Service (KMS) key. This allows the AMI to follow the same process workflow as published RHCOS AMIs.

A region without native support for an RHCOS AMI is not available to select from the terminal during cluster creation because it is not published. However, you can install to this region by configuring the custom AMI in the `install-config.yaml` file.

### 1.8.8.4. Uploading a custom RHCOS AMI in AWS

If you are deploying to a custom Amazon Web Services (AWS) region, you must upload a custom Red Hat Enterprise Linux CoreOS (RHCOS) Amazon Machine Image (AMI) that belongs to that region.

**Prerequisites**

- You configured an AWS account.
- You created an Amazon S3 bucket with the required IAM service role.
- You uploaded your RHCOS VMDK file to Amazon S3. The RHCOS VMDK file must be the highest version that is less than or equal to the OpenShift Container Platform version you are installing.
- You downloaded the AWS CLI and installed it on your computer. See [Install the AWS CLI Using the Bundled Installer](#).

**Procedure**

1. Export your AWS profile as an environment variable:

   ```
   $ export AWS_PROFILE=<aws_profile>  
   ```

   The AWS profile name that holds your AWS credentials, like `govcloud`.

2. Export the region to associate with your custom AMI as an environment variable:

   ```
   $ export AWS_DEFAULT_REGION=<aws_region>  
   ```
The AWS region, like **us-gov-east-1**.

3. Export the version of RHCOS you uploaded to Amazon S3 as an environment variable:

   ```bash
   $ export RHCOS_VERSION=<version>
   ```

   The RHCOS VMDK version, like **4.6.0**.

4. Export the Amazon S3 bucket name as an environment variable:

   ```bash
   $ export VMIMPORT_BUCKET_NAME=<s3_bucket_name>
   ```

5. Create the **containers.json** file and define your RHCOS VMDK file:

   ```bash
   $ cat <<EOF > containers.json
   {
      "Description": "rhcos-$(RHCOS_VERSION)-x86_64-aws.x86_64",
      "Format": "vmdk",
      "UserBucket": {
         "S3Bucket": "$(VMIMPORT_BUCKET_NAME)",
         "S3Key": "rhcos-$(RHCOS_VERSION)-x86_64-aws.x86_64.vmdk"
      }
   }
   EOF
   ```

6. Import the RHCOS disk as an Amazon EBS snapshot:

   ```bash
   $ aws ec2 import-snapshot --region ${AWS_DEFAULT_REGION} \
   --description "<description>" \
   --disk-container "file://<file_path>/containers.json"
   ```

   The description of your RHCOS disk being imported, like **rhcos-$(RHCOS_VERSION)-x86_64-aws.x86_64**.

   The file path to the JSON file describing your RHCOS disk. The JSON file should contain your Amazon S3 bucket name and key.

7. Check the status of the image import:

   ```bash
   $ watch -n 5 aws ec2 describe-import-snapshot-tasks --region ${AWS_DEFAULT_REGION}
   ```

   **Example output**

   ```json
   {
      "ImportSnapshotTasks": [
         {
            "Description": "rhcos-4.6.0-x86_64-aws.x86_64",
            "ImportTaskId": "import-snap-fh6i8uil",
            "SnapshotTaskDetail": {
              "Description": "rhcos-4.6.0-x86_64-aws.x86_64",
              "DiskImageSize": 819056640.0,
          ```
Copy the **SnapshotId** to register the image.

8. Create a custom RHCOS AMI from the RHCOS snapshot:

```bash
$ aws ec2 register-image
  --region ${AWS_DEFAULT_REGION} \
  --architecture x86_64 \
  --description "rhcos-\${RHCOS_VERSION}-x86_64-aws.x86_64" \
  --ena-support \
  --name "rhcos-\${RHCOS_VERSION}-x86_64-aws.x86_64" \
  --virtualization-type hvm \
  --root-device-name '/dev/xvda' \
  --block-device-mappings 'DeviceName=/dev/xvda,Ebs=
  {DeleteOnTermination=true,SnapshotId=<snapshot_ID>}'
```

1. The RHCOS VMDK architecture type, like **x86_64**, **s390x**, or **ppc64le**.
2. The **Description** from the imported snapshot.
3. The name of the RHCOS AMI.
4. The **SnapshotId** from the imported snapshot.

To learn more about these APIs, see the AWS documentation for importing snapshots and creating EBS-backed AMIs.

### 1.8.8.5. Configuring the cluster-wide proxy during installation

Production environments can deny direct access to the Internet and instead have an HTTP or HTTPS proxy available. You can configure a new OpenShift Container Platform cluster to use a proxy by configuring the proxy settings in the `install-config.yaml` file.

**Prerequisites**

- You have an existing `install-config.yaml` file.
- You reviewed the sites that your cluster requires access to and determined whether any of them need to bypass the proxy. By default, all cluster egress traffic is proxied, including calls to hosting cloud provider APIs. You added sites to the **Proxy** object’s **spec.noProxy** field to bypass the proxy if necessary.
NOTE

The Proxy object status.noProxy field is populated with the values of the networking.machineNetwork[].cidr, networking.clusterNetwork[].cidr, and networking.serviceNetwork[] fields from your installation configuration.

For installations on Amazon Web Services (AWS), Google Cloud Platform (GCP), Microsoft Azure, and Red Hat OpenStack Platform (RHOSP), the Proxy object status.noProxy field is also populated with the instance metadata endpoint (169.254.169.254).

- If your cluster is on AWS, you added the ec2.<region>.amazonaws.com, elasticloadbalancing.<region>.amazonaws.com, and s3.<region>.amazonaws.com endpoints to your VPC endpoint. These endpoints are required to complete requests from the nodes to the AWS EC2 API. Because the proxy works on the container level, not the node level, you must route these requests to the AWS EC2 API through the AWS private network. Adding the public IP address of the EC2 API to your allowlist in your proxy server is not sufficient.

Procedure

1. Edit your install-config.yaml file and add the proxy settings. For example:

```
apiVersion: v1
baseDomain: my.domain.com
proxy:
  httpProxy: http://<username>:<pswd>@<ip>:<port>
  httpsProxy: https://<username>:<pswd>@<ip>:<port>
  noProxy: example.com
additionalTrustBundle: |
  -----BEGIN CERTIFICATE-----
  <MY_TRUSTED_CA_CERT>
  -----END CERTIFICATE-----
...
```

1. A proxy URL to use for creating HTTP connections outside the cluster. The URL scheme must be http.

2. A proxy URL to use for creating HTTPS connections outside the cluster.

3. A comma-separated list of destination domain names, IP addresses, or other network CIDRs to exclude from proxying. Preface a domain with . to match subdomains only. For example, .y.com matches x.y.com, but not y.com. Use * to bypass the proxy for all destinations.

4. If provided, the installation program generates a config map that is named user-ca-bundle in the openshift-config namespace to hold the additional CA certificates. If you provide additionalTrustBundle and at least one proxy setting, the Proxy object is configured to reference the user-ca-bundle config map in the trustedCA field. The Cluster Network Operator then creates a trusted-ca-bundle config map that merges the contents specified for the trustedCA parameter with the RHCOS trust bundle. The additionalTrustBundle field is required unless the proxy’s identity certificate is signed by an authority from the RHCOS trust bundle.
The installation program does not support the proxy `readinessEndpoints` field.

2. Save the file and reference it when installing OpenShift Container Platform.

The installation program creates a cluster-wide proxy that is named `cluster` that uses the proxy settings in the provided `install-config.yaml` file. If no proxy settings are provided, a `cluster Proxy` object is still created, but it will have a nil `spec`.

NOTE

Only the `Proxy` object named `cluster` is supported, and no additional proxies can be created.

1.8.9. Deploying the cluster

You can install OpenShift Container Platform on a compatible cloud platform.

IMPORTANT

You can run the `create cluster` command of the installation program only once, during initial installation.

Prerequisites

- Configure an account with the cloud platform that hosts your cluster.
- Obtain the OpenShift Container Platform installation program and the pull secret for your cluster.

Procedure

1. Change to the directory that contains the installation program and initialize the cluster deployment:

   ```
   $ ./openshift-install create cluster --dir <installation_directory> \
   --log-level=info
   ```

   1 For `<installation_directory>`, specify the location of your customized `./install-config.yaml` file.

   2 To view different installation details, specify `warn`, `debug`, or `error` instead of `info`.

   NOTE

   If the cloud provider account that you configured on your host does not have sufficient permissions to deploy the cluster, the installation process stops, and the missing permissions are displayed.

When the cluster deployment completes, directions for accessing your cluster, including a link to its web console and credentials for the `kubeadmin` user, display in your terminal.
Example output

...  
INFO Install complete!  
INFO To access the cluster as the system:admin user when using 'oc', run 'export KUBECONFIG=/home/myuser/install_dir/auth/kubeconfig'  
INFO Access the OpenShift web-console here: https://console-openshift-console.apps.mycluster.example.com  
INFO Login to the console with user: "kubeadmin", and password: "4vYBz-Ee6gm-ymBZj-Wt5AL"  
INFO Time elapsed: 36m22s

NOTE

The cluster access and credential information also outputs to <installation_directory>/openshift_install.log when an installation succeeds.

IMPORTANT

The Ignition config files that the installation program generates contain certificates that expire after 24 hours, which are then renewed at that time. If the cluster is shut down before renewing the certificates and the cluster is later restarted after the 24 hours have elapsed, the cluster automatically recovers the expired certificates. The exception is that you must manually approve the pending node-bootstrapper certificate signing requests (CSRs) to recover kublet certificates. See the documentation for Recovering from expired control plane certificates for more information.

IMPORTANT

You must not delete the installation program or the files that the installation program creates. Both are required to delete the cluster.

2. Optional: Remove or disable the AdministratorAccess policy from the IAM account that you used to install the cluster.

NOTE

The elevated permissions provided by the AdministratorAccess policy are required only during installation.

1.8.10. Installing the OpenShift CLI by downloading the binary

You can install the OpenShift CLI (oc) in order to interact with OpenShift Container Platform from a command-line interface. You can install oc on Linux, Windows, or macOS.

IMPORTANT

If you installed an earlier version of oc, you cannot use it to complete all of the commands in OpenShift Container Platform 4.6. Download and install the new version of oc.

1.8.10.1. Installing the OpenShift CLI on Linux
You can install the OpenShift CLI (**oc**) binary on Linux by using the following procedure.

**Procedure**

2. Select the appropriate version in the **Version** drop-down menu.
3. Click **Download Now** next to the **OpenShift v4.6 Linux Client** entry and save the file.
4. Unpack the archive:
   ```bash
   $ tar xvzf <file>
   ```
5. Place the **oc** binary in a directory that is on your **PATH**.
   To check your **PATH**, execute the following command:
   ```bash
   $ echo $PATH
   ```

After you install the OpenShift CLI, it is available using the **oc** command:

```bash
$ oc <command>
```

### 1.8.10.2. Installing the OpenShift CLI on Windows

You can install the OpenShift CLI (**oc**) binary on Windows by using the following procedure.

**Procedure**

2. Select the appropriate version in the **Version** drop-down menu.
3. Click **Download Now** next to the **OpenShift v4.6 Windows Client** entry and save the file.
4. Unzip the archive with a ZIP program.
5. Move the **oc** binary to a directory that is on your **PATH**.
   To check your **PATH**, open the command prompt and execute the following command:
   ```bash
   C:\> path
   ```

After you install the OpenShift CLI, it is available using the **oc** command:

```bash
C:\> oc <command>
```

### 1.8.10.3. Installing the OpenShift CLI on macOS

You can install the OpenShift CLI (**oc**) binary on macOS by using the following procedure.
Procedure


2. Select the appropriate version in the Version drop-down menu.

3. Click Download Now next to the OpenShift v4.6 MacOSX Client entry and save the file.

4. Unpack and unzip the archive.

5. Move the oc binary to a directory on your PATH. To check your PATH, open a terminal and execute the following command:

   $ echo $PATH

   After you install the OpenShift CLI, it is available using the oc command:

   $ oc <command>

1.8.11. Logging in to the cluster by using the CLI

You can log in to your cluster as a default system user by exporting the cluster kubeconfig file. The kubeconfig file contains information about the cluster that is used by the CLI to connect a client to the correct cluster and API server. The file is specific to a cluster and is created during OpenShift Container Platform installation.

Prerequisites

- You deployed an OpenShift Container Platform cluster.
- You installed the oc CLI.

Procedure

1. Export the kubeadmin credentials:

   $ export KUBECONFIG=<installation_directory>/auth/kubeconfig

   For <installation_directory>, specify the path to the directory that you stored the installation files in.

2. Verify you can run oc commands successfully using the exported configuration:

   $ oc whoami

   Example output

   system:admin

1.8.12. Logging in to the cluster by using the web console
The **kubeadmin** user exists by default after an OpenShift Container Platform installation. You can log in to your cluster as the **kubeadmin** user by using the OpenShift Container Platform web console.

**Prerequisites**

- You have access to the installation host.
- You completed a cluster installation and all cluster Operators are available.

**Procedure**

1. Obtain the password for the **kubeadmin** user from the **kubeadmin-password** file on the installation host:

   $ cat <installation_directory>/auth/kubeadmin-password

   **NOTE**

   Alternatively, you can obtain the **kubeadmin** password from the `<installation_directory>/openshift_install.log` log file on the installation host.

2. List the OpenShift Container Platform web console route:

   $ oc get routes -n openshift-console | grep 'console-openshift'

   **NOTE**

   Alternatively, you can obtain the OpenShift Container Platform route from the `<installation_directory>/openshift_install.log` log file on the installation host.

   **Example output**

   ```
   console     console-openshift-console.apps.<cluster_name>.<base_domain>            console
   https   reencrypt/Redirect   None
   ```

3. Navigate to the route detailed in the output of the preceding command in a web browser and log in as the **kubeadmin** user.

**Additional resources**

- See [Accessing the web console](#) for more details about accessing and understanding the OpenShift Container Platform web console.

**1.8.13. Telemetry access for OpenShift Container Platform**

In OpenShift Container Platform 4.6, the Telemetry service, which runs by default to provide metrics about cluster health and the success of updates, requires Internet access. If your cluster is connected to the Internet, Telemetry runs automatically, and your cluster is registered to the Red Hat OpenShift Cluster Manager (OCM).
After you confirm that your Red Hat OpenShift Cluster Manager inventory is correct, either maintained automatically by Telemetry or manually by using OCM, use subscription watch to track your OpenShift Container Platform subscriptions at the account or multi-cluster level.

Additional resources

- See About remote health monitoring for more information about the Telemetry service.

1.8.14. Next steps

- Validating an installation.
- Customize your cluster.
- If necessary, you can opt out of remote health reporting.
- If necessary, you can remove cloud provider credentials.

1.9. INSTALLING A CLUSTER ON AWS IN A RESTRICTED NETWORK WITH USER-PROVISIONED INFRASTRUCTURE

In OpenShift Container Platform version 4.6, you can install a cluster on Amazon Web Services (AWS) using infrastructure that you provide and an internal mirror of the installation release content.

**IMPORTANT**

While you can install an OpenShift Container Platform cluster by using mirrored installation release content, your cluster still requires Internet access to use the AWS APIs.

One way to create this infrastructure is to use the provided CloudFormation templates. You can modify the templates to customize your infrastructure or use the information that they contain to create AWS objects according to your company’s policies.

**IMPORTANT**

The steps for performing a user-provisioned infrastructure installation are provided as an example only. Installing a cluster with infrastructure you provide requires knowledge of the cloud provider and the installation process of OpenShift Container Platform. Several CloudFormation templates are provided to assist in completing these steps or to help model your own. You are also free to create the required resources through other methods; the templates are just an example.

1.9.1. Prerequisites

- You created a mirror registry on your mirror host and obtained the `imageContentSources` data for your version of OpenShift Container Platform.

**IMPORTANT**

Because the installation media is on the mirror host, you can use that computer to complete all installation steps.
You reviewed details about the OpenShift Container Platform installation and update processes.

You configured an AWS account to host the cluster.

**IMPORTANT**

If you have an AWS profile stored on your computer, it must not use a temporary session token that you generated while using a multi-factor authentication device. The cluster continues to use your current AWS credentials to create AWS resources for the entire life of the cluster, so you must use key-based, long-lived credentials. To generate appropriate keys, see Managing Access Keys for IAM Users in the AWS documentation. You can supply the keys when you run the installation program.

You downloaded the AWS CLI and installed it on your computer. See Install the AWS CLI Using the Bundled Installer (Linux, macOS, or Unix) in the AWS documentation.

If you use a firewall and plan to use the Telemetry service, you configured the firewall to allow the sites that your cluster requires access to.

**NOTE**

Be sure to also review this site list if you are configuring a proxy.

If you do not allow the system to manage identity and access management (IAM), then a cluster administrator can manually create and maintain IAM credentials. Manual mode can also be used in environments where the cloud IAM APIs are not reachable.

### 1.9.2. About installations in restricted networks

In OpenShift Container Platform 4.6, you can perform an installation that does not require an active connection to the Internet to obtain software components. Restricted network installations can be completed using installer-provisioned infrastructure or user-provisioned infrastructure, depending on the cloud platform to which you are installing the cluster.

If you choose to perform a restricted network installation on a cloud platform, you still require access to its cloud APIs. Some cloud functions, like Amazon Web Service’s IAM service, require Internet access, so you might still require Internet access. Depending on your network, you might require less Internet access for an installation on bare metal hardware or on VMware vSphere.

To complete a restricted network installation, you must create a registry that mirrors the contents of the OpenShift Container Platform registry and contains the installation media. You can create this registry on a mirror host, which can access both the Internet and your closed network, or by using other methods that meet your restrictions.

**IMPORTANT**

Because of the complexity of the configuration for user-provisioned installations, consider completing a standard user-provisioned infrastructure installation before you attempt a restricted network installation using user-provisioned infrastructure. Completing this test installation might make it easier to isolate and troubleshoot any issues that might arise during your installation in a restricted network.
1.9.2.1. Additional limits

Clusters in restricted networks have the following additional limitations and restrictions:

- The **ClusterVersion** status includes an **Unable to retrieve available updates** error.
- By default, you cannot use the contents of the Developer Catalog because you cannot access the required image stream tags.

1.9.3. Internet access for OpenShift Container Platform

In OpenShift Container Platform 4.6, you require access to the Internet to obtain the images that are necessary to install your cluster.

You must have Internet access to:

- Access the **Red Hat OpenShift Cluster Manager** page to download the installation program and perform subscription management. If the cluster has Internet access and you do not disable Telemetry, that service automatically entitles your cluster.
- Access **Quay.io** to obtain the packages that are required to install your cluster.
- Obtain the packages that are required to perform cluster updates.

**IMPORTANT**

If your cluster cannot have direct Internet access, you can perform a restricted network installation on some types of infrastructure that you provision. During that process, you download the content that is required and use it to populate a mirror registry with the packages that you need to install a cluster and generate the installation program. With some installation types, the environment that you install your cluster in will not require Internet access. Before you update the cluster, you update the content of the mirror registry.

1.9.4. Required AWS infrastructure components

To install OpenShift Container Platform on user-provisioned infrastructure in Amazon Web Services (AWS), you must manually create both the machines and their supporting infrastructure.

For more information about the integration testing for different platforms, see the **OpenShift Container Platform 4.x Tested Integrations** page.

By using the provided CloudFormation templates, you can create stacks of AWS resources that represent the following components:

- An AWS Virtual Private Cloud (VPC)
- Networking and load balancing components
- Security groups and roles
- An OpenShift Container Platform bootstrap node
- OpenShift Container Platform control plane nodes
- An OpenShift Container Platform compute node
Alternatively, you can manually create the components or you can reuse existing infrastructure that meets the cluster requirements. Review the CloudFormation templates for more details about how the components interrelate.

1.9.4.1. Cluster machines

You need AWS::EC2::Instance objects for the following machines:

- A bootstrap machine. This machine is required during installation, but you can remove it after your cluster deploys.
- Three control plane machines. The control plane machines are not governed by a machine set.
- Compute machines. You must create at least two compute machines, which are also known as worker machines, during installation. These machines are not governed by a machine set.

You can use the following instance types for the cluster machines with the provided CloudFormation templates.

**IMPORTANT**

If m4 instance types are not available in your region, such as with eu-west-3, use m5 types instead.

Table 1.26. Instance types for machines

<table>
<thead>
<tr>
<th>Instance type</th>
<th>Bootstrap</th>
<th>Control plane</th>
<th>Compute</th>
</tr>
</thead>
<tbody>
<tr>
<td>i3.large</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>m4.large</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>m4.xlarge</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>m4.2xlarge</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>m4.4xlarge</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>m4.8xlarge</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>m4.10xlarge</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>m4.16xlarge</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>m5.large</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>m5.xlarge</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>m5.2xlarge</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Instance type</td>
<td>Bootstrap</td>
<td>Control plane</td>
<td>Compute</td>
</tr>
<tr>
<td>------------------</td>
<td>-----------</td>
<td>---------------</td>
<td>---------</td>
</tr>
<tr>
<td>m5.4xlarge</td>
<td>x</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>m5.8xlarge</td>
<td>x</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>m5.10xlarge</td>
<td>x</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>m5.16xlarge</td>
<td>x</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>m6i.xlarge</td>
<td>x</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>c4.2xlarge</td>
<td>x</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>c4.4xlarge</td>
<td>x</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>c4.8xlarge</td>
<td>x</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>r4.large</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>r4.xlarge</td>
<td>x</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>r4.2xlarge</td>
<td>x</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>r4.4xlarge</td>
<td>x</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>r4.8xlarge</td>
<td>x</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>r4.16xlarge</td>
<td>x</td>
<td></td>
<td>x</td>
</tr>
</tbody>
</table>

You might be able to use other instance types that meet the specifications of these instance types.

1.9.4.2. Other infrastructure components

- A VPC
- DNS entries
- Load balancers (classic or network) and listeners
- A public and a private Route 53 zone
- Security groups
- IAM roles
- S3 buckets
If you are working in a disconnected environment or use a proxy, you cannot reach the public IP addresses for EC2 and ELB endpoints. To reach these endpoints, you must create a VPC endpoint and attach it to the subnet that the clusters are using. Create the following endpoints:

- ec2.<region>.amazonaws.com
- elasticloadbalancing.<region>.amazonaws.com
- s3.<region>.amazonaws.com

Required VPC components
You must provide a suitable VPC and subnets that allow communication to your machines.

<table>
<thead>
<tr>
<th>Component</th>
<th>AWS type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>VPC</td>
<td>• AWS::EC2::VPC</td>
<td>You must provide a public VPC for the cluster to use. The VPC uses an endpoint that references the route tables for each subnet to improve communication with the registry that is hosted in S3.</td>
</tr>
<tr>
<td></td>
<td>• AWS::EC2::VPCEndpoint</td>
<td></td>
</tr>
<tr>
<td>Public subnets</td>
<td>• AWS::EC2::Subnet</td>
<td>Your VPC must have public subnets for between 1 and 3 availability zones and associate them with appropriate Ingress rules.</td>
</tr>
<tr>
<td></td>
<td>• AWS::EC2::SubnetNetworkAclAssociation</td>
<td></td>
</tr>
<tr>
<td>Internet gateway</td>
<td>• AWS::EC2::InternetGateway</td>
<td>You must have a public Internet gateway, with public routes, attached to the VPC. In the provided templates, each public subnet has a NAT gateway with an EIP address. These NAT gateways allow cluster resources, like private subnet instances, to reach the Internet and are not required for some restricted network or proxy scenarios.</td>
</tr>
<tr>
<td></td>
<td>• AWS::EC2::VPCEGatewayAttachment</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• AWS::EC2::RouteTable</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• AWS::EC2::Route</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• AWS::EC2::SubnetRouteTableAssociation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• AWS::EC2::NatGateway</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• AWS::EC2::EIP</td>
<td></td>
</tr>
<tr>
<td>Network access control</td>
<td>• AWS::EC2::NetworkAcl</td>
<td>You must allow the VPC to access the following ports:</td>
</tr>
<tr>
<td></td>
<td>• AWS::EC2::NetworkAclEntry</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Port</th>
<th>Reason</th>
</tr>
</thead>
<tbody>
<tr>
<td>80</td>
<td>Inbound HTTP traffic</td>
</tr>
</tbody>
</table>
### Inbound HTTPS traffic

<table>
<thead>
<tr>
<th>Port</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>443</td>
<td>Inbound HTTPS traffic</td>
</tr>
</tbody>
</table>

### Inbound SSH traffic

<table>
<thead>
<tr>
<th>Port Range</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>22</td>
<td>Inbound SSH traffic</td>
</tr>
</tbody>
</table>

### Inbound ephemeral traffic

<table>
<thead>
<tr>
<th>Port Range</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1024 - 65535</td>
<td>Inbound ephemeral traffic</td>
</tr>
</tbody>
</table>

### Outbound ephemeral traffic

<table>
<thead>
<tr>
<th>Port Range</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 - 65535</td>
<td>Outbound ephemeral traffic</td>
</tr>
</tbody>
</table>

### Private subnets

- **AWS::EC2::Subnet**
- **AWS::EC2::RouteTable**
- **AWS::EC2::SubnetRouteTableAssociation**

Your VPC can have private subnets. The provided CloudFormation templates can create private subnets for between 1 and 3 availability zones. If you use private subnets, you must provide appropriate routes and tables for them.

### Required DNS and load balancing components

Your DNS and load balancer configuration needs to use a public hosted zone and can use a private hosted zone similar to the one that the installation program uses if it provisions the cluster’s infrastructure. You must create a DNS entry that resolves to your load balancer. An entry for `api.<cluster_name>.<domain>` must point to the external load balancer, and an entry for `api-int.<cluster_name>.<domain>` must point to the internal load balancer.

The cluster also requires load balancers and listeners for port 6443, which are required for the Kubernetes API and its extensions, and port 22623, which are required for the Ignition config files for new machines. The targets will be the control plane nodes (also known as the master nodes). Port 6443 must be accessible to both clients external to the cluster and nodes within the cluster. Port 22623 must be accessible to nodes within the cluster.

### Component

<table>
<thead>
<tr>
<th>Component</th>
<th>AWS type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNS</td>
<td><strong>AWS::Route 53::HostedZone</strong></td>
<td>The hosted zone for your internal DNS.</td>
</tr>
<tr>
<td>etcd record sets</td>
<td><strong>AWS::Route 53::RecordSet</strong></td>
<td>The registration records for etcd for your control plane machines.</td>
</tr>
<tr>
<td>Public load balancer</td>
<td><strong>AWS::Elastic LoadBalancingV2::LoadBalancer</strong></td>
<td>The load balancer for your public subnets.</td>
</tr>
<tr>
<td>Component</td>
<td>AWS type</td>
<td>Description</td>
</tr>
<tr>
<td>----------------------------</td>
<td>---------------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>External API server record</td>
<td>AWS::Route 53::RecordSetGroup</td>
<td>Alias records for the external API server.</td>
</tr>
<tr>
<td>External listener</td>
<td>AWS::ElasticLoadBalancingV2::Listener</td>
<td>A listener on port 6443 for the external load balancer.</td>
</tr>
<tr>
<td>External target group</td>
<td>AWS::ElasticLoadBalancingV2::TargetGroup</td>
<td>The target group for the external load balancer.</td>
</tr>
<tr>
<td>Private load balancer</td>
<td>AWS::ElasticLoadBalancingV2::LoadBalancer</td>
<td>The load balancer for your private subnets.</td>
</tr>
<tr>
<td>Internal API server record</td>
<td>AWS::Route 53::RecordSetGroup</td>
<td>Alias records for the internal API server.</td>
</tr>
<tr>
<td>Internal listener</td>
<td>AWS::ElasticLoadBalancingV2::Listener</td>
<td>A listener on port 22623 for the internal load balancer.</td>
</tr>
<tr>
<td>Internal target group</td>
<td>AWS::ElasticLoadBalancingV2::TargetGroup</td>
<td>The target group for the internal load balancer.</td>
</tr>
<tr>
<td>Internal listener</td>
<td>AWS::ElasticLoadBalancingV2::Listener</td>
<td>A listener on port 6443 for the internal load balancer.</td>
</tr>
<tr>
<td>Internal target group</td>
<td>AWS::ElasticLoadBalancingV2::TargetGroup</td>
<td>The target group for the internal load balancer.</td>
</tr>
</tbody>
</table>

**Security groups**

The control plane and worker machines require access to the following ports:
<table>
<thead>
<tr>
<th>Group</th>
<th>Type</th>
<th>IP Protocol</th>
<th>Port range</th>
</tr>
</thead>
<tbody>
<tr>
<td>MasterSecurityGroup</td>
<td>AWS::EC2::SecurityGroup</td>
<td>icmp</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>tcp</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td></td>
<td>tcp</td>
<td>6443</td>
</tr>
<tr>
<td></td>
<td></td>
<td>tcp</td>
<td>22623</td>
</tr>
<tr>
<td>WorkerSecurityGroup</td>
<td>AWS::EC2::SecurityGroup</td>
<td>icmp</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>tcp</td>
<td>22</td>
</tr>
<tr>
<td>BootstrapSecurityGroup</td>
<td>AWS::EC2::SecurityGroup</td>
<td>tcp</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td></td>
<td>tcp</td>
<td>19531</td>
</tr>
</tbody>
</table>

**Control plane Ingress**

The control plane machines require the following Ingress groups. Each Ingress group is a **AWS::EC2::SecurityGroupIngress** resource.

<table>
<thead>
<tr>
<th>Ingress group</th>
<th>Description</th>
<th>IP protocol</th>
<th>Port range</th>
</tr>
</thead>
<tbody>
<tr>
<td>MasterIngress Etcd</td>
<td>etcd</td>
<td>tcp</td>
<td>2379-2380</td>
</tr>
<tr>
<td>MasterIngress Vxlan</td>
<td>Vxlan packets</td>
<td>udp</td>
<td>4789</td>
</tr>
<tr>
<td>MasterIngress Vxlan</td>
<td>Vxlan packets</td>
<td>udp</td>
<td>4789</td>
</tr>
<tr>
<td>MasterIngress Internal</td>
<td>Internal cluster communication and Kubernetes proxy metrics</td>
<td>tcp</td>
<td>9000 - 9999</td>
</tr>
<tr>
<td>MasterIngress WorkerInternal</td>
<td>Internal cluster communication</td>
<td>tcp</td>
<td>9000 - 9999</td>
</tr>
<tr>
<td>MasterIngress Kube</td>
<td>Kubernetes kubelet, scheduler and controller manager</td>
<td>tcp</td>
<td>10250 - 10259</td>
</tr>
<tr>
<td>MasterIngress WorkerKube</td>
<td>Kubernetes kubelet, scheduler and controller manager</td>
<td>tcp</td>
<td>10250 - 10259</td>
</tr>
</tbody>
</table>
### Kubernetes Ingress services

<table>
<thead>
<tr>
<th>Ingress group</th>
<th>Description</th>
<th>IP protocol</th>
<th>Port range</th>
</tr>
</thead>
<tbody>
<tr>
<td>MasterIngress Services</td>
<td>Kubernetes Ingress services</td>
<td>tcp</td>
<td>30000 - 32767</td>
</tr>
<tr>
<td>MasterIngress WorkerIngress</td>
<td>Kubernetes Ingress services</td>
<td>tcp</td>
<td>30000 - 32767</td>
</tr>
</tbody>
</table>

### Worker Ingress

The worker machines require the following Ingress groups. Each Ingress group is a **AWS::EC2::SecurityGroupIngress** resource.

<table>
<thead>
<tr>
<th>Ingress group</th>
<th>Description</th>
<th>IP protocol</th>
<th>Port range</th>
</tr>
</thead>
<tbody>
<tr>
<td>WorkerIngress Vxlan</td>
<td>Vxlan packets</td>
<td>udp</td>
<td>4789</td>
</tr>
<tr>
<td>WorkerIngress Vxlan</td>
<td>Vxlan packets</td>
<td>udp</td>
<td>4789</td>
</tr>
<tr>
<td>WorkerIngress Internal</td>
<td>Internal cluster communication</td>
<td>tcp</td>
<td>9000 - 9999</td>
</tr>
<tr>
<td>WorkerIngress Internal</td>
<td>Internal cluster communication</td>
<td>tcp</td>
<td>9000 - 9999</td>
</tr>
<tr>
<td>WorkerIngress Kube</td>
<td>Kubernetes kubelet, scheduler, and controller manager</td>
<td>tcp</td>
<td>10250</td>
</tr>
<tr>
<td>WorkerIngress Kube</td>
<td>Kubernetes kubelet, scheduler, and controller manager</td>
<td>tcp</td>
<td>10250</td>
</tr>
<tr>
<td>WorkerIngress Services</td>
<td>Kubernetes Ingress services</td>
<td>tcp</td>
<td>30000 - 32767</td>
</tr>
<tr>
<td>WorkerIngress Services</td>
<td>Kubernetes Ingress services</td>
<td>tcp</td>
<td>30000 - 32767</td>
</tr>
</tbody>
</table>

### Roles and instance profiles

You must grant the machines permissions in AWS. The provided CloudFormation templates grant the machines **Allow** permissions for the following **AWS::IAM::Role** objects and provide a **AWS::IAM::InstanceProfile** for each set of roles. If you do not use the templates, you can grant the machines the following broad permissions or the following individual permissions.
<table>
<thead>
<tr>
<th>Role</th>
<th>Effect</th>
<th>Action</th>
<th>Resource</th>
</tr>
</thead>
<tbody>
<tr>
<td>Master</td>
<td>Allow</td>
<td>ec2:*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>Allow</td>
<td>elasticloadbalancing:*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>Allow</td>
<td>iam:PassRole</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>Allow</td>
<td>s3:GetObject</td>
<td>*</td>
</tr>
<tr>
<td>Worker</td>
<td>Allow</td>
<td>ec2:Describe*</td>
<td>*</td>
</tr>
<tr>
<td>Bootstrap</td>
<td>Allow</td>
<td>ec2:Describe*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>Allow</td>
<td>ec2:AttachVolume</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>Allow</td>
<td>ec2:DetachVolume</td>
<td>*</td>
</tr>
</tbody>
</table>

### 1.9.4.3. Certificate signing requests management

Because your cluster has limited access to automatic machine management when you use infrastructure that you provision, you must provide a mechanism for approving cluster certificate signing requests (CSRs) after installation. The `kube-controller-manager` only approves the kubelet client CSRs. The `machine-approver` cannot guarantee the validity of a serving certificate that is requested by using kubelet credentials because it cannot confirm that the correct machine issued the request. You must determine and implement a method of verifying the validity of the kubelet serving certificate requests and approving them.

### 1.9.4.4. Required AWS permissions

When you attach the `AdministratorAccess` policy to the IAM user that you create in Amazon Web Services (AWS), you grant that user all of the required permissions. To deploy all components of an OpenShift Container Platform cluster, the IAM user requires the following permissions:

**Example 1.13. Required EC2 permissions for installation**

- tag:TagResources
- tag:UntagResources
- ec2:AllocateAddress
- ec2:AssociateAddress
- ec2:AuthorizeSecurityGroupEgress
- ec2:AuthorizeSecurityGroupIngress
- ec2:CopyImage
- ec2:CreateNetworkInterface
• ec2:AttachNetworkInterface
• ec2:CreateSecurityGroup
• ec2:CreateTags
• ec2:CreateVolume
• ec2:DeleteSecurityGroup
• ec2:DeleteSnapshot
• ec2:DeleteTags
• ec2:DeregisterImage
• ec2:DescribeAccountAttributes
• ec2:DescribeAddresses
• ec2:DescribeAvailabilityZones
• ec2:DescribeDhcpOptions
• ec2:DescribeImages
• ec2:DescribeInstanceAttribute
• ec2:DescribeInstanceCreditSpecifications
• ec2:DescribeInstances
• ec2:DescribeInternetGateways
• ec2:DescribeKeyPairs
• ec2:DescribeNatGateways
• ec2:DescribeNetworkAcls
• ec2:DescribeNetworkInterfaces
• ec2:DescribePrefixLists
• ec2:DescribeRegions
• ec2:DescribeRouteTables
• ec2:DescribeSecurityGroups
• ec2:DescribeSubnets
• ec2:DescribeTags
• ec2:DescribeVolumes
• ec2:DescribeVpcAttribute
Example 1.14. Required permissions for creating network resources during installation

- `ec2:AssociateDhcpOptions`
- `ec2:AssociateRouteTable`
- `ec2:AttachInternetGateway`
- `ec2:CreateDhcpOptions`
- `ec2:CreateInternetGateway`
- `ec2:CreateNatGateway`
- `ec2:CreateRoute`
- `ec2:CreateRouteTable`
- `ec2:CreateSubnet`
- `ec2:CreateVpc`
- `ec2:CreateVpcEndpoint`
- `ec2:ModifySubnetAttribute`
- `ec2:ModifyVpcAttribute`

**NOTE**
If you use an existing VPC, your account does not require these permissions for creating network resources.
Example 1.15. Required Elastic Load Balancing permissions (ELB) for installation

- `elasticloadbalancing:AddTags`
- `elasticloadbalancing:ApplySecurityGroupsToLoadBalancer`
- `elasticloadbalancing:AttachLoadBalancerToSubnets`
- `elasticloadbalancing:ConfigureHealthCheck`
- `elasticloadbalancing:CreateLoadBalancer`
- `elasticloadbalancing:CreateLoadBalancerListeners`
- `elasticloadbalancing:DeleteLoadBalancer`
- `elasticloadbalancing:DeregisterInstancesFromLoadBalancer`
- `elasticloadbalancing:DescribeInstanceHealth`
- `elasticloadbalancing:DescribeLoadBalancerAttributes`
- `elasticloadbalancing:DescribeLoadBalancers`
- `elasticloadbalancing:DescribeTags`
- `elasticloadbalancing:ModifyLoadBalancerAttributes`
- `elasticloadbalancing:RegisterInstancesWithLoadBalancer`
- `elasticloadbalancing:SetLoadBalancerPoliciesOfListener`

Example 1.16. Required Elastic Load Balancing permissions (ELBv2) for installation

- `elasticloadbalancing:AddTags`
- `elasticloadbalancing:CreateListener`
- `elasticloadbalancing:CreateLoadBalancer`
- `elasticloadbalancing:CreateTargetGroup`
- `elasticloadbalancing:DeleteLoadBalancer`
- `elasticloadbalancing:DeregisterTargets`
- `elasticloadbalancing:DescribeListeners`
- `elasticloadbalancing:DescribeLoadBalancerAttributes`
- `elasticloadbalancing:DescribeLoadBalancers`
- `elasticloadbalancing:DescribeTargetGroupAttributes`
Example 1.17. Required IAM permissions for installation

- `iam:AddRoleToInstanceProfile`
- `iam:CreateInstanceProfile`
- `iam:CreateRole`
- `iam:DeleteInstanceProfile`
- `iam:DeleteRole`
- `iam:DeleteRolePolicy`
- `iam:GetInstanceProfile`
- `iam:GetRole`
- `iam:GetRolePolicy`
- `iam:GetUser`
- `iam:ListInstanceProfilesForRole`
- `iam:ListRoles`
- `iam:ListUsers`
- `iam:PassRole`
- `iam:PutRolePolicy`
- `iam:RemoveRoleFromInstanceProfile`
- `iam:SimulatePrincipalPolicy`
- `iam:TagRole`

**NOTE**

If you have not created an elastic load balancer (ELB) in your AWS account, the IAM user also requires the `iam:CreateServiceLinkedRole` permission.

Example 1.18. Required Route 53 permissions for installation
- route53:ChangeResourceRecordSets
- route53:ChangeTagsForResource
- route53:CreateHostedZone
- route53:DeleteHostedZone
- route53:GetChange
- route53:GetHostedZone
- route53:ListHostedZones
- route53:ListHostedZonesByName
- route53:ListResourceRecordSets
- route53:ListTagsForResource
- route53:UpdateHostedZoneComment

Example 1.19. Required S3 permissions for installation

- s3:CreateBucket
- s3:DeleteBucket
- s3:GetAccelerateConfiguration
- s3:GetBucketAcl
- s3:GetBucketCors
- s3:GetBucketLocation
- s3:GetBucketLogging
- s3:GetBucketObjectLockConfiguration
- s3:GetBucketReplication
- s3:GetBucketRequestPayment
- s3:GetBucketTagging
- s3:GetBucketVersioning
- s3:GetBucketWebsite
- s3:GetEncryptionConfiguration
- s3:GetLifecycleConfiguration
- s3:GetReplicationConfiguration
- s3:ListBucket
Example 1.20. S3 permissions that cluster Operators require

- s3:DeleteObject
- s3:GetObject
- s3:GetObjectAcl
- s3:GetObjectTagging
- s3:GetObjectVersion
- s3:PutObject
- s3:PutObjectAcl
- s3:PutObjectTagging

Example 1.21. Required permissions to delete base cluster resources

- autoscaling:DescribeAutoScalingGroups
- ec2:DeleteNetworkInterface
- ec2:DeleteVolume
- elasticloadbalancing:DeleteTargetGroup
- elasticloadbalancing:DescribeTargetGroups
- iam:DeleteAccessKey
- iam:DeleteUser
- iam:ListAttachedRolePolicies
- iam:ListInstanceProfiles
- iam:ListRolePolicies
- iam:ListUserPolicies
- s3:DeleteObject
- s3:ListBucketVersions
- tag:GetResources
Example 1.22. Required permissions to delete network resources

- ec2:DeleteDhcpOptions
- ec2:DeleteInternetGateway
- ec2:DeleteNatGateway
- ec2:DeleteRoute
- ec2:DeleteRouteTable
- ec2:DeleteSubnet
- ec2:DeleteVpc
- ec2:DeleteVpcEndpoints
- ec2:DetachInternetGateway
- ec2:DisassociateRouteTable
- ec2:ReplaceRouteTableAssociation

NOTE
If you use an existing VPC, your account does not require these permissions to delete network resources.

Example 1.23. Additional IAM and S3 permissions that are required to create manifests

- iam:DeleteAccessKey
- iam:DeleteUser
- iam:DeleteUserPolicy
- iam:GetUserPolicy
- iam:ListAccessKeys
- iam:PutUserPolicy
- iam:TagUser
- iam:GetUserPolicy
- iam:ListAccessKeys
- s3:PutBucketPublicAccessBlock
- s3:GetBucketPublicAccessBlock
- s3:PutLifecycleConfiguration
- s3:HeadBucket
• s3:ListBucketMultipartUploads
• s3:AbortMultipartUpload

NOTE
If you are managing your cloud provider credentials with mint mode, the IAM user also requires the `iam:CreateAccessKey` and `iam:CreateUser` permissions.

Example 1.24. Optional permission for quota checks for installation
• servicequotas:ListAWSDefaultServiceQuotas

1.9.5. Generating an SSH private key and adding it to the agent

If you want to perform installation debugging or disaster recovery on your cluster, you must provide an SSH key to both your `ssh-agent` and the installation program. You can use this key to access the bootstrap machine in a public cluster to troubleshoot installation issues.

NOTE
In a production environment, you require disaster recovery and debugging.

You can use this key to SSH into the master nodes as the user `core`. When you deploy the cluster, the key is added to the `core` user’s `~/.ssh/authorized_keys` list.

NOTE
You must use a local key, not one that you configured with platform-specific approaches such as AWS key pairs.

Procedure

1. If you do not have an SSH key that is configured for password-less authentication on your computer, create one. For example, on a computer that uses a Linux operating system, run the following command:

   $ ssh-keygen -t ed25519 -N "\n   -f <path>/<file_name> 1

   Specify the path and file name, such as `~/.ssh/id_rsa`, of the new SSH key. If you have an existing key pair, ensure your public key is in the your `~/.ssh` directory.

   Running this command generates an SSH key that does not require a password in the location that you specified.
NOTE

If you plan to install an OpenShift Container Platform cluster that uses FIPS Validated / Modules in Process cryptographic libraries on the x86_64 architecture, do not create a key that uses the ed25519 algorithm. Instead, create a key that uses the rsa or ecdsa algorithm.

2. Start the ssh-agent process as a background task:

   $ eval "$(ssh-agent -s)"

Example output

   Agent pid 31874

NOTE

If your cluster is in FIPS mode, only use FIPS-compliant algorithms to generate the SSH key. The key must be either RSA or ECDSA.

3. Add your SSH private key to the ssh-agent:

   $ ssh-add <path>/<file_name>  

Example output

   Identity added: /home/<you>/<path>/<file_name> (<computer_name>)

   Specify the path and file name for your SSH private key, such as ~/.ssh/id_rsa

Next steps

- When you install OpenShift Container Platform, provide the SSH public key to the installation program. If you install a cluster on infrastructure that you provision, you must provide this key to your cluster’s machines.

1.9.6. Creating the installation files for AWS

To install OpenShift Container Platform on Amazon Web Services (AWS) using user-provisioned infrastructure, you must generate the files that the installation program needs to deploy your cluster and modify them so that the cluster creates only the machines that it will use. You generate and customize the install-config.yaml file, Kubernetes manifests, and Ignition config files. You also have the option to first set up a separate var partition during the preparation phases of installation.

1.9.6.1. Optional: Creating a separate /var partition

It is recommended that disk partitioning for OpenShift Container Platform be left to the installer. However, there are cases where you might want to create separate partitions in a part of the filesystem that you expect to grow.
OpenShift Container Platform supports the addition of a single partition to attach storage to either the /var partition or a subdirectory of /var. For example:

- /var/lib/containers: Holds container-related content that can grow as more images and containers are added to a system.
- /var/lib/etcd: Holds data that you might want to keep separate for purposes such as performance optimization of etcd storage.
- /var: Holds data that you might want to keep separate for purposes such as auditing.

Storing the contents of a /var directory separately makes it easier to grow storage for those areas as needed and reinstall OpenShift Container Platform at a later date and keep that data intact. With this method, you will not have to pull all your containers again, nor will you have to copy massive log files when you update systems.

Because /var must be in place before a fresh installation of Red Hat Enterprise Linux CoreOS (RHCOS), the following procedure sets up the separate /var partition by creating a machine config that is inserted during the openshift-install preparation phases of an OpenShift Container Platform installation.

**IMPORTANT**

If you follow the steps to create a separate /var partition in this procedure, it is not necessary to create the Kubernetes manifest and Ignition config files again as described later in this section.

**Procedure**

1. Create a directory to hold the OpenShift Container Platform installation files:

   ```
   $ mkdir $HOME/clusterconfig
   ```

2. Run openshift-install to create a set of files in the manifest and openshift subdirectories. Answer the system questions as you are prompted:

   ```
   $ openshift-install create manifests --dir $HOME/clusterconfig
   ```

   **Example output**

   ```
   ? SSH Public Key ...
   INFO Credentials loaded from the "myprofile" profile in file "/home/myuser/.aws/credentials"
   INFO Consuming Install Config from target directory
   INFO Manifests created in: $HOME/clusterconfig/manifests and $HOME/clusterconfig/openshift
   ```

3. Optional: Confirm that the installation program created manifests in the clusterconfig/openshift directory:

   ```
   $ ls $HOME/clusterconfig/openshift/
   ```

   **Example output**

   ```
   99_kubeadmin-password-secret.yaml
   99_openshift-cluster-api_master-machines-0.yaml
   ```
4. Create a MachineConfig object and add it to a file in the openshift directory. For example, name the file 98-var-partition.yaml, change the disk device name to the name of the storage device on the worker systems, and set the storage size as appropriate. This example places the /var directory on a separate partition:

apiVersion: machineconfiguration.openshift.io/v1
kind: MachineConfig
metadata:
  labels:
    machineconfiguration.openshift.io/role: worker
name: 98-var-partition
spec:
  config:
    ignition:
      version: 3.1.0
  storage:
    disks:
      - device: /dev/<device_name>  
        partitions:
          - label: var
            startMiB: <partition_start_offset>  
            sizeMiB: <partition_size>  
        filesystems:
          - device: /dev/disk/by-partlabel/var
            path: /var
            format: xfs
  systemd:
    units:
      - name: var.mount  
        enabled: true
        contents: |
          [Unit]
          Before=local-fs.target
          [Mount]
          What=/dev/disk/by-partlabel/var
          Where=/var
          Options=defaults,prjquota  
          [Install]
          WantedBy=local-fs.target

1. The storage device name of the disk that you want to partition.
2. When adding a data partition to the boot disk, a minimum value of 25000 MiB (Mebibytes) is recommended. The root file system is automatically resized to fill all available space up to the specified offset. If no value is specified, or if the specified value is smaller than the recommended minimum, the resulting root file system will be too small, and future reinstalls of RHCOS might overwrite the beginning of the data partition.
3. The size of the data partition in mebibytes.
4. The name of the mount unit must match the directory specified in the Where= directive.
The name of the mount unit must match the directory specified in the `Where=` directive. For example, for a filesystem mounted on `/var/lib/containers`, the unit must be named `var-`

The `prjquota` mount option must be enabled for filesystems used for container storage.

**NOTE**

When creating a separate `/var` partition, you cannot use different instance types for worker nodes, if the different instance types do not have the same device name.

5. Run `openshift-install` again to create Ignition configs from a set of files in the `manifest` and `openshift` subdirectories:

```bash
$ openshift-install create ignition-configs --dir $HOME/clusterconfig
$ ls $HOME/clusterconfig/
auth  bootstrap.ign  master.ign  metadata.json  worker.ign
```

Now you can use the Ignition config files as input to the installation procedures to install Red Hat Enterprise Linux CoreOS (RHCOS) systems.

### 1.9.6.2. Creating the installation configuration file

Generate and customize the installation configuration file that the installation program needs to deploy your cluster.

**Prerequisites**

- You obtained the OpenShift Container Platform installation program for user-provisioned infrastructure and the pull secret for your cluster. For a restricted network installation, these files are on your mirror host.

- You checked that you are deploying your cluster to a region with an accompanying Red Hat Enterprise Linux CoreOS (RHCOS) AMI published by Red Hat. If you are deploying to a region that requires a custom AMI, such as an AWS GovCloud region, you must create the `install-config.yaml` file manually.

**Procedure**

1. Create the `install-config.yaml` file.

   a. Change to the directory that contains the installation program and run the following command:

   ```bash
   $ ./openshift-install create install-config --dir <installation_directory>
   ```

   For `<installation_directory>`, specify the directory name to store the files that the installation program creates.
b. At the prompts, provide the configuration details for your cloud:
   i. Optional: Select an SSH key to use to access your cluster machines.

   **NOTE**
   For production OpenShift Container Platform clusters on which you want to perform installation debugging or disaster recovery, specify an SSH key that your `ssh-agent` process uses.

   ii. Select **aws** as the platform to target.

   iii. If you do not have an AWS profile stored on your computer, enter the AWS access key ID and secret access key for the user that you configured to run the installation program.

   **NOTE**
   The AWS access key ID and secret access key are stored in `~/.aws/credentials` in the home directory of the current user on the installation host. You are prompted for the credentials by the installation program if the credentials for the exported profile are not present in the file. Any credentials that you provide to the installation program are stored in the file.

   iv. Select the AWS region to deploy the cluster to.

   v. Select the base domain for the Route 53 service that you configured for your cluster.

   vi. Enter a descriptive name for your cluster.

   vii. Paste the pull secret that you obtained from the **Pull Secret** page on the Red Hat OpenShift Cluster Manager site.

2. Edit the `install-config.yaml` file to provide the additional information that is required for an installation in a restricted network.

   a. Update the `pullSecret` value to contain the authentication information for your registry:

   ```yaml
   pullSecret: '{"auths":{"<local_registry>": {"auth": "<credentials>"},"email": "you@example.com"}}'
   ```

   For `<local_registry>`, specify the registry domain name, and optionally the port, that your mirror registry uses to serve content. For example `registry.example.com` or `registry.example.com:5000`. For `<credentials>`, specify the base64-encoded user name
and password for your mirror registry.

b. Add the `additionalTrustBundle` parameter and value. The value must be the contents of the certificate file that you used for your mirror registry, which can be an existing, trusted certificate authority or the self-signed certificate that you generated for the mirror registry.

```
additionalTrustBundle: |
-----BEGIN CERTIFICATE-----
ZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZ
-----END CERTIFICATE-----
```

c. Add the image content resources:

```
imageContentSources:
- mirrors:
  - <local_registry>/<local_repository_name>/release
    source: quay.io/openshift-release-dev/ocp-release
- mirrors:
  - <local_registry>/<local_repository_name>/release
    source: quay.io/openshift-release-dev/ocp-v4.0-art-dev
```

Use the `imageContentSources` section from the output of the command to mirror the repository or the values that you used when you mirrored the content from the media that you brought into your restricted network.

d. Optional: Set the publishing strategy to `Internal`:

```
publish: Internal
```

By setting this option, you create an internal Ingress Controller and a private load balancer.

3. Optional: Back up the `install-config.yaml` file.

**IMPORTANT**

The `install-config.yaml` file is consumed during the installation process. If you want to reuse the file, you must back it up now.

**Additional resources**

- See [Configuration and credential file settings](#) in the AWS documentation for more information about AWS profile and credential configuration.

**1.9.6.3. Configuring the cluster-wide proxy during installation**

Production environments can deny direct access to the Internet and instead have an HTTP or HTTPS proxy available. You can configure a new OpenShift Container Platform cluster to use a proxy by configuring the proxy settings in the `install-config.yaml` file.

**Prerequisites**

- You have an existing `install-config.yaml` file.
You reviewed the sites that your cluster requires access to and determined whether any of them need to bypass the proxy. By default, all cluster egress traffic is proxied, including calls to hosting cloud provider APIs. You added sites to the Proxy object’s `spec.noProxy` field to bypass the proxy if necessary.

**NOTE**

The Proxy object `status.noProxy` field is populated with the values of the `networking.machineNetwork[].cidr`, `networking.clusterNetwork[].cidr`, and `networking.serviceNetwork[]` fields from your installation configuration.

For installations on Amazon Web Services (AWS), Google Cloud Platform (GCP), Microsoft Azure, and Red Hat OpenStack Platform (RHOSP), the Proxy object `status.noProxy` field is also populated with the instance metadata endpoint (169.254.169.254).

If your cluster is on AWS, you added the `ec2.<region>.amazonaws.com`, `elasticloadbalancing.<region>.amazonaws.com`, and `s3.<region>.amazonaws.com` endpoints to your VPC endpoint. These endpoints are required to complete requests from the nodes to the AWS EC2 API. Because the proxy works on the container level, not the node level, you must route these requests to the AWS EC2 API through the AWS private network. Adding the public IP address of the EC2 API to your allowlist in your proxy server is not sufficient.

**Procedure**

1. Edit your `install-config.yaml` file and add the proxy settings. For example:

   ```yaml
   apiVersion: v1
   baseDomain: my.domain.com
   proxy:
     httpProxy: http://<username>:<pswd>@<ip>:<port> ¹
     httpsProxy: https://<username>:<pswd>@<ip>:<port> ²
     noProxy: example.com ³
   additionalTrustBundle: |
     -----BEGIN CERTIFICATE-----
     <MY_TRUSTED_CA_CERT>
     -----END CERTIFICATE-----
   ...
   ```

   ¹ A proxy URL to use for creating HTTP connections outside the cluster. The URL scheme must be http.

   ² A proxy URL to use for creating HTTPS connections outside the cluster.

   ³ A comma-separated list of destination domain names, IP addresses, or other network CIDRs to exclude from proxying. Preface a domain with . to match subdomains only. For example, .y.com matches x.y.com, but not y.com. Use * to bypass the proxy for all destinations.

   ⁴ If provided, the installation program generates a config map that is named `user-ca-bundle` in the `openshift-config` namespace to hold the additional CA certificates. If you provide `additionalTrustBundle` and at least one proxy setting, the `Proxy` object is configured to reference the `user-ca-bundle` config map in the `trustedCA` field. The Cluster Network Operator then creates a `trusted-ca-bundle` config map that merges the contents specified for the `trustedCA` parameter with the RHCOS trust bundle. The
additionalTrustBundle field is required unless the proxy’s identity certificate is signed by an authority from the RHCOS trust bundle.

**NOTE**

The installation program does not support the proxy readinessEndpoints field.

2. Save the file and reference it when installing OpenShift Container Platform.

The installation program creates a cluster-wide proxy that is named cluster that uses the proxy settings in the provided install-config.yaml file. If no proxy settings are provided, a cluster Proxy object is still created, but it will have a nil spec.

**NOTE**

Only the Proxy object named cluster is supported, and no additional proxies can be created.

1.9.6.4. Creating the Kubernetes manifest and Ignition config files

Because you must modify some cluster definition files and manually start the cluster machines, you must generate the Kubernetes manifest and Ignition config files that the cluster needs to make its machines.

The installation configuration file transforms into the Kubernetes manifests. The manifests wrap into the Ignition configuration files, which are later used to create the cluster.

**IMPORTANT**

The Ignition config files that the installation program generates contain certificates that expire after 24 hours, which are then renewed at that time. If the cluster is shut down before renewing the certificates and the cluster is later restarted after the 24 hours have elapsed, the cluster automatically recovers the expired certificates. The exception is that you must manually approve the pending node-bootstrapper certificate signing requests (CSRs) to recover kubelet certificates. See the documentation for Recovering from expired control plane certificates for more information.

**Prerequisites**

- You obtained the OpenShift Container Platform installation program. For a restricted network installation, these files are on your mirror host.

- You created the install-config.yaml installation configuration file.

**Procedure**

1. Change to the directory that contains the installation program and generate the Kubernetes manifests for the cluster:

   $ ./openshift-install create manifests --dir <installation_directory>  

   **Example output**
For `<installation_directory>`, specify the installation directory that contains the `install-config.yaml` file you created.

2. Remove the Kubernetes manifest files that define the control plane machines:

   ```bash
   $ rm -f <installation_directory>/openshift/99_openshift-cluster-api_master-machines-*.yaml
   
   By removing these files, you prevent the cluster from automatically generating control plane machines.
   
3. Remove the Kubernetes manifest files that define the worker machines:

   ```bash
   $ rm -f <installation_directory>/openshift/99_openshift-cluster-api_worker-machineset-*.yaml
   
   Because you create and manage the worker machines yourself, you do not need to initialize these machines.
   
4. Check that the `mastersSchedulable` parameter in the `<installation_directory>/manifests/cluster-scheduler-02-config.yml` Kubernetes manifest file is set to `false`. This setting prevents pods from being scheduled on the control plane machines:

   a. Open the `<installation_directory>/manifests/cluster-scheduler-02-config.yml` file.

   b. Locate the `mastersSchedulable` parameter and ensure that it is set to `false`.

   c. Save and exit the file.

5. Optional: If you do not want the Ingress Operator to create DNS records on your behalf, remove the `privateZone` and `publicZone` sections from the `<installation_directory>/manifests/cluster-dns-02-config.yml` DNS configuration file:

   ```yaml
   apiVersion: config.openshift.io/v1
   kind: DNS
   metadata:
     creationTimestamp: null
   name: cluster
   spec:
     baseDomain: example.openshift.com
     privateZone: 1
       id: mycluster-100419-private-zone
     publicZone: 2
       id: example.openshift.com
   status: {}
   ```

   Remove this section completely.

   If you do so, you must add ingress DNS records manually in a later step.
6. To create the Ignition configuration files, run the following command from the directory that contains the installation program:

```
$ ./openshift-install create ignition-configs --dir <installation_directory>
```

For `<installation_directory>`, specify the same installation directory.

The following files are generated in the directory:

```
├── auth
│   ├── kubeadmin-password
│   └── kubecfg
├── bootstrap.ign
├── master.ign
├── metadata.json
└── worker.ign
```

1.9.7. Extracting the infrastructure name

The Ignition config files contain a unique cluster identifier that you can use to uniquely identify your cluster in Amazon Web Services (AWS). The infrastructure name is also used to locate the appropriate AWS resources during an OpenShift Container Platform installation. The provided CloudFormation templates contain references to this infrastructure name, so you must extract it.

**Prerequisites**

- You obtained the OpenShift Container Platform installation program and the pull secret for your cluster.
- You generated the Ignition config files for your cluster.
- You installed the `jq` package.

**Procedure**

- To extract and view the infrastructure name from the Ignition config file metadata, run the following command:

```
$ jq -r .infraID <installation_directory>/metadata.json
```

For `<installation_directory>`, specify the path to the directory that you stored the installation files in.

**Example output**

```
openshift-vw9j6
```

The output of this command is your cluster name and a random string.
1.9.8. Creating a VPC in AWS

You must create a Virtual Private Cloud (VPC) in Amazon Web Services (AWS) for your OpenShift Container Platform cluster to use. You can customize the VPC to meet your requirements, including VPN and route tables.

You can use the provided CloudFormation template and a custom parameter file to create a stack of AWS resources that represent the VPC.

**NOTE**

If you do not use the provided CloudFormation template to create your AWS infrastructure, you must review the provided information and manually create the infrastructure. If your cluster does not initialize correctly, you might have to contact Red Hat support with your installation logs.

**Prerequisites**

- You configured an AWS account.
- You added your AWS keys and region to your local AWS profile by running `aws configure`.
- You generated the Ignition config files for your cluster.

**Procedure**

1. Create a JSON file that contains the parameter values that the template requires:

   ```json
   [  
     {  
       "ParameterKey": "VpcCidr", 1  
       "ParameterValue": "10.0.0.0/16" 2  
     },  
     {  
       "ParameterKey": "AvailabilityZoneCount", 3  
       "ParameterValue": "1" 4  
     },  
     {  
       "ParameterKey": "SubnetBits", 5  
       "ParameterValue": "12" 6  
     }  
   ]
   ``

1. The CIDR block for the VPC.
2. Specify a CIDR block in the format `x.x.x.x/16-24`.
3. The number of availability zones to deploy the VPC in.
4. Specify an integer between 1 and 3.
5. The size of each subnet in each availability zone.
6. Specify an integer between 5 and 13, where 5 is /27 and 13 is /19.
2. Copy the template from the CloudFormation template for the VPC section of this topic and save it as a YAML file on your computer. This template describes the VPC that your cluster requires.

3. Launch the CloudFormation template to create a stack of AWS resources that represent the VPC:

   IMPORTANT

   You must enter the command on a single line.

   ```
   $ aws cloudformation create-stack --stack-name <name>  
   --template-body file://<template>.yaml  
   --parameters file://<parameters>.json
   ```

   1. `<name>` is the name for the CloudFormation stack, such as `cluster-vpc`. You need the name of this stack if you remove the cluster.

   2. `<template>` is the relative path to and name of the CloudFormation template YAML file that you saved.

   3. `<parameters>` is the relative path to and name of the CloudFormation parameters JSON file.

   Example output

   ```
   arn:aws:cloudformation:us-east-1:269333783861:stack/cluster-vpc/dbedae40-2fd3-11eb-820e-12a48460849f
   ```

4. Confirm that the template components exist:

   ```
   $ aws cloudformation describe-stacks --stack-name <name>
   ```

   After the StackStatus displays CREATE_COMPLETE, the output displays values for the following parameters. You must provide these parameter values to the other CloudFormation templates that you run to create your cluster:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>VpcId</td>
<td>The ID of your VPC.</td>
</tr>
<tr>
<td>PublicSubnetIds</td>
<td>The IDs of the new public subnets.</td>
</tr>
<tr>
<td>PrivateSubnetIds</td>
<td>The IDs of the new private subnets.</td>
</tr>
</tbody>
</table>

1.9.8.1. CloudFormation template for the VPC

You can use the following CloudFormation template to deploy the VPC that you need for your OpenShift Container Platform cluster.
Example 1.25. CloudFormation template for the VPC

AWSTemplateFormatVersion: 2010-09-09
Description: Template for Best Practice VPC with 1-3 AZs

Parameters:
VpcCidr:
  AllowedPattern: ^((\[0-9]?(\[1-9]?)?\[0-9]?)?\[2-4]?)?\[0-9]?)?\[25-0-5]?)\/(\[1-6-9]?\[2-4]?)?$
  ConstraintDescription: CIDR block parameter must be in the form x.x.x.x/16-24.
  Default: 10.0.0.0/16
  Description: CIDR block for VPC.
  Type: String
AvailabilityZoneCount:
  ConstraintDescription: "The number of availability zones. (Min: 1, Max: 3)"
  MinValue: 1
  MaxValue: 3
  Default: 1
  Description: "How many AZs to create VPC subnets for. (Min: 1, Max: 3)"
  Type: Number
SubnetBits:
  ConstraintDescription: CIDR block parameter must be in the form x.x.x.x/19-27.
  MinValue: 5
  MaxValue: 13
  Default: 12
  Description: "Size of each subnet to create within the availability zones. (Min: 5 = /27, Max: 13 = /19)"
  Type: Number

Metadata:
AWS::CloudFormation::Interface:
  ParameterGroups:
  - Label: default: "Network Configuration"
    Parameters:
      - VpcCidr
      - SubnetBits
      - Label: default: "Availability Zones"
        Parameters:
          - AvailabilityZoneCount
  ParameterLabels:
    AvailabilityZoneCount: default: "Availability Zone Count"
    VpcCidr: default: "VPC CIDR"
    SubnetBits: default: "Bits Per Subnet"

Conditions:
DoAz3: !Equals [3, !Ref AvailabilityZoneCount]
DoAz2: !Or [!Equals [2, !Ref AvailabilityZoneCount], Condition: DoAz3]

Resources:
VPC:
  Type: "AWS::EC2::VPC"
Properties:
   EnableDnsSupport: "true"
   EnableDnsHostnames: "true"
   CidrBlock: !Ref VpcCidr
PublicSubnet:
   Type: "AWS::EC2::Subnet"
   Properties:
      VpcId: !Ref VPC
      CidrBlock: !Select [0, !Cidr [!Ref VpcCidr, 6, !Ref SubnetBits]]
      AvailabilityZone: !Select
         - 0
         - Fn::GetAZs: !Ref "AWS::Region"
PublicSubnet2:
   Type: "AWS::EC2::Subnet"
   Condition: DoAz2
   Properties:
      VpcId: !Ref VPC
      CidrBlock: !Select [1, !Cidr [!Ref VpcCidr, 6, !Ref SubnetBits]]
      AvailabilityZone: !Select
         - 1
      - Fn::GetAZs: !Ref "AWS::Region"
PublicSubnet3:
   Type: "AWS::EC2::Subnet"
   Condition: DoAz3
   Properties:
      VpcId: !Ref VPC
      CidrBlock: !Select [2, !Cidr [!Ref VpcCidr, 6, !Ref SubnetBits]]
      AvailabilityZone: !Select
         - 2
      - Fn::GetAZs: !Ref "AWS::Region"
InternetGateway:
   Type: "AWS::EC2::InternetGateway"
GatewayToInternet:
   Type: "AWS::EC2::VPCGatewayAttachment"
   Properties:
      VpcId: !Ref VPC
      InternetGatewayId: !Ref InternetGateway
PublicRouteTable:
   Type: "AWS::EC2::RouteTable"
   Properties:
      VpcId: !Ref VPC
PublicRoute:
   Type: "AWS::EC2::Route"
   DependsOn: GatewayToInternet
   Properties:
      RouteTableId: !Ref PublicRouteTable
      DestinationCidrBlock: 0.0.0.0/0
      GatewayId: !Ref InternetGateway
PublicSubnetRouteTableAssociation:
   Type: "AWS::EC2::SubnetRouteTableAssociation"
   Properties:
      SubnetId: !Ref PublicSubnet
      RouteTableId: !Ref PublicRouteTable
PublicSubnetRouteTableAssociation2:
   Type: "AWS::EC2::SubnetRouteTableAssociation"
   Condition: DoAz2
Properties:
  SubnetId: !Ref PublicSubnet2
  RouteTableId: !Ref PublicRouteTable
PublicSubnetRouteTableAssociation3:
  Condition: DoAz3
  Type: "AWS::EC2::SubnetRouteTableAssociation"
  Properties:
    SubnetId: !Ref PublicSubnet3
    RouteTableId: !Ref PublicRouteTable
PrivateSubnet:
  Type: "AWS::EC2::Subnet"
  Properties:
    VpcId: !Ref VPC
    CidrBlock:
      !Select [3, !Cidr [!Ref VpcCidr, 6, !Ref SubnetBits]]
      AvailabilityZone: !Select
      - 0
      - Fn::GetAZs: !Ref "AWS::Region"
PrivateRouteTable:
  Type: "AWS::EC2::RouteTable"
  Properties:
    VpcId: !Ref VPC
PrivateSubnetRouteTableAssociation:
  Type: "AWS::EC2::SubnetRouteTableAssociation"
  Properties:
    SubnetId: !Ref PrivateSubnet
    RouteTableId: !Ref PrivateRouteTable
NAT:
  DependsOn:
  - GatewayToInternet
  Type: "AWS::EC2::NatGateway"
  Properties:
    AllocationId:
      "Fn::GetAtt":
      - EIP
      - AllocationId
    SubnetId: !Ref PublicSubnet
EIP:
  Type: "AWS::EC2::EIP"
  Properties:
    Domain: vpc
Route:
  Type: "AWS::EC2::Route"
  Properties:
    RouteTableId:
      Ref: PrivateRouteTable
    DestinationCidrBlock: 0.0.0.0/0
    NatGatewayId:
      Ref: NAT
PrivateSubnet2:
  Type: "AWS::EC2::Subnet"
  Condition: DoAz2
  Properties:
    VpcId: !Ref VPC
    CidrBlock:
      !Select [4, !Cidr [!Ref VpcCidr, 6, !Ref SubnetBits]]
      AvailabilityZone: !Select
      - 1
- Fn::GetAZs: !Ref "AWS::Region"

PrivateRouteTable2:
  Type: "AWS::EC2::RouteTable"
  Condition: DoAz2
  Properties:
    VpcId: !Ref VPC

PrivateSubnetRouteTableAssociation2:
  Type: "AWS::EC2::SubnetRouteTableAssociation"
  Condition: DoAz2
  Properties:
    SubnetId: !Ref PrivateSubnet2
    RouteTableId: !Ref PrivateRouteTable2

NAT2:
  DependsOn:
  - GatewayToInternet
  Type: "AWS::EC2::NatGateway"
  Condition: DoAz2
  Properties:
    AllocationId:
      "Fn::GetAtt":
      - EIP2
      - AllocationId
    SubnetId: !Ref PublicSubnet2

EIP2:
  Type: "AWS::EC2::EIP"
  Condition: DoAz2
  Properties:
    Domain: vpc

Route2:
  Type: "AWS::EC2::Route"
  Condition: DoAz2
  Properties:
    RouteTableId:
      Ref: PrivateRouteTable2
    DestinationCidrBlock: 0.0.0.0/0
    NatGatewayId:
      Ref: NAT2

PrivateSubnet3:
  Type: "AWS::EC2::Subnet"
  Condition: DoAz3
  Properties:
    VpcId: !Ref VPC
    CidrBlock: !Select [5, !Cidr [!Ref VpcCidr, 6, !Ref SubnetBits]]
    AvailabilityZone: !Select
      - 2
      - Fn::GetAZs: !Ref "AWS::Region"

PrivateRouteTable3:
  Type: "AWS::EC2::RouteTable"
  Condition: DoAz3
  Properties:
    VpcId: !Ref VPC

PrivateSubnetRouteTableAssociation3:
  Type: "AWS::EC2::SubnetRouteTableAssociation"
  Condition: DoAz3
  Properties:
    SubnetId: !Ref PrivateSubnet3

OpenShift Container Platform 4.6 Installing on AWS
RouteTableId: !Ref PrivateRouteTable3

NAT3:
  DependsOn:
    - GatewayToInternet
  Type: "AWS::EC2::NatGateway"
  Condition: DoAz3
  Properties:
    AllocationId:
      "Fn::GetAtt":
        - EIP3
        - AllocationId
    SubnetId: !Ref PublicSubnet3

EIP3:
  Type: "AWS::EC2::EIP"
  Condition: DoAz3
  Properties:
    Domain: vpc

Route3:
  Type: "AWS::EC2::Route"
  Condition: DoAz3
  Properties:
    RouteTableId:
      Ref: PrivateRouteTable3
    DestinationCidrBlock: 0.0.0.0/0
    NatGatewayId:
      Ref: NAT3

S3Endpoint:
  Type: AWS::EC2::VPCEndpoint
  Properties:
    PolicyDocument:
      Version: 2012-10-17
      Statement:
        - Effect: Allow
        - Principal: '*'
        - Action:
          - '*'
        - Resource:
          - '*'
    RouteTableIds:
      - !Ref PublicRouteTable
      - !Ref PrivateRouteTable
      - !If [DoAz2, !Ref PrivateRouteTable2, !Ref "AWS::NoValue"]
      - !If [DoAz3, !Ref PrivateRouteTable3, !Ref "AWS::NoValue"]
    ServiceName: !Join
      - 
        - com.amazonaws.
        - !Ref 'AWS::Region'
        - .s3
    VpcId: !Ref VPC

Outputs:
  VpcId:
    Description: ID of the new VPC.
    Value: !Ref VPC
  PublicSubnetIds:
    Description: Subnet IDs of the public subnets.
1.9.9. Creating networking and load balancing components in AWS

You must configure networking and classic or network load balancing in Amazon Web Services (AWS) that your OpenShift Container Platform cluster can use.

You can use the provided CloudFormation template and a custom parameter file to create a stack of AWS resources. The stack represents the networking and load balancing components that your OpenShift Container Platform cluster requires. The template also creates a hosted zone and subnet tags.

You can run the template multiple times within a single Virtual Private Cloud (VPC).

NOTE

If you do not use the provided CloudFormation template to create your AWS infrastructure, you must review the provided information and manually create the infrastructure. If your cluster does not initialize correctly, you might have to contact Red Hat support with your installation logs.

Prerequisites

- You configured an AWS account.
- You added your AWS keys and region to your local AWS profile by running `aws configure`.
- You generated the Ignition config files for your cluster.
- You created and configured a VPC and associated subnets in AWS.

Procedure

1. Obtain the hosted zone ID for the Route 53 base domain that you specified in the `install-config.yaml` file for your cluster. You can obtain details about your hosted zone by running the following command:

   
   ```bash
   $ aws route53 list-hosted-zones-by-name --dns-name <route53_domain>
   ```
For the `<route53_domain>`, specify the Route 53 base domain that you used when you generated the `install-config.yaml` file for the cluster.

**Example output**

mycluster.example.com. False 100
HOSTEDZONES 65F8F38E-2268-B835-E15C-AB55336FCBFA
/hostedzone/Z21IXYZABCZ2A4 mycluster.example.com. 10

In the example output, the hosted zone ID is `Z21IXYZABCZ2A4`.

2. Create a JSON file that contains the parameter values that the template requires:

```json
[
    {
        "ParameterKey": "ClusterName",  
        "ParameterValue": "mycluster"  
    },
    {
        "ParameterKey": "InfrastructureName",  
        "ParameterValue": "mycluster-<random_string>"  
    },
    {
        "ParameterKey": "HostedZoneId",  
        "ParameterValue": "<random_string>"  
    },
    {
        "ParameterKey": "HostedZoneName",  
        "ParameterValue": "example.com"  
    },
    {
        "ParameterKey": "PublicSubnets",  
        "ParameterValue": "subnet-<random_string>"  
    },
    {
        "ParameterKey": "PrivateSubnets",  
        "ParameterValue": "subnet-<random_string>"  
    },
    {
        "ParameterKey": "VpcId",  
        "ParameterValue": "vpc-<random_string>"  
    }
]
```

1. A short, representative cluster name to use for hostnames, etc.
2. Specify the cluster name that you used when you generated the `install-config.yaml` file for the cluster.
3. The name for your cluster infrastructure that is encoded in your Ignition config files for the cluster.
4. Specify the infrastructure name that you extracted from the Ignition config file metadata, which has the format `<cluster-name>-<random-string>`.

5. The Route 53 public zone ID to register the targets with.

6. Specify the Route 53 public zone ID, which has a format similar to `Z21IXYZABCZ2A4`. You can obtain this value from the AWS console.

7. The Route 53 zone to register the targets with.

8. Specify the Route 53 base domain that you used when you generated the `install-config.yaml` file for the cluster. Do not include the trailing period (.) that is displayed in the AWS console.

9. The public subnets that you created for your VPC.

10. Specify the `PublicSubnetIds` value from the output of the CloudFormation template for the VPC.

11. The private subnets that you created for your VPC.

12. Specify the `PrivateSubnetIds` value from the output of the CloudFormation template for the VPC.

13. The VPC that you created for the cluster.

14. Specify the `VpcId` value from the output of the CloudFormation template for the VPC.

3. Copy the template from the CloudFormation template for the network and load balancers section of this topic and save it as a YAML file on your computer. This template describes the networking and load balancing objects that your cluster requires.

   **IMPORTANT**
   If you are deploying your cluster to an AWS government region, you must update the `InternalApiServerRecord` in the CloudFormation template to use CNAME records. Records of type ALIAS are not supported for AWS government regions.

4. Launch the CloudFormation template to create a stack of AWS resources that provide the networking and load balancing components:

   **IMPORTANT**
   You must enter the command on a single line.

   ```bash
   $ aws cloudformation create-stack --stack-name <name>  
   --template-body file://<template>.yaml  
   --parameters file://<parameters>.json  
   --capabilities CAPABILITY_NAMED_IAM
   ```

   `<name>` is the name for the CloudFormation stack, such as `cluster-dns`. You need the name of this stack if you remove the cluster.
<template> is the relative path to and name of the CloudFormation template YAML file that you saved.

<parameters> is the relative path to and name of the CloudFormation parameters JSON file.

You must explicitly declare the `CAPABILITY_NAMED_IAM` capability because the provided template creates some `AWS::IAM::Role` resources.

**Example output**

```
arn:aws:cloudformation:us-east-1:269333783861:stack/cluster-dns/cd3e5de0-2fd4-11eb-5cf0-12be5c33a183
```

5. Confirm that the template components exist:

```
$ aws cloudformation describe-stacks --stack-name <name>
```

After the `StackStatus` displays `CREATE_COMPLETE`, the output displays values for the following parameters. You must provide these parameter values to the other CloudFormation templates that you run to create your cluster:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PrivateHostedZoneId</td>
<td>Hosted zone ID for the private DNS.</td>
</tr>
<tr>
<td>ExternalApiLoadBalancerName</td>
<td>Full name of the external API load balancer.</td>
</tr>
<tr>
<td>InternalApiLoadBalancerName</td>
<td>Full name of the internal API load balancer.</td>
</tr>
<tr>
<td>ApiServerDnsName</td>
<td>Full hostname of the API server.</td>
</tr>
<tr>
<td>RegisterNlbIpTargetsLambda</td>
<td>Lambda ARN useful to help register/deregister IP targets for these load balancers.</td>
</tr>
<tr>
<td>ExternalApiTargetGroupArn</td>
<td>ARN of external API target group.</td>
</tr>
<tr>
<td>InternalApiTargetGroupArn</td>
<td>ARN of internal API target group.</td>
</tr>
<tr>
<td>InternalServiceTargetGroupArn</td>
<td>ARN of internal service target group.</td>
</tr>
</tbody>
</table>

1.9.9.1. CloudFormation template for the network and load balancers

You can use the following CloudFormation template to deploy the networking objects and load balancers that you need for your OpenShift Container Platform cluster.

**Example 1.26. CloudFormation template for the network and load balancers**

```yaml
AWSTemplateFormatVersion: 2010-09-09
Description: Template for OpenShift Cluster Network Elements (Route53 & LBs)

Parameters:
ClusterName:
  AllowedPattern: ^([a-zA-Z][a-zA-Z0-9-]{0,26})$  
  MaxLength: 27
  MinLength: 1
  ConstraintDescription: Cluster name must be alphanumeric, start with a letter, and have a maximum of 27 characters.
  Description: A short, representative cluster name to use for host names and other identifying names.
  Type: String

InfrastructureName:
  AllowedPattern: ^([a-zA-Z][a-zA-Z0-9-]{0,26})$  
  MaxLength: 27
  MinLength: 1
  ConstraintDescription: Infrastructure name must be alphanumeric, start with a letter, and have a maximum of 27 characters.
  Description: A short, unique cluster ID used to tag cloud resources and identify items owned or used by the cluster.
  Type: String

HostedZoneId:
  Description: The Route53 public zone ID to register the targets with, such as Z21IXYZABCZ2A4.
  Type: String

HostedZoneName:
  Description: The Route53 zone to register the targets with, such as example.com. Omit the trailing period.
  Type: String

Default: "example.com"

PublicSubnets:
  Description: The internet-facing subnets.
  Type: List<AWS::EC2::Subnet::Id>

PrivateSubnets:
  Description: The internal subnets.
  Type: List<AWS::EC2::Subnet::Id>

VpcId:
  Description: The VPC-scoped resources will belong to this VPC.
  Type: AWS::EC2::VPC::Id

Metadata:
```
AWS::CloudFormation::Interface:
  ParameterGroups:
  - Label:
    default: "Cluster Information"
  Parameters:
    - ClusterName
    - InfrastructureName
  - Label:
    default: "Network Configuration"
  Parameters:
    - VpcId
    - PublicSubnets
    - PrivateSubnets
  - Label:
    default: "DNS"
  Parameters:
    - HostedZoneName
    - HostedZoneId
  ParameterLabels:
    ClusterName:
      default: "Cluster Name"
    InfrastructureName:
      default: "Infrastructure Name"
    VpcId:
      default: "VPC ID"
    PublicSubnets:
      default: "Public Subnets"
    PrivateSubnets:
      default: "Private Subnets"
    HostedZoneName:
      default: "Public Hosted Zone Name"
    HostedZoneId:
      default: "Public Hosted Zone ID"

Resources:
ExtApiElb:
  Type: AWS::ElasticLoadBalancingV2::LoadBalancer
  Properties:
    Name: !Join ["-", [!Ref InfrastructureName, "ext"]]
    IpAddressType: ipv4
    Subnets: !Ref PublicSubnets
    Type: network

IntApiElb:
  Type: AWS::ElasticLoadBalancingV2::LoadBalancer
  Properties:
    Name: !Join ["-", [!Ref InfrastructureName, "int"]]
    Scheme: internal
    IpAddressType: ipv4
    Subnets: !Ref PrivateSubnets
    Type: network

IntDns:
  Type: "AWS::Route53::HostedZone"
  Properties:
    HostedZoneConfig:
Comment: "Managed by CloudFormation"
Name: !Join [".", [!Ref ClusterName, !Ref HostedZoneName]]
HostedZoneTags:
  - Key: Name
    Value: !Join ["-", [!Ref InfrastructureName, "int"]]
  - Key: !Join ["", ["kubernetes.io/cluster/", !Ref InfrastructureName]]
    Value: "owned"
VPCs:
  - VPCId: !Ref VpcId
    VPCRegion: !Ref "AWS::Region"

ExternalApiServerRecord:
  Type: AWS::Route53::RecordSetGroup
  Properties:
    Comment: Alias record for the API server
    HostedZoneId: !Ref HostedZoneId
    RecordSets:
      - Name:
          !Join [".", [
            "api", !Ref ClusterName, !Join ["", [!Ref HostedZoneName, "."]]
          ]
        Type: A
        AliasTarget:
          HostedZoneId: !GetAtt ExtApiElb.CanonicalHostedZoneID
          DNSName: !GetAtt ExtApiElb.DNSName

InternalApiServerRecord:
  Type: AWS::Route53::RecordSetGroup
  Properties:
    Comment: Alias record for the API server
    HostedZoneId: !Ref IntDns
    RecordSets:
      - Name:
          !Join [".", [
            "api", !Ref ClusterName, !Join ["", [!Ref HostedZoneName, "."]]
          ]
        Type: A
        AliasTarget:
          HostedZoneId: !GetAtt IntApiElb.CanonicalHostedZoneID
          DNSName: !GetAtt IntApiElb.DNSName

      - Name:
          !Join [".", [
            "api-int", !Ref ClusterName, !Join ["", [!Ref HostedZoneName, "."]]
          ]
        Type: A
        AliasTarget:
          HostedZoneId: !GetAtt IntApiElb.CanonicalHostedZoneID
          DNSName: !GetAtt IntApiElb.DNSName

ExternalApiListener:
  Type: AWS::ElasticLoadBalancingV2::Listener
  Properties:
    DefaultActions:
- Type: forward
  TargetGroupArn: Ref: ExternalApiTargetGroup
  LoadBalancerArn: Ref: ExtApiElb
  Port: 6443
  Protocol: TCP

ExternalApiTargetGroup:
  Type: AWS::ElasticLoadBalancingV2::TargetGroup
  Properties:
    HealthCheckIntervalSeconds: 10
    HealthCheckPath: "/readyz"
    HealthCheckPort: 6443
    HealthCheckProtocol: HTTPS
    HealthyThresholdCount: 2
    UnhealthyThresholdCount: 2
    Port: 6443
    Protocol: TCP
    TargetType: ip
    VpcId: Ref: VpcId
    TargetGroupAttributes:
      - Key: deregistration_delay.timeout_seconds
        Value: 60

InternalApiListener:
  Type: AWS::ElasticLoadBalancingV2::Listener
  Properties:
    DefaultActions:
      - Type: forward
        TargetGroupArn: Ref: InternalApiTargetGroup
        LoadBalancerArn: Ref: IntApiElb
        Port: 6443
        Protocol: TCP

InternalApiTargetGroup:
  Type: AWS::ElasticLoadBalancingV2::TargetGroup
  Properties:
    HealthCheckIntervalSeconds: 10
    HealthCheckPath: "/readyz"
    HealthCheckPort: 6443
    HealthCheckProtocol: HTTPS
    HealthyThresholdCount: 2
    UnhealthyThresholdCount: 2
    Port: 6443
    Protocol: TCP
    TargetType: ip
    VpcId: Ref: VpcId
    TargetGroupAttributes:
      - Key: deregistration_delay.timeout_seconds
        Value: 60
InternalServiceInternalListener:
  Type: AWS::ElasticLoadBalancingV2::Listener
  Properties:
  - DefaultActions:
    - Type: forward
    TargetGroupArn: !Ref InternalServiceTargetGroup
  LoadBalancerArn: !Ref IntApiElb
  Port: 22623
  Protocol: TCP

InternalServiceTargetGroup:
  Type: AWS::ElasticLoadBalancingV2::TargetGroup
  Properties:
  - HealthCheckIntervalSeconds: 10
    HealthCheckPath: /healthz
    HealthCheckPort: 22623
    HealthCheckProtocol: HTTPS
    HealthyThresholdCount: 2
    UnhealthyThresholdCount: 2
    Port: 22623
    Protocol: TCP
    TargetType: ip
    VpcId: !Ref VpcId
    TargetGroupAttributes:
    - Key: deregistration_delay.timeout_seconds
      Value: 60

RegisterTargetLambdaIamRole:
  Type: AWS::IAM::Role
  Properties:
  - RoleName: !Join ["-", [!Ref InfrastructureName, "nlb", "lambda", "role"]]
  AssumeRolePolicyDocument:
    Version: "2012-10-17"
    Statement:
    - Effect: "Allow"
      Principal:
        Service:
        - "lambda.amazonaws.com"
      Action:
      - "sts:AssumeRole"
    Path: "/"
  Policies:
  - PolicyName: !Join ["-", [!Ref InfrastructureName, "master", "policy"]]
    PolicyDocument:
      Version: "2012-10-17"
      Statement:
      - Effect: "Allow"
        Action:
        - "elasticloadbalancing:RegisterTargets",
        "elasticloadbalancing:DeregisterTargets",
      Resource: !Ref InternalApiTargetGroup
- Effect: "Allow"
  Action:
  ```
  ["elasticloadbalancing:RegisterTargets",
   "elasticloadbalancing:DeregisterTargets",
  ]
  ```
  Resource: !Ref InternalServiceTargetGroup
- Effect: "Allow"
  Action:
  ```
  ["elasticloadbalancing:RegisterTargets",
   "elasticloadbalancing:DeregisterTargets",
  ]
  ```
  Resource: !Ref ExternalApiTargetGroup

RegisterNlbIpTargets:
Type: "AWS::Lambda::Function"
Properties:
  Handler: "index.handler"
  Role:
    Fn::GetAtt:
    - "RegisterTargetLambdaIamRole"
    - "Arn"
  Code:
    ZipFile: |
      import json
      import boto3
      import cfnresponse
      def handler(event, context):
        elb = boto3.client('elbv2')
        if event['RequestType'] == 'Delete':
          elb.deregister_targets(TargetGroupArn=event['ResourceProperties']['TargetArn'], Targets=[{'Id': event['ResourceProperties']['TargetIp']}])
        elif event['RequestType'] == 'Create':
          elb.register_targets(TargetGroupArn=event['ResourceProperties']['TargetArn'], Targets=[{'Id': event['ResourceProperties']['TargetIp']}])
        responseData = {}
        cfnresponse.send(event, context, cfnresponse.SUCCESS, responseData,
        event['ResourceProperties']['TargetArn']+event['ResourceProperties']['TargetIp'])
  Runtime: "python3.7"
  Timeout: 120

RegisterSubnetTagsLambdaIamRole:
Type: AWS::IAM::Role
Properties:
  RoleName: !Join ["-", [!Ref InfrastructureName, "subnet-tags-lambda-role"]]
  AssumeRolePolicyDocument:
    Version: "2012-10-17"
    Statement:
      - Effect: "Allow"
        Principal:
          Service:
            - "lambda.amazonaws.com"
        Action:
          - "sts:AssumeRole"
    Path: "/"
Policies:
- PolicyName: !Join ["-", [!Ref InfrastructureName, "subnet-tagging-policy"]]

PolicyDocument:
  Version: "2012-10-17"
  Statement:
  - Effect: "Allow"
    Action:
    - "ec2:DeleteTags",
    - "ec2:CreateTags"
    Resource: "arn:aws:ec2:*:*:subnet/*"
  - Effect: "Allow"
    Action:
    - "ec2:DescribeSubnets",
    - "ec2:DescribeTags"
    Resource: "*

RegisterSubnetTags:
  Type: "AWS::Lambda::Function"
  Properties:
    Handler: "index.handler"
    Role:
      Fn::GetAtt:
      - "RegisterSubnetTagsLambdaIamRole"
      - "Arn"
  Code:
    ZipFile: |
      import json
      import boto3
      import cfnresponse
      def handler(event, context):
        ec2_client = boto3.client('ec2')
        if event['RequestType'] == 'Delete':
          for subnet_id in event['ResourceProperties']['Subnets']:
            ec2_client.delete_tags(Resources=[subnet_id], Tags=[{'Key': 'kubernetes.io/cluster/' + event['ResourceProperties']['InfrastructureName']}]));
        elif event['RequestType'] == 'Create':
          for subnet_id in event['ResourceProperties']['Subnets']:
            ec2_client.create_tags(Resources=[subnet_id], Tags=[{'Key': 'kubernetes.io/cluster/' + event['ResourceProperties']['InfrastructureName'], 'Value': 'shared'}]);
        responseData = {}
        cfnresponse.send(event, context, cfnresponse.SUCCESS, responseData,
        event['ResourceProperties']['InfrastructureName']+event['ResourceProperties']['Subnets'][0])
  Runtime: "python3.7"
  Timeout: 120

RegisterPublicSubnetTags:
  Type: Custom::SubnetRegister
  Properties:
    ServiceToken: !GetAtt RegisterSubnetTags.Arn
    InfrastructureName: !Ref InfrastructureName
    Subnets: !Ref PublicSubnets
If you are deploying your cluster to an AWS government region, you must update the `InternalApiServerRecord` to use CNAME records. Records of type ALIAS are not supported for AWS government regions. For example:

```
Type: CNAME
TTL: 10
ResourceRecords:
  - !GetAtt IntApiElb.DNSName
```

Additional resources

- See [Listing public hosted zones](#) in the AWS documentation for more information about listing public hosted zones.

1.9.10. Creating security group and roles in AWS
You must create security groups and roles in Amazon Web Services (AWS) for your OpenShift Container Platform cluster to use.

You can use the provided CloudFormation template and a custom parameter file to create a stack of AWS resources. The stack represents the security groups and roles that your OpenShift Container Platform cluster requires.

**NOTE**

If you do not use the provided CloudFormation template to create your AWS infrastructure, you must review the provided information and manually create the infrastructure. If your cluster does not initialize correctly, you might have to contact Red Hat support with your installation logs.

**Prerequisites**

- You configured an AWS account.
- You added your AWS keys and region to your local AWS profile by running `aws configure`.
- You generated the Ignition config files for your cluster.
- You created and configured a VPC and associated subnets in AWS.

**Procedure**

1. Create a JSON file that contains the parameter values that the template requires:

   ```json
   [
   
   
   
   "ParameterKey": "InfrastructureName", 1
   "ParameterValue": "mycluster-<random_string>" 2
   
   
   "ParameterKey": "VpcCidr", 3
   "ParameterValue": "10.0.0.0/16" 4
   
   
   "ParameterKey": "PrivateSubnets", 5
   "ParameterValue": "subnet-<random_string>" 6
   
   
   "ParameterKey": "VpcId", 7
   "ParameterValue": "vpc-<random_string>" 8
   
   ]
   
   1 The name for your cluster infrastructure that is encoded in your Ignition config files for the cluster.
   2 Specify the infrastructure name that you extracted from the Ignition config file metadata, which has the format `<cluster-name-<random-string>`.  
   3 The CIDR block for the VPC.
4. Specify the CIDR block parameter that you used for the VPC that you defined in the form x.x.x.x/16-24.

5. The private subnets that you created for your VPC.

6. Specify the `PrivateSubnetsId` value from the output of the CloudFormation template for the VPC.

7. The VPC that you created for the cluster.

8. Specify the `VpcId` value from the output of the CloudFormation template for the VPC.

2. Copy the template from the **CloudFormation template for security objects** section of this topic and save it as a YAML file on your computer. This template describes the security groups and roles that your cluster requires.

3. Launch the CloudFormation template to create a stack of AWS resources that represent the security groups and roles:

   **IMPORTANT**
   
   You must enter the command on a single line.

   ```
   $ aws cloudformation create-stack --stack-name <name> 1
   --template-body file://<template>.yaml 2
   --parameters file://<parameters>.json 3
   --capabilities CAPABILITY_NAMED_IAM 4
   
   1 <name> is the name for the CloudFormation stack, such as `cluster-sec`. You need the name of this stack if you remove the cluster.

   2 <template> is the relative path to and name of the CloudFormation template YAML file that you saved.

   3 <parameters> is the relative path to and name of the CloudFormation parameters JSON file.

   4 You must explicitly declare the CAPABILITY_NAMED_IAM capability because the provided template creates some AWS::IAM::Role and AWS::IAM::InstanceProfile resources.
   ```

   **Example output**

   ```
   arn:aws:cloudformation:us-east-1:269333783861:stack/cluster-sec/03bd4210-2ed7-11eb-6d7a-13fc0b61e9db
   ```

4. Confirm that the template components exist:

   ```
   $ aws cloudformation describe-stacks --stack-name <name>
   ```
After the **StackStatus** displays **CREATE_COMPLETE**, the output displays values for the following parameters. You must provide these parameter values to the other CloudFormation templates that you run to create your cluster:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MasterSecurityGroupId</strong></td>
<td>Master Security Group ID</td>
</tr>
<tr>
<td><strong>WorkerSecurityGroupId</strong></td>
<td>Worker Security Group ID</td>
</tr>
<tr>
<td><strong>MasterInstanceProfile</strong></td>
<td>Master IAM Instance Profile</td>
</tr>
<tr>
<td><strong>WorkerInstanceProfile</strong></td>
<td>Worker IAM Instance Profile</td>
</tr>
</tbody>
</table>

### 1.9.10.1. CloudFormation template for security objects

You can use the following CloudFormation template to deploy the security objects that you need for your OpenShift Container Platform cluster.

```
AWSTemplateFormatVersion: 2010-09-09
Description: Template for OpenShift Cluster Security Elements (Security Groups & IAM)

Parameters:

- **InfrastructureName**
  - AllowedPattern: `^[a-zA-Z][a-zA-Z0-9-]{0,26}$`
  - MaxLength: 27
  - MinLength: 1
  - ConstraintDescription: Infrastructure name must be alphanumeric, start with a letter, and have a maximum of 27 characters.
  - Description: A short, unique cluster ID used to tag cloud resources and identify items owned or used by the cluster.
  - Type: String

- **VpcCidr**
  - AllowedPattern: `^((\d{1-3})\.(\d{1-3})\.(\d{1-3})\.(\d{1-3})|\d{1-4})((\d{1-3})|0)$/16-24$`
  - ConstraintDescription: CIDR block parameter must be in the form x.x.x.x/16-24.
  - Default: 10.0.0.0/16
  - Description: CIDR block for VPC.
  - Type: String

- **VpcId**
  - Description: The VPC-scoped resources will belong to this VPC.
  - Type: AWS::EC2::VPC::Id

- **PrivateSubnets**
  - Description: The internal subnets.
  - Type: List<AWS::EC2::Subnet::Id>
```
Metadata:
AWS::CloudFormation::Interface:
  ParameterGroups:
  - Label:
    - default: "Cluster Information"
  Parameters:
    - InfrastructureName
  - Label:
    - default: "Network Configuration"
  Parameters:
    - VpcId
    - VpcCidr
    - PrivateSubnets
  ParameterLabels:
    InfrastructureName:
    - default: "Infrastructure Name"
    VpcId:
    - default: "VPC ID"
    VpcCidr:
    - default: "VPC CIDR"
    PrivateSubnets:
    - default: "Private Subnets"

Resources:
MasterSecurityGroup:
  Type: AWS::EC2::SecurityGroup
  Properties:
    GroupDescription: Cluster Master Security Group
    SecurityGroupIngress:
    - IpProtocol: icmp
      FromPort: 0
      ToPort: 0
      CidrIp: !Ref VpcCidr
    - IpProtocol: tcp
      FromPort: 22
      ToPort: 22
      CidrIp: !Ref VpcCidr
    - IpProtocol: tcp
      FromPort: 6443
      ToPort: 6443
      CidrIp: !Ref VpcCidr
    - IpProtocol: tcp
      FromPort: 22623
      ToPort: 22623
      CidrIp: !Ref VpcCidr
    VpcId: !Ref VpcId

WorkerSecurityGroup:
  Type: AWS::EC2::SecurityGroup
  Properties:
    GroupDescription: Cluster Worker Security Group
    SecurityGroupIngress:
    - IpProtocol: icmp
      FromPort: 0
      ToPort: 0
CidrIp: !Ref VpcCidr
  IpProtocol: tcp
  FromPort: 22
  ToPort: 22
CidrIp: !Ref VpcCidr
VpcId: !Ref VpcId

MasterIngressEtcd:
  Type: AWS::EC2::SecurityGroupIngress
  Properties:
    GroupId: !GetAtt MasterSecurityGroup.GroupId
    SourceSecurityGroupId: !GetAtt MasterSecurityGroup.GroupId
    Description: etcd
    FromPort: 2379
    ToPort: 2380
    IpProtocol: tcp

MasterIngressVxlan:
  Type: AWS::EC2::SecurityGroupIngress
  Properties:
    GroupId: !GetAtt MasterSecurityGroup.GroupId
    SourceSecurityGroupId: !GetAtt WorkerSecurityGroup.GroupId
    Description: Vxlan packets
    FromPort: 4789
    ToPort: 4789
    IpProtocol: udp

MasterIngressWorkerVxlan:
  Type: AWS::EC2::SecurityGroupIngress
  Properties:
    GroupId: !GetAtt MasterSecurityGroup.GroupId
    SourceSecurityGroupId: !GetAtt WorkerSecurityGroup.GroupId
    Description: Vxlan packets
    FromPort: 4789
    ToPort: 4789
    IpProtocol: udp

MasterIngressGeneve:
  Type: AWS::EC2::SecurityGroupIngress
  Properties:
    GroupId: !GetAtt MasterSecurityGroup.GroupId
    SourceSecurityGroupId: !GetAtt MasterSecurityGroup.GroupId
    Description: Geneve packets
    FromPort: 6081
    ToPort: 6081
    IpProtocol: udp

MasterIngressWorkerGeneve:
  Type: AWS::EC2::SecurityGroupIngress
  Properties:
    GroupId: !GetAtt MasterSecurityGroup.GroupId
    SourceSecurityGroupId: !GetAtt WorkerSecurityGroup.GroupId
    Description: Geneve packets
    FromPort: 6081
    ToPort: 6081
    IpProtocol: udp
MasterIngressInternal:
Type: AWS::EC2::SecurityGroupIngress
Properties:
  GroupId: !GetAtt MasterSecurityGroup.GroupId
  SourceSecurityGroupId: !GetAtt MasterSecurityGroup.GroupId
  Description: Internal cluster communication
  FromPort: 9000
  ToPort: 9999
  IpProtocol: tcp

MasterIngressWorkerInternal:
Type: AWS::EC2::SecurityGroupIngress
Properties:
  GroupId: !GetAtt MasterSecurityGroup.GroupId
  SourceSecurityGroupId: !GetAtt WorkerSecurityGroup.GroupId
  Description: Internal cluster communication
  FromPort: 9000
  ToPort: 9999
  IpProtocol: tcp

MasterIngressInternalUDP:
Type: AWS::EC2::SecurityGroupIngress
Properties:
  GroupId: !GetAtt MasterSecurityGroup.GroupId
  SourceSecurityGroupId: !GetAtt MasterSecurityGroup.GroupId
  Description: Internal cluster communication
  FromPort: 9000
  ToPort: 9999
  IpProtocol: udp

MasterIngressWorkerInternalUDP:
Type: AWS::EC2::SecurityGroupIngress
Properties:
  GroupId: !GetAtt MasterSecurityGroup.GroupId
  SourceSecurityGroupId: !GetAtt WorkerSecurityGroup.GroupId
  Description: Internal cluster communication
  FromPort: 9000
  ToPort: 9999
  IpProtocol: udp

MasterIngressKube:
Type: AWS::EC2::SecurityGroupIngress
Properties:
  GroupId: !GetAtt MasterSecurityGroup.GroupId
  SourceSecurityGroupId: !GetAtt MasterSecurityGroup.GroupId
  Description: Kubernetes kubelet, scheduler and controller manager
  FromPort: 10250
  ToPort: 10259
  IpProtocol: tcp

MasterIngressWorkerKube:
Type: AWS::EC2::SecurityGroupIngress
Properties:
  GroupId: !GetAtt MasterSecurityGroup.GroupId
  SourceSecurityGroupId: !GetAtt WorkerSecurityGroup.GroupId

CHAPTER 1. INSTALLING ON AWS
Description: Kubernetes kubelet, scheduler and controller manager
FromPort: 10250
ToPort: 10259
IpProtocol: tcp

MasterIngressIngressServices:
Type: AWS::EC2::SecurityGroupIngress
Properties:
  GroupId: !GetAtt MasterSecurityGroup.GroupId
  SourceSecurityGroupId: !GetAtt MasterSecurityGroup.GroupId
Description: Kubernetes ingress services
FromPort: 30000
ToPort: 32767
IpProtocol: tcp

MasterIngressWorkerIngressServices:
Type: AWS::EC2::SecurityGroupIngress
Properties:
  GroupId: !GetAtt MasterSecurityGroup.GroupId
  SourceSecurityGroupId: !GetAtt WorkerSecurityGroup.GroupId
Description: Kubernetes ingress services
FromPort: 30000
ToPort: 32767
IpProtocol: tcp

MasterIngressIngressServicesUDP:
Type: AWS::EC2::SecurityGroupIngress
Properties:
  GroupId: !GetAtt MasterSecurityGroup.GroupId
  SourceSecurityGroupId: !GetAtt MasterSecurityGroup.GroupId
Description: Kubernetes ingress services
FromPort: 30000
ToPort: 32767
IpProtocol: udp

MasterIngressWorkerIngressServicesUDP:
Type: AWS::EC2::SecurityGroupIngress
Properties:
  GroupId: !GetAtt MasterSecurityGroup.GroupId
  SourceSecurityGroupId: !GetAtt WorkerSecurityGroup.GroupId
Description: Kubernetes ingress services
FromPort: 30000
ToPort: 32767
IpProtocol: udp

WorkerIngressVxlan:
Type: AWS::EC2::SecurityGroupIngress
Properties:
  GroupId: !GetAtt WorkerSecurityGroup.GroupId
  SourceSecurityGroupId: !GetAtt WorkerSecurityGroup.GroupId
Description: Vxlan packets
FromPort: 4789
ToPort: 4789
IpProtocol: udp

WorkerIngressMasterVxlan:
Type: AWS::EC2::SecurityGroupIngress
Properties:
  GroupId: !GetAtt WorkerSecurityGroup.GroupId
  SourceSecurityGroupId: !GetAtt MasterSecurityGroup.GroupId
  Description: Vxlan packets
  FromPort: 4789
  ToPort: 4789
  IpProtocol: udp

WorkerIngressGeneve:
Type: AWS::EC2::SecurityGroupIngress
Properties:
  GroupId: !GetAtt WorkerSecurityGroup.GroupId
  SourceSecurityGroupId: !GetAtt WorkerSecurityGroup.GroupId
  Description: Geneve packets
  FromPort: 6081
  ToPort: 6081
  IpProtocol: udp

WorkerIngressMasterGeneve:
Type: AWS::EC2::SecurityGroupIngress
Properties:
  GroupId: !GetAtt WorkerSecurityGroup.GroupId
  SourceSecurityGroupId: !GetAtt MasterSecurityGroup.GroupId
  Description: Geneve packets
  FromPort: 6081
  ToPort: 6081
  IpProtocol: udp

WorkerIngressInternal:
Type: AWS::EC2::SecurityGroupIngress
Properties:
  GroupId: !GetAtt WorkerSecurityGroup.GroupId
  SourceSecurityGroupId: !GetAtt WorkerSecurityGroup.GroupId
  Description: Internal cluster communication
  FromPort: 9000
  ToPort: 9999
  IpProtocol: tcp

WorkerIngressMasterInternal:
Type: AWS::EC2::SecurityGroupIngress
Properties:
  GroupId: !GetAtt WorkerSecurityGroup.GroupId
  SourceSecurityGroupId: !GetAtt MasterSecurityGroup.GroupId
  Description: Internal cluster communication
  FromPort: 9000
  ToPort: 9999
  IpProtocol: tcp

WorkerIngressInternalUDP:
Type: AWS::EC2::SecurityGroupIngress
Properties:
  GroupId: !GetAtt WorkerSecurityGroup.GroupId
  SourceSecurityGroupId: !GetAtt WorkerSecurityGroup.GroupId
  Description: Internal cluster communication
  FromPort: 9000
WorkerIngressMasterInternalUDP:
Type: AWS::EC2::SecurityGroupIngress
Properties:
  GroupId: !GetAtt WorkerSecurityGroup.GroupId
  SourceSecurityGroupId: !GetAtt MasterSecurityGroup.GroupId
  Description: Internal cluster communication
  FromPort: 9000
  ToPort: 9999
  IpProtocol: udp

WorkerIngressKube:
Type: AWS::EC2::SecurityGroupIngress
Properties:
  GroupId: !GetAtt WorkerSecurityGroup.GroupId
  SourceSecurityGroupId: !GetAtt WorkerSecurityGroup.GroupId
  Description: Kubernetes secure kubelet port
  FromPort: 10250
  ToPort: 10250
  IpProtocol: tcp

WorkerIngressWorkerKube:
Type: AWS::EC2::SecurityGroupIngress
Properties:
  GroupId: !GetAtt WorkerSecurityGroup.GroupId
  SourceSecurityGroupId: !GetAtt MasterSecurityGroup.GroupId
  Description: Internal Kubernetes communication
  FromPort: 10250
  ToPort: 10250
  IpProtocol: tcp

WorkerIngressIngressServices:
Type: AWS::EC2::SecurityGroupIngress
Properties:
  GroupId: !GetAtt WorkerSecurityGroup.GroupId
  SourceSecurityGroupId: !GetAtt WorkerSecurityGroup.GroupId
  Description: Kubernetes ingress services
  FromPort: 30000
  ToPort: 32767
  IpProtocol: tcp

WorkerIngressMasterIngressServices:
Type: AWS::EC2::SecurityGroupIngress
Properties:
  GroupId: !GetAtt WorkerSecurityGroup.GroupId
  SourceSecurityGroupId: !GetAtt MasterSecurityGroup.GroupId
  Description: Kubernetes ingress services
  FromPort: 30000
  ToPort: 32767
  IpProtocol: tcp

WorkerIngressIngressServicesUDP:
Type: AWS::EC2::SecurityGroupIngress
Properties:
WorkerIngressMasterIngressServicesUDP:
Type: AWS::EC2::SecurityGroupIngress
Properties:
  GroupId: !GetAtt WorkerSecurityGroup.GroupId
  SourceSecurityGroupId: !GetAtt MasterSecurityGroup.GroupId
  Description: Kubernetes ingress services
  FromPort: 30000
  ToPort: 32767
  IpProtocol: udp

MasterIamRole:
Type: AWS::IAM::Role
Properties:
  AssumeRolePolicyDocument:
    Version: "2012-10-17"
    Statement:
      - Effect: "Allow"
        Principal:
          Service:
            - "ec2.amazonaws.com"
        Action:
          - "sts:AssumeRole"
    Policies:
      - PolicyName: !Join ["-", [!Ref InfrastructureName, "master", "policy"]]
        PolicyDocument:
          Version: "2012-10-17"
          Statement:
            - Effect: "Allow"
              Action:
                "ec2:AttachVolume"
                "ec2:AuthorizeSecurityGroupIngress"
                "ec2:CreateSecurityGroup"
                "ec2:CreateTags"
                "ec2:CreateVolume"
                "ec2:DeleteSecurityGroup"
                "ec2:DeleteVolume"
                "ec2:Describe**"
                "ec2:DetachVolume"
                "ec2:ModifyInstanceAttribute"
                "ec2:ModifyVolume"
                "ec2:RevokeSecurityGroupIngress"
                "elasticloadbalancing:AddTags"
                "elasticloadbalancing:AttachLoadBalancerToSubnets"
                "elasticloadbalancing:ApplySecurityGroupsToLoadBalancer"
                "elasticloadbalancing:CreateListener"
                "elasticloadbalancing:CreateLoadBalancer"
                "elasticloadbalancing:CreateLoadBalancerPolicy"
                "elasticloadbalancing:CreateLoadBalancerListeners"
                "elasticloadbalancing:CreateTargetGroup"
- "elasticloadbalancing:ConfigureHealthCheck"
- "elasticloadbalancing:DeleteListener"
- "elasticloadbalancing:DeleteLoadBalancer"
- "elasticloadbalancing:DeleteLoadBalancerListeners"
- "elasticloadbalancing:DeleteTargetGroup"
- "elasticloadbalancing:DeregisterInstancesFromLoadBalancer"
- "elasticloadbalancing:DeregisterTargets"
- "elasticloadbalancing:Describe"
- "elasticloadbalancing:DetachLoadBalancerFromSubnets"
- "elasticloadbalancing:ModifyListener"
- "elasticloadbalancing:ModifyLoadBalancerAttributes"
- "elasticloadbalancing:ModifyTargetGroup"
- "elasticloadbalancing:ModifyTargetGroupAttributes"
- "elasticloadbalancing:RegisterInstancesWithLoadBalancer"
- "elasticloadbalancing:RegisterTargets"
- "elasticloadbalancing:SetLoadBalancerPoliciesForBackendServer"
- "elasticloadbalancing:SetLoadBalancerPoliciesOfListener"
- "kms:DescribeKey"

Resource: "**"

**MasterInstanceProfile**:
- Type: "AWS::IAM::InstanceProfile"
- Properties:
  - Roles:
    - Ref: "MasterIamRole"

**WorkerIamRole**:
- Type: AWS::IAM::Role
- Properties:
  - AssumeRolePolicyDocument:
    - Version: "2012-10-17"
    - Statement:
      - Effect: "Allow"
      - Principal:
        - Service:
          - "ec2.amazonaws.com"
      - Action:
        - "sts:AssumeRole"
      - Policies:
        - PolicyName: !Join ["-", [!Ref InfrastructureName, "worker", "policy"]]
          - PolicyDocument:
            - Version: "2012-10-17"
            - Statement:
              - Effect: "Allow"
              - Action:
                - "ec2:DescribeInstances"
                - "ec2:DescribeRegions"
            - Resource: "**"

**WorkerInstanceProfile**:
- Type: "AWS::IAM::InstanceProfile"
- Properties:
  - Roles:
    - Ref: "WorkerIamRole"

**Outputs:**
1.9.11. RHCOS AMIs for the AWS infrastructure

Red Hat provides Red Hat Enterprise Linux CoreOS (RHCOS) AMIs valid for the various Amazon Web Services (AWS) zones you can specify for your OpenShift Container Platform nodes.

NOTE

You can also install to regions that do not have a RHCOS AMI published by importing your own AMI.

Table 1.27. RHCOS AMIs

<table>
<thead>
<tr>
<th>AWS zone</th>
<th>AWS AMI</th>
</tr>
</thead>
<tbody>
<tr>
<td>af-south-1</td>
<td>ami-09921c9c1c36e695c</td>
</tr>
<tr>
<td>ap-east-1</td>
<td>ami-01ee8446e9af6b197</td>
</tr>
<tr>
<td>ap-northeast-1</td>
<td>ami-04e5b5722a55846ea</td>
</tr>
<tr>
<td>ap-northeast-2</td>
<td>ami-0fdc25c8a0273a742</td>
</tr>
<tr>
<td>ap-south-1</td>
<td>ami-09e3deb397cc526a8</td>
</tr>
<tr>
<td>ap-southeast-1</td>
<td>ami-0630e03f75e02eec4</td>
</tr>
<tr>
<td>ap-southeast-2</td>
<td>ami-069450613262ba03c</td>
</tr>
<tr>
<td>ca-central-1</td>
<td>ami-012518cddb3057dfd</td>
</tr>
<tr>
<td>eu-central-1</td>
<td>ami-0bd7175ff5b1ae0f0c</td>
</tr>
<tr>
<td>eu-north-1</td>
<td>ami-06c9ec42d0a839ad2</td>
</tr>
</tbody>
</table>
### 1.9.12. Creating the bootstrap node in AWS

You must create the bootstrap node in Amazon Web Services (AWS) to use during OpenShift Container Platform cluster initialization.

You can use the provided CloudFormation template and a custom parameter file to create a stack of AWS resources. The stack represents the bootstrap node that your OpenShift Container Platform installation requires.

**NOTE**

If you do not use the provided CloudFormation template to create your bootstrap node, you must review the provided information and manually create the infrastructure. If your cluster does not initialize correctly, you might have to contact Red Hat support with your installation logs.

**Prerequisites**

- You configured an AWS account.
- You added your AWS keys and region to your local AWS profile by running `aws configure`.
- You generated the Ignition config files for your cluster.
- You created and configured a VPC and associated subnets in AWS.
- You created and configured DNS, load balancers, and listeners in AWS.

<table>
<thead>
<tr>
<th>AWS zone</th>
<th>AWS AMI</th>
</tr>
</thead>
<tbody>
<tr>
<td>eu-south-1</td>
<td>ami-0614d7440a0363d71</td>
</tr>
<tr>
<td>eu-west-1</td>
<td>ami-01b89df58b5d4d5fa</td>
</tr>
<tr>
<td>eu-west-2</td>
<td>ami-06f6e31ddd554f89d</td>
</tr>
<tr>
<td>eu-west-3</td>
<td>ami-0dc82e2517ded15a1</td>
</tr>
<tr>
<td>me-south-1</td>
<td>ami-07d181e3aa0f76067</td>
</tr>
<tr>
<td>sa-east-1</td>
<td>ami-0cd44e6dd20e6c7fa</td>
</tr>
<tr>
<td>us-east-1</td>
<td>ami-04a16d506e5b0e246</td>
</tr>
<tr>
<td>us-east-2</td>
<td>ami-0a1f868ad58ea59a7</td>
</tr>
<tr>
<td>us-west-1</td>
<td>ami-0a65d76e3a6f6622f</td>
</tr>
<tr>
<td>us-west-2</td>
<td>ami-0dd9008abadc519f1</td>
</tr>
</tbody>
</table>
You created the security groups and roles required for your cluster in AWS.

Procedure

1. Provide a location to serve the `bootstrap.ign` Ignition config file to your cluster. This file is located in your installation directory. One way to do this is to create an S3 bucket in your cluster’s region and upload the Ignition config file to it.

   **IMPORTANT**

   The provided CloudFormation Template assumes that the Ignition config files for your cluster are served from an S3 bucket. If you choose to serve the files from another location, you must modify the templates.

   **IMPORTANT**

   If you are deploying to a region that has endpoints that differ from the AWS SDK, or you are providing your own custom endpoints, you must use a presigned URL for your S3 bucket instead of the `s3://` schema.

   **NOTE**

   The bootstrap Ignition config file does contain secrets, like X.509 keys. The following steps provide basic security for the S3 bucket. To provide additional security, you can enable an S3 bucket policy to allow only certain users, such as the OpenShift IAM user, to access objects that the bucket contains. You can avoid S3 entirely and serve your bootstrap Ignition config file from any address that the bootstrap machine can reach.

   a. Create the bucket:

      ```bash
      $ aws s3 mb s3://<cluster-name>-infra
      ```

   b. Upload the `bootstrap.ign` Ignition config file to the bucket:

      ```bash
      $ aws s3 cp <installation_directory>/bootstrap.ign s3://<cluster-name>-infra/bootstrap.ign
      ```

      For `<installation_directory>`, specify the path to the directory that you stored the installation files in.

   c. Verify that the file uploaded:

      ```bash
      $ aws s3 ls s3://<cluster-name>-infra/
      ```

      **Example output**

      ```
      2019-04-03 16:15:16  314878 bootstrap.ign
      ```
2. Create a JSON file that contains the parameter values that the template requires:

```
[
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  

```
The name for your cluster infrastructure that is encoded in your Ignition config files for the cluster.

Specify the infrastructure name that you extracted from the Ignition config file metadata, which has the format `<cluster-name>-<random-string>`.

Current Red Hat Enterprise Linux CoreOS (RHCOS) AMI to use for the bootstrap node.

Specify a valid `AWS::EC2::Image::Id` value.

CIDR block to allow SSH access to the bootstrap node.

Specify a CIDR block in the format `x.x.x.x/16-24`.

The public subnet that is associated with your VPC to launch the bootstrap node into.

Specify the `PublicSubnetIds` value from the output of the CloudFormation template for the VPC.

The master security group ID (for registering temporary rules)

Specify the `MasterSecurityGroupId` value from the output of the CloudFormation template for the security group and roles.

The VPC created resources will belong to.

Specify the `VpcId` value from the output of the CloudFormation template for the VPC.

Location to fetch bootstrap Ignition config file from.

Specify the S3 bucket and file name in the form `s3://<bucket_name>/bootstrap.ign`.

Whether or not to register a network load balancer (NLB).

Specify `yes` or `no`. If you specify `yes`, you must provide a Lambda Amazon Resource Name (ARN) value.

The ARN for NLB IP target registration lambda group.

Specify the `RegisterNlbIpTargetsLambda` value from the output of the CloudFormation template for DNS and load balancing. Use `arn:aws-us-gov` if deploying the cluster to an AWS GovCloud region.

The ARN for external API load balancer target group.

Specify the `ExternalApiTargetGroupArn` value from the output of the CloudFormation template for DNS and load balancing. Use `arn:aws-us-gov` if deploying the cluster to an AWS GovCloud region.

The ARN for internal API load balancer target group.

Specify the `InternalApiTargetGroupArn` value from the output of the CloudFormation template for DNS and load balancing. Use `arn:aws-us-gov` if deploying the cluster to an AWS GovCloud region.
The ARN for internal service load balancer target group.

Specify the InternalServiceTargetGroupArn value from the output of the CloudFormation template for DNS and load balancing. Use arn:aws-us-gov if deploying the cluster to an AWS GovCloud region.

3. Copy the template from the CloudFormation template for the bootstrap machine section of this topic and save it as a YAML file on your computer. This template describes the bootstrap machine that your cluster requires.

4. Launch the CloudFormation template to create a stack of AWS resources that represent the bootstrap node:

   IMPORTANT
   You must enter the command on a single line.

   $ aws cloudformation create-stack --stack-name <name> 1
   --template-body file:///<template>.yaml 2
   --parameters file:///<parameters>.json 3
   --capabilities CAPABILITY_NAMED_IAM 4

   1 <name> is the name for the CloudFormation stack, such as cluster-bootstrap. You need the name of this stack if you remove the cluster.
   2 <template> is the relative path to and name of the CloudFormation template YAML file that you saved.
   3 <parameters> is the relative path to and name of the CloudFormation parameters JSON file.
   4 You must explicitly declare the CAPABILITY_NAMED_IAM capability because the provided template creates some AWS::IAM::Role and AWS::IAM::InstanceProfile resources.

Example output

   arn:aws:cloudformation:us-east-1:269333783861:stack/cluster-bootstrap/12944486-2add-11eb-9dee-12dace8e3a83

5. Confirm that the template components exist:

   $ aws cloudformation describe-stacks --stack-name <name>

   After the StackStatus displays CREATE_COMPLETE, the output displays values for the following parameters. You must provide these parameter values to the other CloudFormation templates that you run to create your cluster:
### Bootstrap InstanceId
The bootstrap Instance ID.

### Bootstrap PublicIp
The bootstrap node public IP address.

### Bootstrap PrivateIp
The bootstrap node private IP address.

#### 1.9.12.1. CloudFormation template for the bootstrap machine
You can use the following CloudFormation template to deploy the bootstrap machine that you need for your OpenShift Container Platform cluster.

**Example 1.28. CloudFormation template for the bootstrap machine**

```yaml
AWSTemplateFormatVersion: 2010-09-09
Description: Template for OpenShift Cluster Bootstrap (EC2 Instance, Security Groups and IAM)

Parameters:
InfrastructureName:
  AllowedPattern: ^([a-zA-Z][a-zA-Z0-9-]{0,26})$
  MaxLength: 27
  MinLength: 1
  ConstraintDescription: Infrastructure name must be alphanumeric, start with a letter, and have a maximum of 27 characters.
  Description: A short, unique cluster ID used to tag cloud resources and identify items owned or used by the cluster.
  Type: String
RhcosAmi:
  Description: Current Red Hat Enterprise Linux CoreOS AMI to use for bootstrap.
  Type: AWS::EC2::Image::Id
AllowedBootstrapSshCidr:
  AllowedPattern: ^((\[0-9][1-9][0-9]|1[0-9][0-9]|2[0-4][0-9]|25[0-5])\.)\((0-9|1[0-9]|2[0-9]|3[0-22])$\)
  ConstraintDescription: CIDR block parameter must be in the form x.x.x.x/0-32.
  Default: 0.0.0.0/0
  Description: CIDR block to allow SSH access to the bootstrap node.
  Type: String
PublicSubnet:
  Description: The public subnet to launch the bootstrap node into.
  Type: AWS::EC2::Subnet::Id
MasterSecurityGroupId:
  Description: The master security group ID for registering temporary rules.
  Type: AWS::EC2::SecurityGroup::Id
VpcId:
  Description: The VPC-scoped resources will belong to this VPC.
  Type: AWS::EC2::VPC::Id
BootstrapIgnitionLocation:
  Description: Ignition config file location.
  Type: String
```

---

241
AutoRegisterELB:
  Default: "yes"
  AllowedValues:
  - "yes"
  - "no"
  Description: Do you want to invoke NLB registration, which requires a Lambda ARN parameter?
  Type: String
RegisterNlbIpTargetsLambdaArn:
  Description: ARN for NLB IP target registration lambda.
  Type: String
ExternalApiTargetGroupArn:
  Description: ARN for external API load balancer target group.
  Type: String
InternalApiTargetGroupArn:
  Description: ARN for internal API load balancer target group.
  Type: String
InternalServiceTargetGroupArn:
  Description: ARN for internal service load balancer target group.
  Type: String

Metadata:
AWS::CloudFormation::Interface:
  ParameterGroups:
  - Label:
    default: "Cluster Information"
    Parameters:
    - InfrastructureName
    - Label:
      default: "Host Information"
      Parameters:
      - RhcosAmi
      - BootstrapIgnitionLocation
      - MasterSecurityGroupId
      - Label:
        default: "Network Configuration"
        Parameters:
        - VpcId
        - AllowedBootstrapSshCidr
        - PublicSubnet
        - Label:
          default: "Load Balancer Automation"
          Parameters:
          - AutoRegisterELB
          - RegisterNlbIpTargetsLambdaArn
          - ExternalApiTargetGroupArn
          - InternalApiTargetGroupArn
          - InternalServiceTargetGroupArn
  ParameterLabels:
  InfrastructureName:
    default: "Infrastructure Name"
  VpcId:
    default: "VPC ID"
  AllowedBootstrapSshCidr:
    default: "Allowed SSH Source"
  PublicSubnet:
    default: "Public Subnet"
RhcosAmi:
  default: "Red Hat Enterprise Linux CoreOS AMI ID"
BootstrapIgnitionLocation:
  default: "Bootstrap Ignition Source"
MasterSecurityGroupId:
  default: "Master Security Group ID"
AutoRegisterELB:
  default: "Use Provided ELB Automation"

Conditions:
  DoRegistration: !Equals ["yes", !Ref AutoRegisterELB]

Resources:
BootstrapIamRole:
  Type: AWS::IAM::Role
  Properties:
    AssumeRolePolicyDocument:
      Version: "2012-10-17"
      Statement:
        - Effect: "Allow"
          Principal:
            Service:
              - "ec2.amazonaws.com"
          Action:
            - "sts:AssumeRole"
          Path: "/
          Policies:
            - PolicyName: !Join ["-", [!Ref InfrastructureName, "bootstrap", "policy"]]
              PolicyDocument:
                Version: "2012-10-17"
                Statement:
                  - Effect: "Allow"
                    Action: "ec2:Describe*"
                    Resource: "*
                  - Effect: "Allow"
                    Action: "ec2:AttachVolume"
                    Resource: "*
                  - Effect: "Allow"
                    Action: "ec2:DetachVolume"
                    Resource: "*
                  - Effect: "Allow"
                    Action: "s3:GetObject"
                    Resource: "*

BootstrapInstanceProfile:
  Type: "AWS::IAM::InstanceProfile"
  Properties:
    Path: "/
    Roles:
      - Ref: "BootstrapIamRole"

BootstrapSecurityGroup:
  Type: AWS::EC2::SecurityGroup
  Properties:
    GroupDescription: Cluster Bootstrap Security Group
    SecurityGroupIngress:

CHAPTER 1. INSTALLING ON AWS

243
- IpProtocol: tcp
  FromPort: 22
  ToPort: 22
  CidrIp: !Ref AllowedBootstrapSshCidr
- IpProtocol: tcp
  FromPort: 19531
  ToPort: 19531
  CidrIp: 0.0.0.0/0
  VpcId: !Ref VpcId

BootstrapInstance:
  Type: AWS::EC2::Instance
  Properties:
    ImageId: !Ref RhcosAmi
    IamInstanceProfile: !Ref BootstrapInstanceProfile
    InstanceType: "i3.large"
    NetworkInterfaces:
      - AssociatePublicIpAddress: "true"
        DeviceIndex: "0"
        GroupSet:
          - !Ref "BootstrapSecurityGroup"
          - !Ref "MasterSecurityGroupId"
        SubnetId: !Ref "PublicSubnet"
  UserData:
    Fn::Base64: !Sub
    - '{"ignition":{"config":{"replace":{"source":"${S3Loc}"}},"version":"3.1.0"}}'
    - { S3Loc: !Ref BootstrapIgnitionLocation }

RegisterBootstrapApiTarget:
  Condition: DoRegistration
  Type: Custom::NLBRegister
  Properties:
    ServiceToken: !Ref RegisterNlbIpTargetsLambdaArn
    TargetArn: !Ref ExternalApiTargetGroupArn
    TargetIp: !GetAtt BootstrapInstance.PrivateIp

RegisterBootstrapInternalApiTarget:
  Condition: DoRegistration
  Type: Custom::NLBRegister
  Properties:
    ServiceToken: !Ref RegisterNlbIpTargetsLambdaArn
    TargetArn: !Ref InternalApiTargetGroupArn
    TargetIp: !GetAtt BootstrapInstance.PrivateIp

RegisterBootstrapInternalServiceTarget:
  Condition: DoRegistration
  Type: Custom::NLBRegister
  Properties:
    ServiceToken: !Ref RegisterNlbIpTargetsLambdaArn
    TargetArn: !Ref InternalServiceTargetGroupArn
    TargetIp: !GetAtt BootstrapInstance.PrivateIp

Outputs:
  BootstrapInstanceId:
Additional resources

- See RHCOS AMIs for the AWS infrastructure for details about the Red Hat Enterprise Linux CoreOS (RHCOS) AMIs for the AWS zones.

1.9.13. Creating the control plane machines in AWS

You must create the control plane machines in Amazon Web Services (AWS) that your cluster will use.

You can use the provided CloudFormation template and a custom parameter file to create a stack of AWS resources that represent the control plane nodes.

**IMPORTANT**

The CloudFormation template creates a stack that represents three control plane nodes.

**NOTE**

If you do not use the provided CloudFormation template to create your control plane nodes, you must review the provided information and manually create the infrastructure. If your cluster does not initialize correctly, you might have to contact Red Hat support with your installation logs.

Prerequisites

- You configured an AWS account.
- You added your AWS keys and region to your local AWS profile by running `aws configure`.
- You generated the Ignition config files for your cluster.
- You created and configured a VPC and associated subnets in AWS.
- You created and configured DNS, load balancers, and listeners in AWS.
- You created the security groups and roles required for your cluster in AWS.
- You created the bootstrap machine.

Procedure

1. Create a JSON file that contains the parameter values that the template requires:

```json
Description: Bootstrap Instance ID.
Value: !Ref BootstrapInstance

BootstrapPublicIp:
Description: The bootstrap node public IP address.
Value: !GetAtt BootstrapInstance.PublicIp

BootstrapPrivateIp:
Description: The bootstrap node private IP address.
Value: !GetAtt BootstrapInstance.PrivateIp
```
[{
    "ParameterKey": "InfrastructureName",
    "ParameterValue": "mycluster-<random_string>"
},
{
    "ParameterKey": "RhcosAmi",
    "ParameterValue": "ami-<random_string>"
},
{
    "ParameterKey": "AutoRegisterDNS",
    "ParameterValue": "yes"
},
{
    "ParameterKey": "PrivateHostedZoneId",
    "ParameterValue": "<random_string>"
},
{
    "ParameterKey": "PrivateHostedZoneName",
    "ParameterValue": "mycluster.example.com"
},
{
    "ParameterKey": "Master0Subnet",
    "ParameterValue": "subnet-<random_string>"
},
{
    "ParameterKey": "Master1Subnet",
    "ParameterValue": "subnet-<random_string>"
},
{
    "ParameterKey": "Master2Subnet",
    "ParameterValue": "subnet-<random_string>"
},
{
    "ParameterKey": "MasterSecurityGroupId",
    "ParameterValue": "sg-<random_string>"
},
{
    "ParameterKey": "IgnitionLocation",
    "ParameterValue": "https://api-int.<cluster_name>.<domain_name>:22623/config/master"
},
{
    "ParameterKey": "CertificateAuthorities",
    "ParameterValue": "data:text/plain;charset=utf-8;base64,ABC...xYz=="
},
{
    "ParameterKey": "MasterInstanceProfileName",
    "ParameterValue": "<roles_stack>-MasterInstanceProfile-<random_string>"
},
{
    "ParameterKey": "MasterInstanceType",
    "ParameterValue": "<instance_type>"
}]
The name for your cluster infrastructure that is encoded in your Ignition config files for the cluster.

Specify the infrastructure name that you extracted from the Ignition config file metadata, which has the format `<cluster-name>-<random-string>`.

Current Red Hat Enterprise Linux CoreOS (RHCOS) AMI to use for the control plane machines.

Specify an `AWS::EC2::Image::Id` value.

Whether or not to perform DNS etcd registration.

Specify `yes` or `no`. If you specify `yes`, you must provide hosted zone information.

The Route 53 private zone ID to register the etcd targets with.

Specify the `PrivateHostedZoneId` value from the output of the CloudFormation template for DNS and load balancing.

The Route 53 zone to register the targets with.

Specify `<cluster_name>.<domain_name>` where `<domain_name>` is the Route 53 base domain that you used when you generated `install-config.yaml` file for the cluster. Do not include the trailing period (.) that is displayed in the AWS console.

A subnet, preferably private, to launch the control plane machines on.
Specify a subnet from the `PrivateSubnets` value from the output of the CloudFormation template for DNS and load balancing.

The master security group ID to associate with control plane nodes (also known as the master nodes).

Specify the `MasterSecurityGroupId` value from the output of the CloudFormation template for the security group and roles.

The location to fetch control plane Ignition config file from.

Specify the generated Ignition config file location, `https://api-int.<cluster_name>.<domain_name>:22623/config/master`.

The base64 encoded certificate authority string to use.

Specify the value from the `master.ign` file that is in the installation directory. This value is the long string with the format `data:text/plain;charset=utf-8;base64,ABC…xYz==`.

The IAM profile to associate with control plane nodes.

Specify the `MasterInstanceProfile` parameter value from the output of the CloudFormation template for the security group and roles.

The type of AWS instance to use for the control plane machines.

Allowed values:

- `m4.xlarge`
- `m4.2xlarge`
- `m4.4xlarge`
- `m4.8xlarge`
- `m4.10xlarge`
- `m4.16xlarge`
- `m5.xlarge`
- `m5.2xlarge`
- `m5.4xlarge`
- `m5.8xlarge`
- `m5.10xlarge`
- `m5.16xlarge`
- `m6i.xlarge`
- `c4.2xlarge`
- `c4.4xlarge`
- c4.8xlarge
- r4.xlarge
- r4.2xlarge
- r4.4xlarge
- r4.8xlarge
- r4.16xlarge

**IMPORTANT**

If m4 instance types are not available in your region, such as with eu-west-3, specify an m5 type, such as m5.xlarge, instead.

Whether or not to register a network load balancer (NLB).

Specify yes or no. If you specify yes, you must provide a Lambda Amazon Resource Name (ARN) value.

The ARN for NLB IP target registration lambda group.

Specify the RegisterNlbIpTargetsLambda value from the output of the CloudFormation template for DNS and load balancing. Use arn:aws-us-gov if deploying the cluster to an AWS GovCloud region.

The ARN for external API load balancer target group.

Specify the ExternalApiTargetGroupArn value from the output of the CloudFormation template for DNS and load balancing. Use arn:aws-us-gov if deploying the cluster to an AWS GovCloud region.

The ARN for internal API load balancer target group.

Specify the InternalApiTargetGroupArn value from the output of the CloudFormation template for DNS and load balancing. Use arn:aws-us-gov if deploying the cluster to an AWS GovCloud region.

The ARN for internal service load balancer target group.

Specify the InternalServiceTargetGroupArn value from the output of the CloudFormation template for DNS and load balancing. Use arn:aws-us-gov if deploying the cluster to an AWS GovCloud region.

2. Copy the template from the CloudFormation template for control plane machines section of this topic and save it as a YAML file on your computer. This template describes the control plane machines that your cluster requires.

3. If you specified an m5 instance type as the value for MasterInstanceType, add that instance type to the MasterInstanceType.AllowedValues parameter in the CloudFormation template.

4. Launch the CloudFormation template to create a stack of AWS resources that represent the control plane nodes:
IMPORTANT
You must enter the command on a single line.

```
$ aws cloudformation create-stack --stack-name <name>  
   --template-body file://<template>.yaml  
   --parameters file://<parameters>.json
```

1. `<name>` is the name for the CloudFormation stack, such as `cluster-control-plane`. You need the name of this stack if you remove the cluster.

2. `<template>` is the relative path to and name of the CloudFormation template YAML file that you saved.

3. `<parameters>` is the relative path to and name of the CloudFormation parameters JSON file.

Example output

```
arn:aws:cloudformation:us-east-1:269333783861:stack/cluster-control-plane/21c7e2b0-2ee2-11eb-c6f6-0aa34627df4b
```

NOTE
The CloudFormation template creates a stack that represents three control plane nodes.

5. Confirm that the template components exist:

```
$ aws cloudformation describe-stacks --stack-name <name>
```

1.9.13.1. CloudFormation template for control plane machines

You can use the following CloudFormation template to deploy the control plane machines that you need for your OpenShift Container Platform cluster.

Example 1.29. CloudFormation template for control plane machines

```
AWSTemplateFormatVersion: 2010-09-09
Description: Template for OpenShift Cluster Node Launch (EC2 master instances)

Parameters:
  InfrastructureName:
    AllowedPattern: \^[a-zA-Z][a-zA-Z0-9-]{0,26}\$
    MaxLength: 27
    MinLength: 1
    ConstraintDescription: Infrastructure name must be alphanumeric, start with a letter, and have a maximum of 27 characters.
    Description: A short, unique cluster ID used to tag nodes for the kubelet cloud provider.
    Type: String
  RhcosAmi:
```
Description: Current Red Hat Enterprise Linux CoreOS AMI to use for bootstrap.
Type: AWS::EC2::Image::Id
AutoRegisterDNS:
  Default: "yes"
  AllowedValues:
    - "yes"
    - "no"
  Description: Do you want to invoke DNS etcd registration, which requires Hosted Zone information?
  Type: String
PrivateHostedZoneId:
  Description: The Route53 private zone ID to register the etcd targets with, such as Z21IXYZABCZ2A4.
  Type: String
PrivateHostedZoneName:
  Description: The Route53 zone to register the targets with, such as cluster.example.com. Omit the trailing period.
  Type: String
Master0Subnet:
  Description: The subnets, recommend private, to launch the master nodes into.
  Type: AWS::EC2::Subnet::Id
Master1Subnet:
  Description: The subnets, recommend private, to launch the master nodes into.
  Type: AWS::EC2::Subnet::Id
Master2Subnet:
  Description: The subnets, recommend private, to launch the master nodes into.
  Type: AWS::EC2::Subnet::Id
MasterSecurityGroupId:
  Description: The master security group ID to associate with master nodes.
  Type: AWS::EC2::SecurityGroup::Id
IgnitionLocation:
  Default: https://api-int.$CLUSTER_NAME.$DOMAIN:22623/config/master
  Description: Ignition config file location.
  Type: String
CertificateAuthorities:
  Default: data:text/plain;charset=utf-8;base64,ABC...xYz==
  Description: Base64 encoded certificate authority string to use.
  Type: String
MasterInstanceProfileName:
  Description: IAM profile to associate with master nodes.
  Type: String
MasterInstanceType:
  Default: m5.xlarge
  Type: String
  AllowedValues:
    - "m4.xlarge"
    - "m4.2xlarge"
    - "m4.4xlarge"
    - "m4.10xlarge"
    - "m4.16xlarge"
    - "m5.xlarge"
    - "m5.2xlarge"
    - "m5.4xlarge"
    - "m5.8xlarge"
    - "m5.12xlarge"
    - "m5.16xlarge"
- "m5a.xlarge"
- "m5a.2xlarge"
- "m5a.4xlarge"
- "m5a.8xlarge"
- "m5a.10xlarge"
- "m5a.16xlarge"
- "c4.2xlarge"
- "c4.4xlarge"
- "c4.8xlarge"
- "c5.2xlarge"
- "c5.4xlarge"
- "c5.9xlarge"
- "c5.12xlarge"
- "c5.18xlarge"
- "c5.24xlarge"
- "c5a.2xlarge"
- "c5a.4xlarge"
- "c5a.8xlarge"
- "c5a.12xlarge"
- "c5a.16xlarge"
- "c5a.24xlarge"
- "r4.xlarge"
- "r4.2xlarge"
- "r4.4xlarge"
- "r4.8xlarge"
- "r4.16xlarge"
- "r5.xlarge"
- "r5.2xlarge"
- "r5.4xlarge"
- "r5.8xlarge"
- "r5.12xlarge"
- "r5.16xlarge"
- "r5.24xlarge"
- "r5a.xlarge"
- "r5a.2xlarge"
- "r5a.4xlarge"
- "r5a.8xlarge"
- "r5a.12xlarge"
- "r5a.16xlarge"
- "r5a.24xlarge"

AutoRegisterELB:
  Default: "yes"
  AllowedValues:
  - "yes"
  - "no"
  Description: Do you want to invoke NLB registration, which requires a Lambda ARN parameter?
  Type: String

RegisterNlbIpTargetsLambdaArn:
  Description: ARN for NLB IP target registration lambda. Supply the value from the cluster infrastructure or select "no" for AutoRegisterELB.
  Type: String

ExternalApiTargetGroupArn:
  Description: ARN for external API load balancer target group. Supply the value from the cluster infrastructure or select "no" for AutoRegisterELB.
  Type: String
InternalApiTargetGroupArn:
Description: ARN for internal API load balancer target group. Supply the value from the cluster infrastructure or select "no" for AutoRegisterELB.
Type: String

InternalServiceTargetGroupArn:
Description: ARN for internal service load balancer target group. Supply the value from the cluster infrastructure or select "no" for AutoRegisterELB.
Type: String

Metadata:
AWS::CloudFormation::Interface:
ParameterGroups:
- Label: "default"::"Cluster Information"
Parameters:
- InfrastructureName
- Label: "default"::"Host Information"
Parameters:
- MasterInstanceType
- RhcosAmi
- IgnitionLocation
- CertificateAuthorities
- MasterSecurityGroupId
- MasterInstanceProfileName
- Label: "default"::"Network Configuration"
Parameters:
- VpcId
- AllowedBootstrapSshCidr
- Master0Subnet
- Master1Subnet
- Master2Subnet
- Label: "default"::"DNS"
Parameters:
- AutoRegisterDNS
- PrivateHostedZoneName
- PrivateHostedZoneId
- Label: "default"::"Load Balancer Automation"
Parameters:
- AutoRegisterELB
- RegisterNlbIpTargetsLambdaArn
- ExternalApiTargetGroupArn
- InternalApiTargetGroupArn
- InternalServiceTargetGroupArn

ParameterLabels:
InfrastructureName:
- default: "Infrastructure Name"
VpcId:
- default: "VPC ID"
Master0Subnet:
- default: "Master-0 Subnet"
Master1Subnet:
- default: "Master-1 Subnet"
Master2Subnet:
  default: "Master-2 Subnet"
MasterInstanceType:
  default: "Master Instance Type"
MasterInstanceProfileName:
  default: "Master Instance Profile Name"
RhcosAmi:
  default: "Red Hat Enterprise Linux CoreOS AMI ID"
BootstrapIgnitionLocation:
  default: "Master Ignition Source"
CertificateAuthorities:
  default: "Ignition CA String"
MasterSecurityGroupId:
  default: "Master Security Group ID"
AutoRegisterDNS:
  default: "Use Provided DNS Automation"
AutoRegisterELB:
  default: "Use Provided ELB Automation"
PrivateHostedZoneName:
  default: "Private Hosted Zone Name"
PrivateHostedZoneId:
  default: "Private Hosted Zone ID"

Conditions:
DoRegistration: !Equals ["yes", !Ref AutoRegisterELB]
DoDns: !Equals ["yes", !Ref AutoRegisterDNS]

Resources:
  Master0:
    Type: AWS::EC2::Instance
    Properties:
      ImageId: !Ref RhcosAmi
      BlockDeviceMappings:
        - DeviceName: /dev/xvda
          Ebs:
            VolumeSize: "120"
            VolumeType: "gp2"
      IamInstanceProfile: !Ref MasterInstanceProfileName
      InstanceType: !Ref MasterInstanceType
      NetworkInterfaces:
        - AssociatePublicIpAddress: "false"
          DeviceIndex: "0"
          GroupSet:
            - !Ref "MasterSecurityGroupId"
          SubnetId: !Ref "Master2Subnet"
      UserData:
        Fn::Base64: !Sub
          - "\"ignition\":\"{\"config\":\"{\"merge\":\"{\"source\":\"${SOURCE}\"}\"},\"security\":\"{\"tls\":\"{\"certificateAuthorities\":\"{\"source\":\"${CA_BUNDLE}\"}\"},\"version\":\"3.1.0\"}\"}\"\"}
            - "SOURCE: !Ref IgnitionLocation,
              CA_BUNDLE: !Ref CertificateAuthorities,
              \"version\": \"3.1.0\"
        \"
      Tags:
        - Key: !Join ["", [\"kubernetes.io/cluster/\", !Ref InfrastructureName]]
          Value: "shared"
RegisterMaster0:
Condition: DoRegistration
Type: Custom::NLBRegister
Properties:
  ServiceToken: !Ref RegisterNlbIpTargetsLambdaArn
  TargetArn: !Ref ExternalApiTargetGroupArn
  TargetIp: !GetAtt Master0.PrivateIp

RegisterMaster0InternalApiTarget:
Condition: DoRegistration
Type: Custom::NLBRegister
Properties:
  ServiceToken: !Ref RegisterNlbIpTargetsLambdaArn
  TargetArn: !Ref InternalApiTargetGroupArn
  TargetIp: !GetAtt Master0.PrivateIp

RegisterMaster0InternalServiceTarget:
Condition: DoRegistration
Type: Custom::NLBRegister
Properties:
  ServiceToken: !Ref RegisterNlbIpTargetsLambdaArn
  TargetArn: !Ref InternalServiceTargetGroupArn
  TargetIp: !GetAtt Master0.PrivateIp

Master1:
Type: AWS::EC2::Instance
Properties:
  ImageId: !Ref RhcosAmi
  BlockDeviceMappings:
    - DeviceName: /dev/xvda
      Ebs:
        VolumeSize: "120"
        VolumeType: "gp2"
  IamInstanceProfile: !Ref MasterInstanceProfileName
  InstanceType: !Ref MasterInstanceType
  NetworkInterfaces:
    - AssociatePublicIpAddress: "false"
      DeviceIndex: "0"
      GroupSet:
        - !Ref "MasterSecurityGroupId"
      SubnetId: !Ref "Master1Subnet"
  UserData:
    Fn::Base64: !Sub
      - '{{"ignition":{"config":{"merge":["source":"${SOURCE}"],"security":{"tls":{"certificateAuthorities":["source":"${CA_BUNDLE}"],"version":"3.1.0"}}},"version":"3.1.0"}}
        -
          SOURCE: !Ref IgnitionLocation,
          CA_BUNDLE: !Ref CertificateAuthorities,
    Tags:
      - Key: !Join ["", ["kubernetes.io/cluster/", !Ref InfrastructureName]]
        Value: "shared"

RegisterMaster1:
Condition: DoRegistration
Type: Custom::NLBRegister
Properties:
  ServiceToken: !Ref RegisterNlbIpTargetsLambdaArn
  TargetArn: !Ref ExternalApiTargetGroupArn
  TargetIp: !GetAtt Master1.PrivateIp

RegisterMaster1InternalApiTarget:
Condition: DoRegistration
Type: Custom::NLBRegister
Properties:
  ServiceToken: !Ref RegisterNlbIpTargetsLambdaArn
  TargetArn: !Ref InternalApiTargetGroupArn
  TargetIp: !GetAtt Master1.PrivateIp

RegisterMaster1InternalServiceTarget:
Condition: DoRegistration
Type: Custom::NLBRegister
Properties:
  ServiceToken: !Ref RegisterNlbIpTargetsLambdaArn
  TargetArn: !Ref InternalServiceTargetGroupArn
  TargetIp: !GetAtt Master1.PrivateIp

Master2:
Type: AWS::EC2::Instance
Properties:
  ImageId: !Ref RhcosAmi
  BlockDeviceMappings:
    - DeviceName: /dev/xvda
      Ebs:
        VolumeSize: "120"
        VolumeType: "gp2"
  IamInstanceProfile: !Ref MasterInstanceProfileName
  InstanceType: !Ref MasterInstanceType
  NetworkInterfaces:
    - AssociatePublicIpAddress: "false"
      GroupSet:
        - !Ref "MasterSecurityGroupId"
      SubnetId: !Ref "Master2Subnet"
  UserData:
    Fn::Base64:
      - '{"ignition":{"config":{"merge":[{"source":"${SOURCE}"}]},"security":{"tls":{"certificateAuthorities":[{"source":"${CA_BUNDLE}"}],"version":"3.1.0"}}}
        - {
            SOURCE: !Ref IgnitionLocation,
            CA_BUNDLE: !Ref CertificateAuthorities,
        }
  Tags:
    - Key: !Join ["", ["kubernetes.io/cluster/", !Ref InfrastructureName]]
      Value: "shared"

RegisterMaster2:
Condition: DoRegistration
Type: Custom::NLBRegister
Properties:
  ServiceToken: !Ref RegisterNlbIpTargetsLambdaArn
TargetArn: !Ref ExternalApiTargetGroupArn
TargetIp: !GetAtt Master2.PrivateIp

RegisterMaster2InternalApiTarget:
Condition: DoRegistration
Type: Custom::NLBRegister
Properties:
  ServiceToken: !Ref RegisterNlbIpTargetsLambdaArn
  TargetArn: !Ref InternalApiTargetGroupArn
  TargetIp: !GetAtt Master2.PrivateIp

RegisterMaster2InternalServiceTarget:
Condition: DoRegistration
Type: Custom::NLBRegister
Properties:
  ServiceToken: !Ref RegisterNlbIpTargetsLambdaArn
  TargetArn: !Ref InternalServiceTargetGroupArn
  TargetIp: !GetAtt Master2.PrivateIp

EtcdSrvRecords:
Condition: DoDns
Type: AWS::Route53::RecordSet
Properties:
  HostedZoneId: !Ref PrivateHostedZoneId
  Name: !Join [".", ["_etcd-server-ssl._tcp", !Ref PrivateHostedZoneName]]
  ResourceRecords:
    - !Join [
      " ",
      ["0 10 2380", !Join [".", ["etcd-0", !Ref PrivateHostedZoneName]]],
    ]
    - !Join [
      " ",
      ["0 10 2380", !Join [".", ["etcd-1", !Ref PrivateHostedZoneName]]],
    ]
    - !Join [
      " ",
      ["0 10 2380", !Join [".", ["etcd-2", !Ref PrivateHostedZoneName]]],
    ]
  TTL: 60
  Type: SRV

Etcd0Record:
Condition: DoDns
Type: AWS::Route53::RecordSet
Properties:
  HostedZoneId: !Ref PrivateHostedZoneId
  Name: !Join [".", ["etcd-0", !Ref PrivateHostedZoneName]]
  ResourceRecords:
    - !GetAtt Master0.PrivateIp
  TTL: 60
  Type: A

Etcd1Record:
Condition: DoDns
Type: AWS::Route53::RecordSet
Properties:
1.9.14. Creating the worker nodes in AWS

You can create worker nodes in Amazon Web Services (AWS) for your cluster to use.

You can use the provided CloudFormation template and a custom parameter file to create a stack of AWS resources that represent a worker node.

**IMPORTANT**

The CloudFormation template creates a stack that represents one worker node. You must create a stack for each worker node.

**NOTE**

If you do not use the provided CloudFormation template to create your worker nodes, you must review the provided information and manually create the infrastructure. If your cluster does not initialize correctly, you might have to contact Red Hat support with your installation logs.

**Prerequisites**

- You configured an AWS account.
- You added your AWS keys and region to your local AWS profile by running `aws configure`.
- You generated the Ignition config files for your cluster.
• You created and configured a VPC and associated subnets in AWS.
• You created and configured DNS, load balancers, and listeners in AWS.
• You created the security groups and roles required for your cluster in AWS.
• You created the bootstrap machine.
• You created the control plane machines.

Procedure

1. Create a JSON file that contains the parameter values that the CloudFormation template requires:

```json
[
  {
    "ParameterKey": "InfrastructureName",  
    "ParameterValue": "mycluster-<random_string>"  
  },
  {
    "ParameterKey": "RhcosAmi",  
    "ParameterValue": "ami-<random_string>"  
  },
  {
    "ParameterKey": "Subnet",  
    "ParameterValue": "subnet-<random_string>"  
  },
  {
    "ParameterKey": "WorkerSecurityGroupId",  
    "ParameterValue": "sg-<random_string>"  
  },
  {
    "ParameterKey": "IgnitionLocation",  
    "ParameterValue": "https://api-int.<cluster_name>.<domain_name>:22623/config/worker"  
  },
  {
    "ParameterKey": "CertificateAuthorities",  
    "ParameterValue": ""  
  },
  {
    "ParameterKey": "WorkerInstanceProfileName",  
    "ParameterValue": ""  
  },
  {
    "ParameterKey": "WorkerInstanceType",  
    "ParameterValue": "m4.large"  
  }
]
```

1. **The name for your cluster infrastructure that is encoded in your Ignition config files for the cluster.**
Specify the infrastructure name that you extracted from the Ignition config file metadata, which has the format `<cluster-name>-<random-string>`.

Current Red Hat Enterprise Linux CoreOS (RHCOS) AMI to use for the worker nodes.

Specify an AWS::EC2::Image::Id value.

A subnet, preferably private, to launch the worker nodes on.

Specify a subnet from the PrivateSubnets value from the output of the CloudFormation template for DNS and load balancing.

The worker security group ID to associate with worker nodes.

Specify the WorkerSecurityGroupId value from the output of the CloudFormation template for the security group and roles.

The location to fetch bootstrap Ignition config file from.

Specify the generated Ignition config location, `https://api-int.<cluster_name>.<domain_name>:22623/config/worker`.

Base64 encoded certificate authority string to use.

Specify the value from the worker.ign file that is in the installation directory. This value is the long string with the format `data:text/plain;charset=utf-8;base64,ABC…xYz==`.

The IAM profile to associate with worker nodes.

Specify the WorkerInstanceProfile parameter value from the output of the CloudFormation template for the security group and roles.

The type of AWS instance to use for the control plane machines.

Allowed values:

- m4.large
- m4.xlarge
- m4.2xlarge
- m4.4xlarge
- m4.8xlarge
- m4.10xlarge
- m4.16xlarge
- m5.large
- m5.xlarge
- m5.2xlarge
- m5.4xlarge
• m5.8xlarge
• m5.10xlarge
• m5.16xlarge
• m6i.xlarge
• c4.2xlarge
• c4.4xlarge
• c4.8xlarge
• r4.large
• r4.xlarge
• r4.2xlarge
• r4.4xlarge
• r4.8xlarge
• r4.16xlarge

IMPORTANT
If m4 instance types are not available in your region, such as with eu-west-3, use m5 types instead.

2. Copy the template from the CloudFormation template for worker machines section of this topic and save it as a YAML file on your computer. This template describes the networking objects and load balancers that your cluster requires.

3. If you specified an m5 instance type as the value for WorkerInstanceType, add that instance type to the WorkerInstanceType.AllowedValues parameter in the CloudFormation template.

4. Launch the CloudFormation template to create a stack of AWS resources that represent a worker node:

   IMPORTANT
   You must enter the command on a single line.

   $ aws cloudformation create-stack --stack-name <name> 1
   --template-body:file://<template>.yaml 
   --parameters file://<parameters>.json 3

   1 <name> is the name for the CloudFormation stack, such as cluster-worker-1. You need the name of this stack if you remove the cluster.

   2 <template> is the relative path to and name of the CloudFormation template YAML file that you saved.
3. `<parameters>` is the relative path to and name of the CloudFormation parameters JSON file.

**Example output**

```
arn:aws:cloudformation:us-east-1:269333783861:stack/cluster-worker-1/729ee301-1c2a-11eb-348f-3d9888c65b59
```

**NOTE**

The CloudFormation template creates a stack that represents one worker node.

5. Confirm that the template components exist:

```
$ aws cloudformation describe-stacks --stack-name <name>
```

6. Continue to create worker stacks until you have created enough worker machines for your cluster. You can create additional worker stacks by referencing the same template and parameter files and specifying a different stack name.

**IMPORTANT**

You must create at least two worker machines, so you must create at least two stacks that use this CloudFormation template.

1.9.15. Initializing the bootstrap sequence on AWS with user-provisioned infrastructure

After you create all of the required infrastructure in Amazon Web Services (AWS), you can start the bootstrap sequence that initializes the OpenShift Container Platform control plane.

**Prerequisites**

- You configured an AWS account.
- You added your AWS keys and region to your local AWS profile by running `aws configure`.
- You generated the Ignition config files for your cluster.
- You created and configured a VPC and associated subnets in AWS.
- You created and configured DNS, load balancers, and listeners in AWS.
- You created the security groups and roles required for your cluster in AWS.
- You created the bootstrap machine.
- You created the control plane machines.
- You created the worker nodes.

**Procedure**

1. Change to the directory that contains the installation program and start the bootstrap process that initializes the OpenShift Container Platform control plane:

```
$ ./openshift-install wait-for bootstrap-complete --dir <installation_directory> \
   --log-level=info
```

1. For `<installation_directory>`, specify the path to the directory that you stored the installation files in.

2. To view different installation details, specify `warn`, `debug`, or `error` instead of `info`.

**Example output**

```
INFO Waiting up to 20m0s for the Kubernetes API at https://api.mycluster.example.com:6443...
INFO API v1.19.0+9f84db3 up
INFO Waiting up to 30m0s for bootstrapping to complete...
INFO It is now safe to remove the bootstrap resources
INFO Time elapsed: 1s
```

If the command exits without a **FATAL** warning, your OpenShift Container Platform control plane has initialized.

**NOTE**

After the control plane initializes, it sets up the compute nodes and installs additional services in the form of Operators.

**Additional resources**

- See [Monitoring installation progress](#) for details about monitoring the installation, bootstrap, and control plane logs as an OpenShift Container Platform installation progresses.

- See [Gathering bootstrap node diagnostic data](#) for information about troubleshooting issues related to the bootstrap process.

### 1.9.16. Logging in to the cluster by using the CLI

You can log in to your cluster as a default system user by exporting the cluster **kubeconfig** file. The **kubeconfig** file contains information about the cluster that is used by the CLI to connect a client to the correct cluster and API server. The file is specific to a cluster and is created during OpenShift Container Platform installation.

**Prerequisites**

- You deployed an OpenShift Container Platform cluster.

- You installed the **oc** CLI.

**Procedure**

1. Export the **kubeadmin** credentials:
1. Export the KUBECONFIG configuration file:

```bash
$ export KUBECONFIG=<installation_directory>/auth/kubeconfig
```

For `<installation_directory>`, specify the path to the directory that you stored the installation files in.

2. Verify you can run `oc` commands successfully using the exported configuration:

```bash
$ oc whoami
```

Example output

```
system:admin
```

### 1.9.17. Approving the certificate signing requests for your machines

When you add machines to a cluster, two pending certificate signing requests (CSRs) are generated for each machine that you added. You must confirm that these CSRs are approved or, if necessary, approve them yourself. The client requests must be approved first, followed by the server requests.

**Prerequisites**

- You added machines to your cluster.

**Procedure**

1. Confirm that the cluster recognizes the machines:

```bash
$ oc get nodes
```

Example output

```
NAME    STATUS    ROLES   AGE   VERSION
master-0 Ready master 63m  v1.19.0
master-1 Ready master 63m  v1.19.0
master-2 Ready master 64m  v1.19.0
```

The output lists all of the machines that you created.

**NOTE**

The preceding output might not include the compute nodes, also known as worker nodes, until some CSRs are approved.

2. Review the pending CSRs and ensure that you see the client requests with the **Pending** or **Approved** status for each machine that you added to the cluster:

```bash
$ oc get csr
```

Example output

...
In this example, two machines are joining the cluster. You might see more approved CSRs in the list.

3. If the CSRs were not approved, after all of the pending CSRs for the machines you added are in **Pending** status, approve the CSRs for your cluster machines:

**NOTE**

Because the CSRs rotate automatically, approve your CSRs within an hour of adding the machines to the cluster. If you do not approve them within an hour, the certificates will rotate, and more than two certificates will be present for each node. You must approve all of these certificates. Once the client CSR is approved, the Kubelet creates a secondary CSR for the serving certificate, which requires manual approval. Then, subsequent serving certificate renewal requests are automatically approved by the **machine-approver** if the Kubelet requests a new certificate with identical parameters.

**NOTE**

For clusters running on platforms that are not machine API enabled, such as bare metal and other user-provisioned infrastructure, you must implement a method of automatically approving the kubelet serving certificate requests (CSRs). If a request is not approved, then the **oc exec, oc rsh, and oc logs** commands cannot succeed, because a serving certificate is required when the API server connects to the kubelet. Any operation that contacts the Kubelet endpoint requires this certificate approval to be in place. The method must watch for new CSRs, confirm that the CSR was submitted by the **node-bootstrapper** service account in the **system:node** or **system:admin** groups, and confirm the identity of the node.

- To approve them individually, run the following command for each valid CSR:

  ```
  $ oc adm certificate approve <csr_name> 1
  
  1 <csr_name> is the name of a CSR from the list of current CSRs.
  ```

- To approve all pending CSRs, run the following command:

  ```
  $ oc get csr -o go-template='{{range .items}}{{if not .status}}{{.metadata.name}}{{"\n"}}{{end}}{{end}}' | xargs --no-run-if-empty oc adm certificate approve
  ```

**NOTE**

Some Operators might not become available until some CSRs are approved.
4. Now that your client requests are approved, you must review the server requests for each machine that you added to the cluster:

```bash
$ oc get csr
```

**Example output**

<table>
<thead>
<tr>
<th>NAME</th>
<th>AGE</th>
<th>REQUESTOR</th>
<th>CONDITION</th>
</tr>
</thead>
<tbody>
<tr>
<td>csr-bfd72</td>
<td>5m26s</td>
<td>system:node:ip-10-0-50-126.us-east-2.compute.internal</td>
<td>Pending</td>
</tr>
<tr>
<td>csr-c57lv</td>
<td>5m26s</td>
<td>system:node:ip-10-0-95-157.us-east-2.compute.internal</td>
<td>Pending</td>
</tr>
</tbody>
</table>

5. If the remaining CSRs are not approved, and are in the **Pending** status, approve the CSRs for your cluster machines:

- To approve them individually, run the following command for each valid CSR:

```bash
$ oc adm certificate approve <csr_name>  
```

<csr_name> is the name of a CSR from the list of current CSRs.

- To approve all pending CSRs, run the following command:

```bash
$ oc get csr -o go-template='{{range .items}}{{if not .status}}{{.metadata.name}}{{"\n"}}{{end}}{{end}}' | xargs oc adm certificate approve
```

6. After all client and server CSRs have been approved, the machines have the **Ready** status. Verify this by running the following command:

```bash
$ oc get nodes
```

**Example output**

<table>
<thead>
<tr>
<th>NAME</th>
<th>STATUS</th>
<th>ROLES</th>
<th>AGE</th>
<th>VERSION</th>
</tr>
</thead>
<tbody>
<tr>
<td>master-0</td>
<td>Ready</td>
<td>master</td>
<td>73m</td>
<td>v1.20.0</td>
</tr>
<tr>
<td>master-1</td>
<td>Ready</td>
<td>master</td>
<td>73m</td>
<td>v1.20.0</td>
</tr>
<tr>
<td>master-2</td>
<td>Ready</td>
<td>master</td>
<td>74m</td>
<td>v1.20.0</td>
</tr>
<tr>
<td>worker-0</td>
<td>Ready</td>
<td>worker</td>
<td>11m</td>
<td>v1.20.0</td>
</tr>
<tr>
<td>worker-1</td>
<td>Ready</td>
<td>worker</td>
<td>11m</td>
<td>v1.20.0</td>
</tr>
</tbody>
</table>

**NOTE**

It can take a few minutes after approval of the server CSRs for the machines to transition to the **Ready** status.

**Additional information**

- For more information on CSRs, see [Certificate Signing Requests](#).
1.9.18. Initial Operator configuration

After the control plane initializes, you must immediately configure some Operators so that they all become available.

Prerequisites

- Your control plane has initialized.

Procedure

1. Watch the cluster components come online:

   ```
   $ watch -n5 oc get clusteroperators
   ```

Example output

<table>
<thead>
<tr>
<th>NAME</th>
<th>VERSION</th>
<th>AVAILABLE</th>
<th>PROGRESSING</th>
<th>DEGRADED</th>
</tr>
</thead>
<tbody>
<tr>
<td>authentication</td>
<td>4.6.0</td>
<td>True</td>
<td>False</td>
<td>False</td>
</tr>
<tr>
<td>cloud-credential</td>
<td>4.6.0</td>
<td>True</td>
<td>False</td>
<td>False</td>
</tr>
<tr>
<td>cluster-autoscaler</td>
<td>4.6.0</td>
<td>True</td>
<td>False</td>
<td>False</td>
</tr>
<tr>
<td>config-operator</td>
<td>4.6.0</td>
<td>True</td>
<td>False</td>
<td>False</td>
</tr>
<tr>
<td>console</td>
<td>4.6.0</td>
<td>True</td>
<td>False</td>
<td>False</td>
</tr>
<tr>
<td>csi-snapshot-controller</td>
<td>4.6.0</td>
<td>True</td>
<td>False</td>
<td>False</td>
</tr>
<tr>
<td>dns</td>
<td>4.6.0</td>
<td>True</td>
<td>False</td>
<td>False</td>
</tr>
<tr>
<td>etcd</td>
<td>4.6.0</td>
<td>True</td>
<td>False</td>
<td>False</td>
</tr>
<tr>
<td>image-registry</td>
<td>4.6.0</td>
<td>True</td>
<td>False</td>
<td>False</td>
</tr>
<tr>
<td>ingress</td>
<td>4.6.0</td>
<td>True</td>
<td>False</td>
<td>False</td>
</tr>
<tr>
<td>insights</td>
<td>4.6.0</td>
<td>True</td>
<td>False</td>
<td>False</td>
</tr>
<tr>
<td>kube-apiserver</td>
<td>4.6.0</td>
<td>True</td>
<td>False</td>
<td>False</td>
</tr>
<tr>
<td>kube-controller-manager</td>
<td>4.6.0</td>
<td>True</td>
<td>False</td>
<td>False</td>
</tr>
<tr>
<td>kube-scheduler</td>
<td>4.6.0</td>
<td>True</td>
<td>False</td>
<td>False</td>
</tr>
<tr>
<td>kube-storage-version-migrator</td>
<td>4.6.0</td>
<td>True</td>
<td>False</td>
<td>False</td>
</tr>
<tr>
<td>machine-api</td>
<td>4.6.0</td>
<td>True</td>
<td>False</td>
<td>False</td>
</tr>
<tr>
<td>machine-approver</td>
<td>4.6.0</td>
<td>True</td>
<td>False</td>
<td>False</td>
</tr>
<tr>
<td>machine-config</td>
<td>4.6.0</td>
<td>True</td>
<td>False</td>
<td>False</td>
</tr>
<tr>
<td>marketplace</td>
<td>4.6.0</td>
<td>True</td>
<td>False</td>
<td>False</td>
</tr>
<tr>
<td>monitoring</td>
<td>4.6.0</td>
<td>True</td>
<td>False</td>
<td>False</td>
</tr>
<tr>
<td>network</td>
<td>4.6.0</td>
<td>True</td>
<td>False</td>
<td>False</td>
</tr>
<tr>
<td>node-tuning</td>
<td>4.6.0</td>
<td>True</td>
<td>False</td>
<td>False</td>
</tr>
<tr>
<td>openshift-apiserver</td>
<td>4.6.0</td>
<td>True</td>
<td>False</td>
<td>False</td>
</tr>
<tr>
<td>openshift-controller-manager</td>
<td>4.6.0</td>
<td>True</td>
<td>False</td>
<td>False</td>
</tr>
<tr>
<td>openshift-samples</td>
<td>4.6.0</td>
<td>True</td>
<td>False</td>
<td>False</td>
</tr>
<tr>
<td>operator-lifecycle-manager</td>
<td>4.6.0</td>
<td>True</td>
<td>False</td>
<td>False</td>
</tr>
<tr>
<td>operator-lifecycle-manager-catalog</td>
<td>4.6.0</td>
<td>True</td>
<td>False</td>
<td>False</td>
</tr>
<tr>
<td>operator-lifecycle-manager-packageserver</td>
<td>4.6.0</td>
<td>True</td>
<td>False</td>
<td>False</td>
</tr>
<tr>
<td>service-ca</td>
<td>4.6.0</td>
<td>True</td>
<td>False</td>
<td>False</td>
</tr>
<tr>
<td>storage</td>
<td>4.6.0</td>
<td>True</td>
<td>False</td>
<td>False</td>
</tr>
</tbody>
</table>

2. Configure the Operators that are not available.

1.9.18.1. Disabling the default OperatorHub sources
Operator catalogs that source content provided by Red Hat and community projects are configured for OperatorHub by default during an OpenShift Container Platform installation. In a restricted network environment, you must disable the default catalogs as a cluster administrator.

**Procedure**

- Disable the sources for the default catalogs by adding `disableAllDefaultSources: true` to the `OperatorHub` object:

  ```
  $ oc patch OperatorHub cluster --type json \
  -p '[{"op": "add", "path": "/spec/disableAllDefaultSources", "value": true}]'
  ```

**TIP**

Alternatively, you can use the web console to manage catalog sources. From the **Administration → Cluster Settings → Global Configuration → OperatorHub** page, click the **Sources** tab, where you can create, delete, disable, and enable individual sources.

### 1.9.18.2. Image registry storage configuration

Amazon Web Services provides default storage, which means the Image Registry Operator is available after installation. However, if the Registry Operator cannot create an S3 bucket and automatically configure storage, you must manually configure registry storage.

Instructions are shown for configuring a persistent volume, which is required for production clusters. Where applicable, instructions are shown for configuring an empty directory as the storage location, which is available for only non-production clusters.

Additional instructions are provided for allowing the image registry to use block storage types by using the **Recreate** rollout strategy during upgrades.

#### 1.9.18.2.1. Configuring registry storage for AWS with user-provisioned infrastructure

During installation, your cloud credentials are sufficient to create an Amazon S3 bucket and the Registry Operator will automatically configure storage.

If the Registry Operator cannot create an S3 bucket and automatically configure storage, you can create an S3 bucket and configure storage with the following procedure.

**Prerequisites**

- You have a cluster on AWS with user-provisioned infrastructure.

- For Amazon S3 storage, the secret is expected to contain two keys:
  
  - **REGISTRY_STORAGE_S3_ACCESSKEY**
  
  - **REGISTRY_STORAGE_S3_SECRETKEY**

**Procedure**

Use the following procedure if the Registry Operator cannot create an S3 bucket and automatically configure storage.

1. Set up a **Bucket Lifecycle Policy** to abort incomplete multipart uploads that are one day old.
2. Fill in the storage configuration in `configs.imageregistry.operator.openshift.io/cluster`:

```
$ oc edit configs.imageregistry.operator.openshift.io/cluster
```

**Example configuration**

```
storage:
  s3:
    bucket: <bucket-name>
    region: <region-name>
```

**WARNING**
To secure your registry images in AWS, **block public access** to the S3 bucket.

1.9.18.2. Configuring storage for the image registry in non-production clusters

You must configure storage for the Image Registry Operator. For non-production clusters, you can set the image registry to an empty directory. If you do so, all images are lost if you restart the registry.

**Procedure**

- To set the image registry storage to an empty directory:

  ```
  $ oc patch configs.imageregistry.operator.openshift.io cluster --type merge --patch '{{"spec":{{"storage":{{"emptyDir":[]}}}}}'}
  ```

**WARNING**
Configure this option for only non-production clusters.

If you run this command before the Image Registry Operator initializes its components, the `oc patch` command fails with the following error:

```
Error from server (NotFound): configs.imageregistry.operator.openshift.io "cluster" not found
```

Wait a few minutes and run the command again.

1.9.19. Deleting the bootstrap resources

After you complete the initial Operator configuration for the cluster, remove the bootstrap resources from Amazon Web Services (AWS).
Prerequisites

- You completed the initial Operator configuration for your cluster.

Procedure

1. Delete the bootstrap resources. If you used the CloudFormation template, delete its stack:
   - Delete the stack by using the AWS CLI:
     $$ aws cloudformation delete-stack --stack-name <name>$$
   - <name> is the name of your bootstrap stack.
   - Delete the stack by using the AWS CloudFormation console.

1.9.20. Creating the Ingress DNS Records

If you removed the DNS Zone configuration, manually create DNS records that point to the Ingress load balancer. You can create either a wildcard record or specific records. While the following procedure uses A records, you can use other record types that you require, such as CNAME or alias.

Prerequisites

- You deployed an OpenShift Container Platform cluster on Amazon Web Services (AWS) that uses infrastructure that you provisioned.
- You installed the OpenShift CLI (oc).
- You installed the jq package.
- You downloaded the AWS CLI and installed it on your computer. See Install the AWS CLI Using the Bundled Installer (Linux, macOS, or Unix).

Procedure

1. Determine the routes to create.
   - To create a wildcard record, use *.apps.<cluster_name>.<domain_name>, where <cluster_name> is your cluster name, and <domain_name> is the Route 53 base domain for your OpenShift Container Platform cluster.
   - To create specific records, you must create a record for each route that your cluster uses, as shown in the output of the following command:
     $$ oc get --all-namespaces -o jsonpath="\{range .items[*]\}{range .status.ingress[*]\}{.host\{"\n\}"\{end\}\{end\}\} routes$$

Example output

oauth-openshift.apps.<cluster_name>.<domain_name>
console-openshift-console.apps.<cluster_name>.<domain_name>
downloads-openshift-console.apps.<cluster_name>.<domain_name>
alertmanager-main-openshift-monitoring.apps.<cluster_name>.<domain_name>
grafana-openshift-monitoring.apps.<cluster_name>.<domain_name>
prometheus-k8s-openshift-monitoring.apps.<cluster_name>.<domain_name>

2. Retrieve the Ingress Operator load balancer status and note the value of the external IP address that it uses, which is shown in the EXTERNAL-IP column:

   $ oc -n openshift-ingress get service router-default

   **Example output**
   
<table>
<thead>
<tr>
<th>NAME</th>
<th>TYPE</th>
<th>CLUSTER-IP</th>
<th>EXTERNAL-IP</th>
<th>PORT(S)</th>
</tr>
</thead>
</table>
   | router-default | LoadBalancer   | 172.30.62.215 | ab3...28.us-east-2.elb.amazonaws.com   | 80:31499/TCP,443:30693/TCP   | 5m

3. Locate the hosted zone ID for the load balancer:

   $ aws elb describe-load-balancers | jq -r '.LoadBalancerDescriptions[] | select(.DNSName == "<external_ip>").CanonicalHostedZoneNameID'

   1 For `<external_ip>`, specify the value of the external IP address of the Ingress Operator load balancer that you obtained.

   **Example output**
   
   Z3AADJGX6KTTL2

   The output of this command is the load balancer hosted zone ID.

4. Obtain the public hosted zone ID for your cluster’s domain:

   $ aws route53 list-hosted-zones-by-name \
   --dns-name "<domain_name>"  
   --query 'HostedZones[? Config.PrivateZone != `true` && Name == `<domain_name>.`].Id'  
   --output text

   1 2 For `<domain_name>`, specify the Route 53 base domain for your OpenShift Container Platform cluster.

   **Example output**
   
   /hostedzone/Z3URY6TWQ91KVV

   The public hosted zone ID for your domain is shown in the command output. In this example, it is Z3URY6TWQ91KVV.

5. Add the alias records to your private zone:

   $ aws route53 change-resource-record-sets --hosted-zone-id "<private_hosted_zone_id>" --
For `<private_hosted_zone_id>`, specify the value from the output of the CloudFormation template for DNS and load balancing.

For `<cluster_domain>`, specify the domain or subdomain that you use with your OpenShift Container Platform cluster.

For `<hosted_zone_id>`, specify the public hosted zone ID for the load balancer that you obtained.

For `<external_ip>`, specify the value of the external IP address of the Ingress Operator load balancer. Ensure that you include the trailing period (\.) in this parameter value.

6. Add the records to your public zone:

```bash
$ aws route53 change-resource-record-sets --hosted-zone-id "<public_hosted_zone_id>" --change-batch '{
  "Changes": [
    {
      "Action": "CREATE",
      "ResourceRecordSet": {
        "Name": "\052.apps.<cluster_domain>",
        "Type": "A",
        "AliasTarget": {
          "HostedZoneId": "<hosted_zone_id>",
          "DNSName": "<external_ip>.",
          "EvaluateTargetHealth": false
        }
      }
    }
  ]
}'
```

1 For `<private_hosted_zone_id>`, specify the value from the output of the CloudFormation template for DNS and load balancing.

2 For `<cluster_domain>`, specify the domain or subdomain that you use with your OpenShift Container Platform cluster.

3 For `<hosted_zone_id>`, specify the public hosted zone ID for the load balancer that you obtained.

4 For `<external_ip>`, specify the value of the external IP address of the Ingress Operator load balancer. Ensure that you include the trailing period (\.) in this parameter value.
For `<hosted_zone_id>`, specify the public hosted zone ID for the load balancer that you obtained.

For `<external_ip>`, specify the value of the external IP address of the Ingress Operator load balancer. Ensure that you include the trailing period (.) in this parameter value.

1.9.21. Completing an AWS installation on user-provisioned infrastructure

After you start the OpenShift Container Platform installation on Amazon Web Service (AWS) user-provisioned infrastructure, monitor the deployment to completion.

**Prerequisites**

- You removed the bootstrap node for an OpenShift Container Platform cluster on user-provisioned AWS infrastructure.
- You installed the `oc` CLI.

**Procedure**

1. From the directory that contains the installation program, complete the cluster installation:

   ```
   $ ./openshift-install --dir <installation_directory> wait-for install-complete
   ```

   For `<installation_directory>`, specify the path to the directory that you stored the installation files in.

**Example output**

```
INFO Waiting up to 40m0s for the cluster at https://api.mycluster.example.com:6443 to initialize...
INFO Waiting up to 10m0s for the openshift-console route to be created...
INFO Install complete!
INFO To access the cluster as the system:admin user when using 'oc', run 'export KUBECONFIG=/home/myuser/install_dir/auth/kubeconfig'
INFO Access the OpenShift web-console here: https://console-openshift-console.apps.mycluster.example.com
INFO Login to the console with user: "kubeadmin", and password: "4vYBz-Fe5en-ymBEc-Wt6NL"
INFO Time elapsed: 1s
```

**IMPORTANT**

The Ignition config files that the installation program generates contain certificates that expire after 24 hours, which are then renewed at that time. If the cluster is shut down before renewing the certificates and the cluster is later restarted after the 24 hours have elapsed, the cluster automatically recovers the expired certificates. The exception is that you must manually approve the pending `node-bootstrapper` certificate signing requests (CSRs) to recover `kubelet` certificates. See the documentation for *Recovering from expired control plane certificates* for more information.
2. Register your cluster on the Cluster registration page.

1.9.22. Logging in to the cluster by using the web console

The kubeadmin user exists by default after an OpenShift Container Platform installation. You can log in to your cluster as the kubeadmin user by using the OpenShift Container Platform web console.

Prerequisites

- You have access to the installation host.
- You completed a cluster installation and all cluster Operators are available.

Procedure

1. Obtain the password for the kubeadmin user from the kubeadmin-password file on the installation host:

   ```
   $ cat <installation_directory>/auth/kubeadmin-password
   ```

   **NOTE**
   Alternatively, you can obtain the kubeadmin password from the `<installation_directory>/openshift_install.log` log file on the installation host.

2. List the OpenShift Container Platform web console route:

   ```
   $ oc get routes -n openshift-console | grep 'console-openshift'
   ```

   **NOTE**
   Alternatively, you can obtain the OpenShift Container Platform route from the `<installation_directory>/openshift_install.log` log file on the installation host.

   **Example output**

   ```
   console     console-openshift-console.apps.<cluster_name>.<base_domain>            console
   https   reencrypt/Redirect   None
   ```

3. Navigate to the route detailed in the output of the preceding command in a web browser and log in as the kubeadmin user.

Additional resources

- See Accessing the web console for more details about accessing and understanding the OpenShift Container Platform web console.

1.9.23. Telemetry access for OpenShift Container Platform

In OpenShift Container Platform 4.6, the Telemetry service, which runs by default to provide metrics about cluster health and the success of updates, requires Internet access. If your cluster is connected to the Internet, Telemetry runs automatically, and your cluster is registered to the Red Hat OpenShift
After you confirm that your Red Hat OpenShift Cluster Manager inventory is correct, either maintained automatically by Telemetry or manually by using OCM, use subscription watch to track your OpenShift Container Platform subscriptions at the account or multi-cluster level.

Additional resources

- See About remote health monitoring for more information about the Telemetry service

1.9.24. Additional resources

- See Working with stacks in the AWS documentation for more information about AWS CloudFormation stacks.

1.9.25. Next steps

- Validate an installation.
- Customize your cluster.
- Configure image streams for the Cluster Samples Operator and the must-gather tool.
- Learn how to use Operator Lifecycle Manager (OLM) on restricted networks.
- If the mirror registry that you used to install your cluster has a trusted CA, add it to the cluster by configuring additional trust stores.
- If necessary, you can opt out of remote health reporting.
- If necessary, you can remove cloud provider credentials.

1.10. UNINSTALLING A CLUSTER ON AWS

You can remove a cluster that you deployed to Amazon Web Services (AWS).

1.10.1. Removing a cluster that uses installer-provisioned infrastructure

You can remove a cluster that uses installer-provisioned infrastructure from your cloud.

NOTE

After uninstallation, check your cloud provider for any resources not removed properly, especially with User Provisioned Infrastructure (UPI) clusters. There might be resources that the installer did not create or that the installer is unable to access.

Prerequisites

- Have a copy of the installation program that you used to deploy the cluster.
- Have the files that the installation program generated when you created your cluster.

Procedure
1. From the directory that contains the installation program on the computer that you used to install the cluster, run the following command:

```
$ ./openshift-install destroy cluster \
--dir <installation_directory> --log-level info
```

1. For `<installation_directory>`, specify the path to the directory that you stored the installation files in.

2. To view different details, specify `warn`, `debug`, or `error` instead of `info`.

**NOTE**

You must specify the directory that contains the cluster definition files for your cluster. The installation program requires the `metadata.json` file in this directory to delete the cluster.

2. Optional: Delete the `<installation_directory>` directory and the OpenShift Container Platform installation program.