Private Cluster Support (Preview)

Date published: 2022-01-06
Date modified: 2023-08-02

Legal Notice

The documentation is and contains Cloudera proprietary information protected by copyright and other intellectual property rights. No license under copyright or any other intellectual property right is granted herein.

Copyright information for Cloudera software may be found within the documentation accompanying each component in a particular release.

Cloudera software includes software from various open source or other third party projects, and may be released under the Apache Software License 2.0 (“ASLv2”), the Affero General Public License version 3 (AGPLv3), or other license terms.

Other software included may be released under the terms of alternative open source licenses. Please review the license and notice files accompanying the software for additional licensing information.

Please visit the Cloudera software product page for more information on Cloudera software. For more information on Cloudera support services, please visit either the Support or Sales page. Feel free to contact us directly to discuss your specific needs.

Cloudera reserves the right to change any products at any time, and without notice. Cloudera assumes no responsibility nor liability arising from the use of products, except as expressly agreed to in writing by Cloudera.

Cloudera, Cloudera Altus, HUE, Impala, Cloudera Impala, and other Cloudera marks are registered or unregistered trademarks in the United States and other countries. All other trademarks are the property of their respective owners. Disclaimer: EXCEPT AS EXPRESSLY PROVIDED IN A WRITTEN AGREEMENT WITH CLOUDERA, CLOUDERA DOES NOT MAKE NOR GIVE ANY REPRESENTATION, WARRANTY, NOR COVENANT OF ANY KIND, WHETHER EXPRESS OR IMPLIED, IN CONNECTION WITH CLOUDERA TECHNOLOGY OR RELATED SUPPORT PROVIDED IN CONNECTION THEREWITH. CLOUDERA DOES NOT WARRANT THAT CLOUDERA PRODUCTS NOR SOFTWARE WILL OPERATE UNINTERRUPTED NOR THAT IT WILL BE FREE FROM DEFECTS NOR ERRORS, THAT IT WILL PROTECT YOUR DATA FROM LOSS, CORRUPTION NOR UNAVAILABILITY, NOR THAT IT WILL MEET ALL OF CUSTOMER’S BUSINESS REQUIREMENTS. WITHOUT LIMITING THE FOREGOING, AND TO THE MAXIMUM EXTENT PERMITTED BY APPLICABLE LAW, CLOUDERA EXPRESSLY DISCLAIMS ANY AND ALL IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO IMPLIED WARRANTIES OF MERCHANTABILITY, QUALITY, NON-INFRINGEMENT, TITLE, AND FITNESS FOR A PARTICULAR PURPOSE AND ANY REPRESENTATION, WARRANTY, OR COVENANT BASED ON COURSE OF DEALING OR USAGE IN TRADE.

Contents

This document has been released as part of a technical preview for features described herein. Technical preview components are provided as a convenience to our customers for their evaluation and trial usage. These components are provided ‘as is’ without warranty or support. Further, Cloudera assumes no liability for the usage of technical preview components, which should be used by customers at their own risk.

Private Cluster Support

Each type of network architecture supported by CDP has a unique set of tradeoffs among ease of setup, security, workloads (Experiences) supported, and so on.

Private Clusters, which are available on both AWS (GA) and Azure (Preview), provide a simple way to create a secure cluster, where the API server and the workloads themselves only rely on private IP addresses that are not accessible from the internet. Connectivity to the cluster from the CDP control place is provided by the Cluster Connectivity Manager v2 (CCM v2). CCMv2 uses an agent running in the cluster, and an inverting proxy running on CDP, which creates a HTTPS tunnel between the workload and the control plane.

Requirements

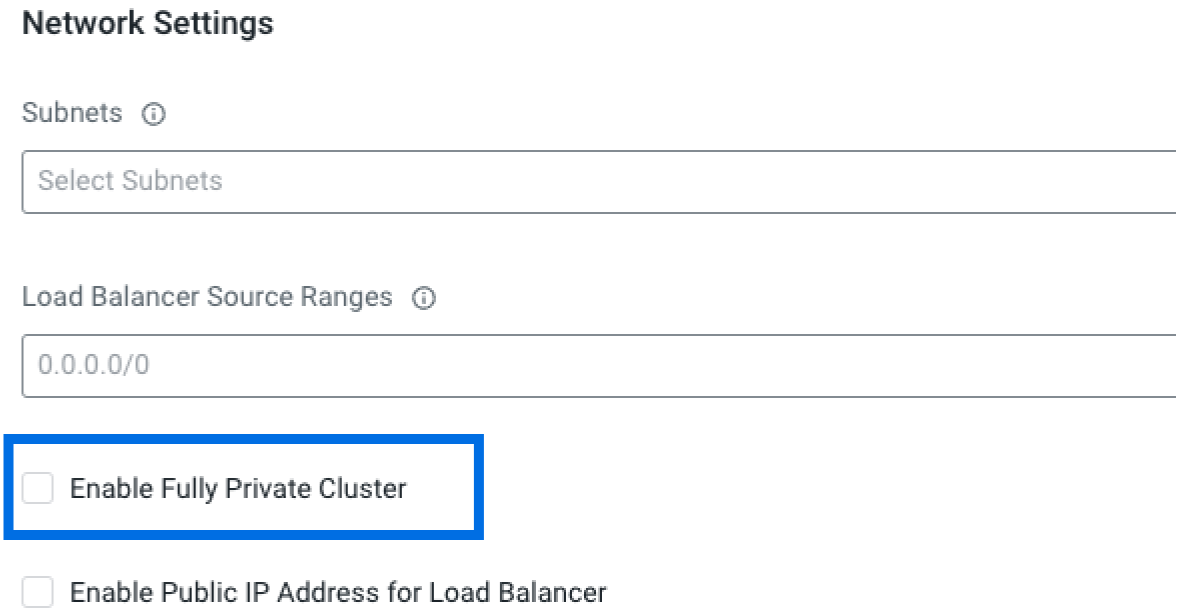
Ensure the following entitlements are needed to enable provisioning of private Kubernetes clusters. Customers should file SRE Jira tickets to request these entitlements, if necessary.

- CDP_CCM_V2_JUMPGATE
- CCMV2 JUMPGATE
- ML_ENABLE_PRIVATE_CLUSTER

Enable the Private Cluster

To enable a private cluster, select the option when provisioning the workspace.

1. In **ML Workspaces**, select **Provision Workspace.**
2. Enter a **Workspace Name**, and select **Environment**.
3. Select the **Advanced Options** toggle.
4. In **Network Settings**, select **Enable Fully Private Cluster**.
5. Make any other settings needed, and select **Provision Workspace**.



Network Settings

Subnets ⓘ

Select Subnets

Load Balancer Source Ranges ⓘ

0.0.0.0/0

□ Enable Fully Private Cluster

□ Enable Public IP Address for Load Balancer

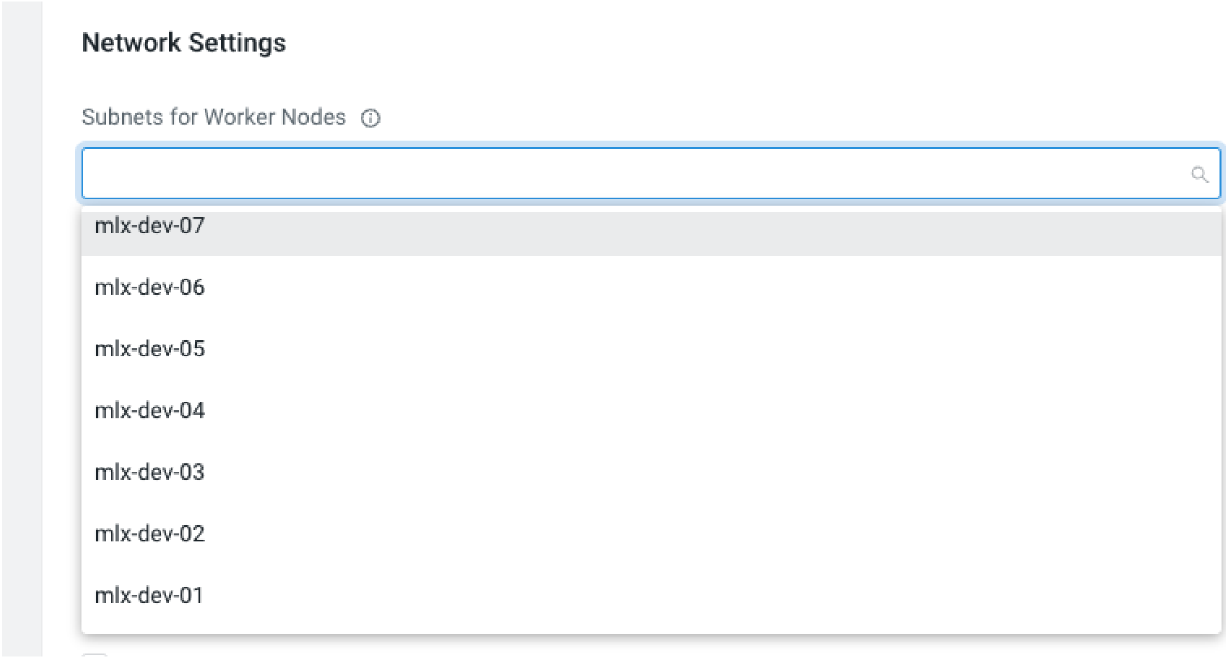
The workspace is provisioned using a fully private cluster.

User Defined Routing (Azure only)

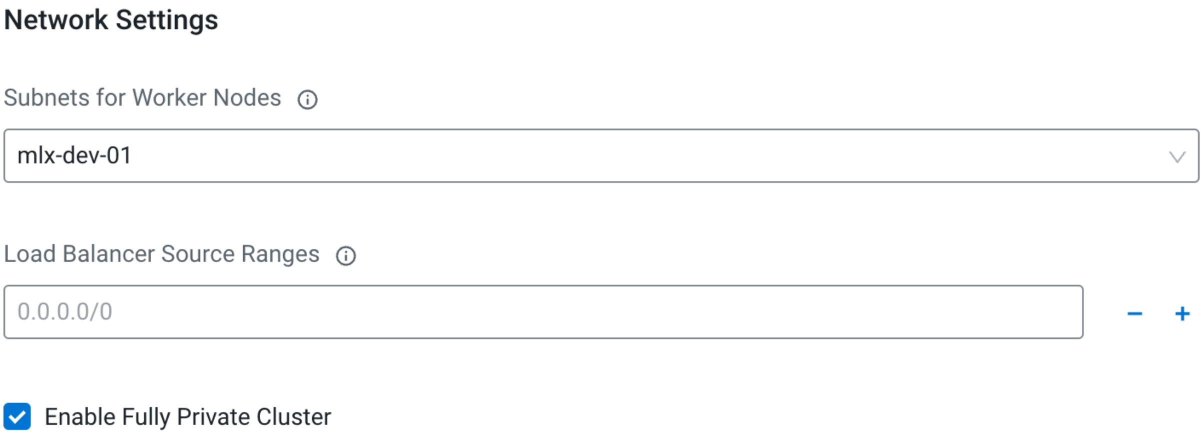
With the Fully Private Cluster configuration, Azure still creates some public IP resources to support load balancer egress. If necessary, you can avoid creating public IP addresses in the CML cluster by using a User Defined Routing (UDR) table. A UDR table can be configured in the cluster subnet to route packets to a customer-configured firewall, for example to limit internet access or analyze traffic. For more information on setting up UDR, see the Microsoft articles Virtual appliance scenario or Virtual network traffic routing.

To utilize a UDR and firewall in the Azure CML private cluster, select the following when setting up the cluster.

1. Select a subnet with a default route configuration to forward the traffic to the network appliance or firewall.



2. Create load balancers with private IP addresses. This is the default choice when creating clusters in CML.



Network Settings

Subnets for Worker Nodes

mlx-dev-01

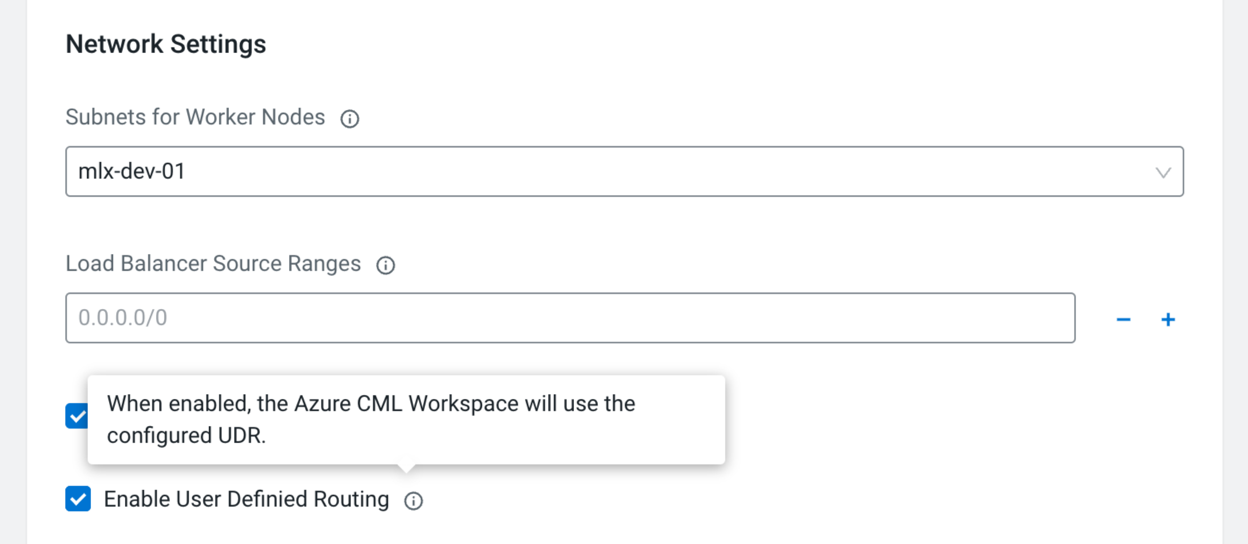
Load Balancer Source Ranges

0.0.0.0/0

☑ Enable Fully Private Cluster

This document has been released as part of a technical preview for features described herein. Technical preview components are provided as a convenience to our customers for their evaluation and trial usage. These components are provided 'as is' without warranty or support. Further, Cloudera assumes no liability for the usage of technical preview components, which should be used by customers at their own risk.

3. Select **Enable User Defined Routing**.



Network Settings

Subnets for Worker Nodes: mlx-dev-01

Load Balancer Source Ranges: 0.0.0.0/0

When enabled, the Azure CML Workspace will use the configured UDR.

☑ Enable User Definied Routing

Limitations of private cluster configuration in Azure

There are several limitations to note:

- **While using a Private cluster**, you must disable the network acceleration in the freeIPA VM. In the Azure portal, go to **Virtual machines**, and identify the freeIPA VM (usually, <environment name>-freeipa<random string>). In **Settings** > **Networking**, open the network interface, select **Edit accelerated networking**, and disable it.
- There is a limitation in Azure (tracking# 2212090040005952) where more than 15 tags fail private-DNS-Zone provisioning. Because of this limitation, CML won't add the following tags to the cluster during cluster provisioning. If necessary, you can add them manually to the resources in the managed cluster.
    - WorkspaceCrn
    - TenantID
    - Any custom tags that are configured during the CML cluster provision.
- If a customer script uses **tenantID**, it needs to be derived from one of the CRN tags.
- Instead of workspace CRN, the workspace name tag needs to be used.
- No tenant tags can be configured in the environment service.
- If the cluster creation fails due to TCP issues, customers should delete the cluster, clean up the cloud resources and retry the cluster creation.

Creating a Model Registry on an Azure UDR Private Cluster

Use the following template CDP CLI command to create a UDR private cluster on Azure with a Model Registry. You must replace the following template items with your own information.

- <environment CRN>
- <environment name> (in two places)
- <existing NFS name>
- <subnet>

For more information on Model Registry, see the Preview Feature document.

If you have not yet downloaded the CDP CLI tool, see the documentation.

The required CDP CLI version is version 0.9.93 or higher.

CDP CLI command

This CDP CLI command has three key sections:

1. Enables support for private clusters in Azure ("privateCluster": true,)
2. Enables UDR for the private cluster ("outboundTypes": ["OUTBOUND_TYPE_UDR"],)
3. Specifies the subnet for the UDR-enabled private cluster ("subnets")

```
cdp ml create-model-registry --cli-input-json '{
  "environmentCrn": "<environment CRN>",
  "environmentName": "<environment name>",
  "createWorkspacePayload": {
    "environmentName": "<environment name>",
    "workspaceName": "modelregistry",
    "privateCluster": true, # This setting enables the support for private cluster in azure.
    "outboundTypes": ["OUTBOUND_TYPE_UDR"], # Required for enabling UDR.
    "skipValidation": true,
    "disableTLS": false,
    "disableSSO": false,
    "existingNFS": "<existing NFS name>",
    "nfsVersion": "3",
    "xEntitlements": [
      "ML_MODEL_REGISTRY",
      "ML_ENABLE_PRIVATE_CLUSTER"
    ],
   "provisionK8sRequest": {
    "environmentName": "<environment name>",
    "network": {
     "topology": {
      "subnets": [
       "<subnet>" # subnet with a default route configuration to forward the traffic to the
network
appliance or firewall. This is required to enable UDR.
       ]
     }
    },
    "instanceGroups": [{
     "autoscaling": {
       "minInstances": 1,
       "maxInstances": 5
      },
      "instanceType": "Standard_DS5_v2",
      "rootVolume": {
       "size": 256
      }
     }
     ]
   }
  }
}' --profile eu-stage
```

Notes

- In a private cluster configuration, the Kubernetes API server has a private IP address, and it is not routable from the internet, thus reducing the attack surface for the cluster. If a customer wants to have access to the cluster, they need to peer their on-premises network to the CDP environment through VPN peering, for example.
- In a private cluster setup, Azure may still create public IP resources for egress. To avoid having any public IP addresses, enable User Defined Routing (UDR).
- An alternative configuration that is supported is to provide a non-transparent proxy server for the private cluster (for AWS only). For more information, see https://docs.cloudera.com/machine-learning/cloud/requirements-aws/topics/ml-non-transparent-proxy-aws.html

This document has been released as part of a technical preview for features described herein. Technical preview components are provided as a convenience to our customers for their evaluation and trial usage. These components are provided 'as is' without warranty or support. Further, Cloudera assumes no liability for the usage of technical preview components, which should be used by customers at their own risk.