Public Endpoint Access Gateway for Azure (Preview)

Date published: 2021-07-23
Date modified: 2021-07-27

Legal Notice

Contents

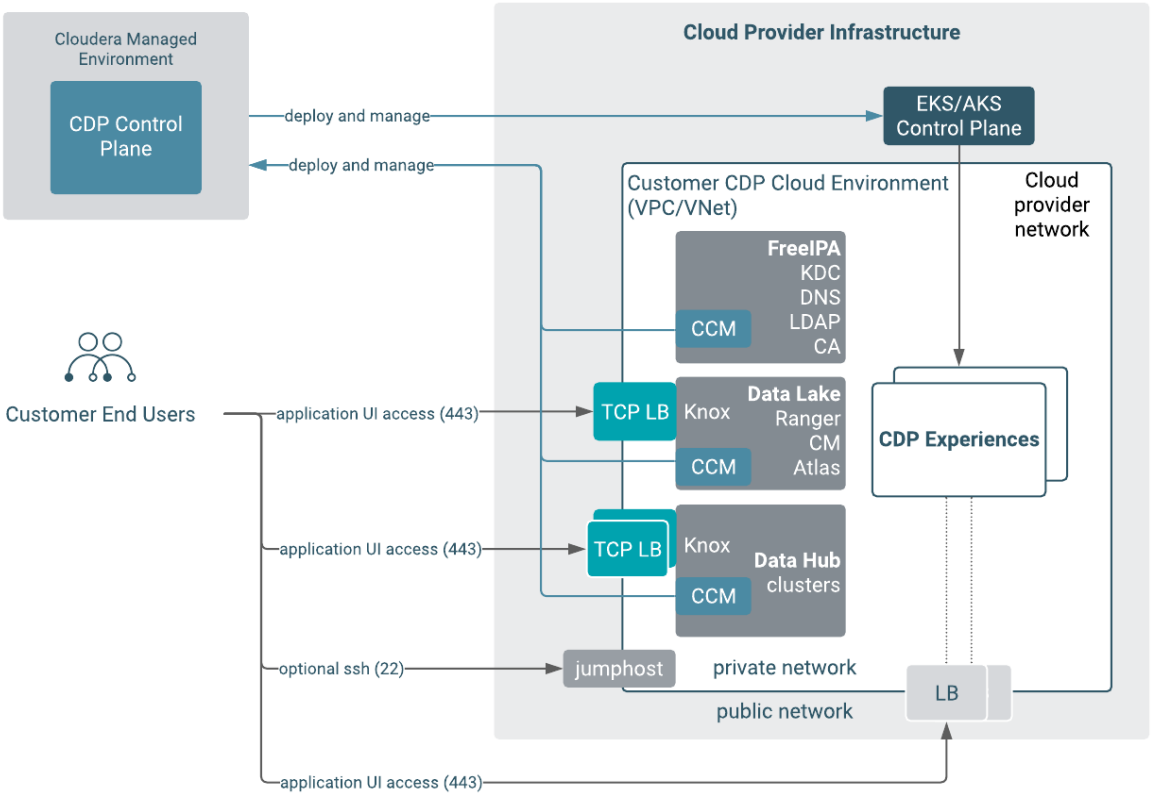
Public Endpoint Access Gateway for Azure

If the network into which you are deploying your CDP environment does not have pre-established connectivity with your corporate network, enabling the Public Access Gateway can reduce the complexity users face when interacting with the CDP endpoints.

The recommended way to deploy production-ready CDP environments is to deploy them on private networks, but this additional security makes it difficult for users to access UIs and APIs without configuring complex network connectivity between users and internal cloud provider networks. The Public Endpoint Access Gateway provides secure connectivity to UIs and APIs in Data Lake and Data Hub clusters deployed using private networking, allowing users to access these resources without complex changes to their networking or creating direct connections to cloud provider networks.

You can enable the Public Endpoint Access Gateway when registering your Azure environment in CDP. The gateway interfaces the Knox service, which is automatically integrated with your identity provider configured in CDP, allowing you to authenticate using your SSO credentials without any additional configuration. All communication with the gateway is over TLS, so connections are secure. You can control the IP ranges from where connections to the gateway can be established by configuring your security groups.

The following diagram illustrates this setup:



Note: The gateway provides secure connectivity to UIs and APIs. All Knox-enabled endpoints are supported. The gateway does not cover SSH or data access paths (such as Kafka Broker and NiFi Site2Site endpoints). Cloudera recommends that you set up connectivity between private networks in the public cloud and internal customer networks for secure and fast Kafka and NiFi deployments.

Enabling Public Endpoint Access Gateway for Azure

You can enable Public Endpoint Access Gateway for Azure during Azure environment registration after enabling Cluster Connectivity Manager (CCM). Once activated, the gateway will be used for the Data Lake and all the Data Hubs within the environment. There is no way to activate it on a per Data Lake or per Data Hub level. Once it is enabled for an environment, there is no way to deactivate it. The gateway can be used either with an existing VNet or with a new VNet created by CDP.

If you choose to enable Public Endpoint Access Gateway, CDP will create two Azure load balancers per cluster (that is, two for each Data Lake and Data Hub).

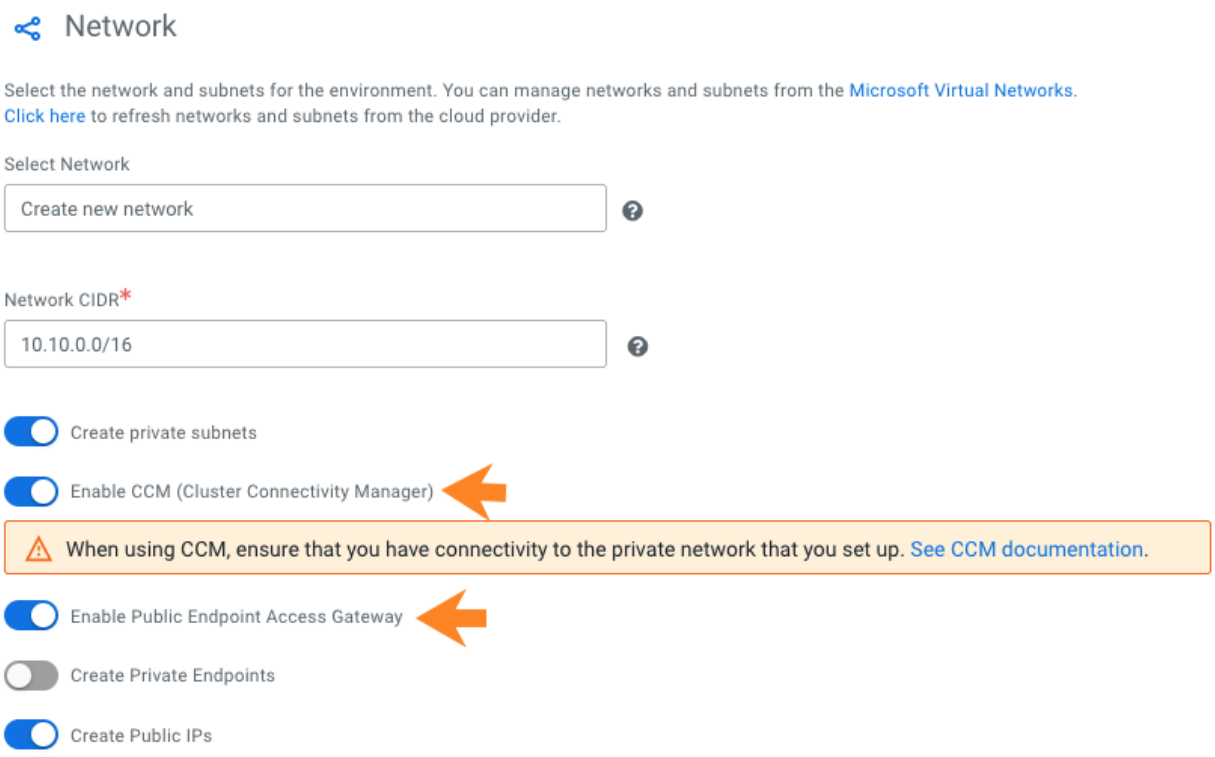
Steps: CDP web interface

When registering your Azure environment, make sure to do the following:

1. On the **Region, Networking, and Security** page, select your existing VNet or select to have a new VNet created.
2. If you selected an existing VNet, select at least one existing private subnet (or at least three subnets if you would like to provision Data Warehouse instances).
3. The **Enable Cluster Connectivity Manager** option is enabled by default to enable communication via private subnets.

   Note: In spite of the warning that you see on the UI, with the Public Endpoint Gateway enabled, you do not need to set up any additional connectivity in order to use CCM.

4. Click on **Enable Public Endpoint Access Gateway** to enable it. This enables UIs and APIs of the Data Lake and Data Hub clusters to be accessible over the internet.



5. Under **Security Access Settings**, make sure to restrict access to only be accepted from sources coming from your external network range.

   Note: The security access settings do not apply to the load balancer used by the Public Endpoint Access Gateway, but they apply to the instances that are running in private subnets and to which the Public Endpoint Access Gateway routes traffic. Therefore the

security access settings should allow the users' public IP ranges to be able to connect through the public load balancer.

6. Finish registering your environment.

Steps: CDP CLI

During Azure environment registration via CDP CLI, you can optionally enable public endpoint access gateway using the following CLI parameter:

```
--endpoint-access-gateway-scheme PUBLIC
```

When enabling Endpoint Access Gateway, you should also enable CCM using the `--enable-tunnel` parameter.

For example:

```
cdp environments create-azure-environment
...
--enable-tunnel
--endpoint-access-gateway-scheme PUBLIC
```

Equivalent CLI JSON for an environment request looks like this:

```
cdp environments create-azure-environment
...
"enableTunnel": true,
"endpointAccessGatewayScheme": "PUBLIC"
```