Dedicated Cluster and Private Link

Set up and configure a dedicated cluster

# Objectives

Dedicated clusters are designed for production-ready functionality that requires private endpoint networking capabilities. In this lab, you will create a dedicated cluster and connect it via a private link to an AWS VPC.

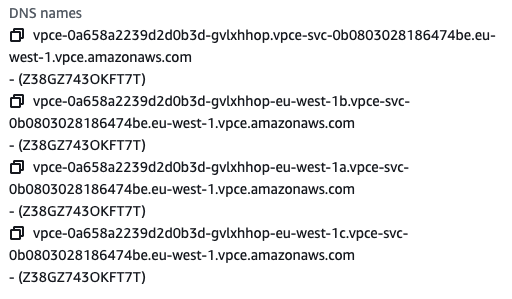
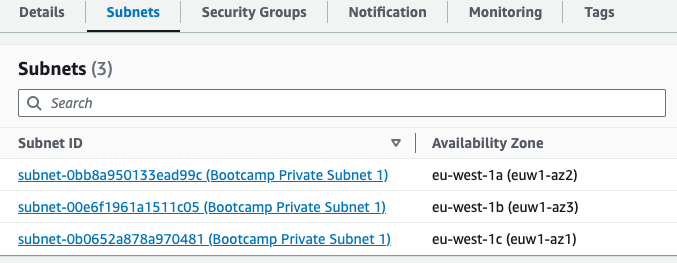
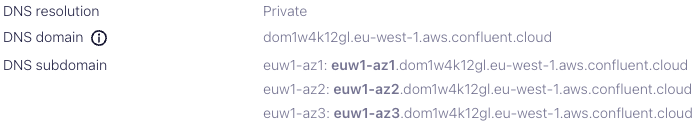
# Labs

## Create or verify your AWS VPC

* You can reuse the VPC you utilised in [Lab 7: Enterprise Cluster](https://docs.google.com/document/u/1/d/1KmredwALkwNgiW3-rnVXPi6zcoAcX_cKs98TU-SExn0/edit). Verify you can still access the jumphost in that VPC.

## Connect the cluster to AWS

You should have a running cluster you created at the beginning of the previous lab. Now, it is time to connect this cluster to your VPC.

* In your new cluster, navigate to the Network Management panel (it will be marked with an exclamation mark and a yellow triangle)
* Create an Ingress connection
* Create a PrivateLink access
* Enter the 12-digit AWS Account Number of your AWS account. You can find this on the top right of your AWS console.
* Copy the VPC endpoint service name somewhere safe (like a local editor). You will need this later (but it is recoverable, so do not worry if you forgot it).
* Click **Add**
* This process can take a few minutes. Before you can continue, you need to wait until the connection status transitions from “Pending” to “Ready.”
* All the remaining required steps can be found in the documentation. You will need to perform the following manual steps:
  + Create a VPC Endpoint and link it up with the Confluent Cloud endpoint, following the same steps as you did for the Enterprise cluster.
    - Give the endpoint a name
    - Choose the type “Endpoint services that use NLBs and GWLBs”
    - Ensure you pick the correct VPC.
    - Pick your three Availability Zones
    - Pick the private subnets
    - Create or use a security group with permission to access the endpoint from your jumphost, at least on TCP ports 80, 443 and 9092. If you used **simple-vpc**, you can use **all-bootcamp-sg**.
    - Press Create Endpoint
  + Now, you need to set up the [DNS](https://docs.confluent.io/cloud/current/networking/private-links/aws-privatelink.html#set-up-dns-records-in-aws) records, which will be tricky if done manually. Remember that you are using a private DNS resolution.   
    The interesting part is the naming of the zones, which is not consistent across different regions.   
    Look at the Details page of the endpoint you just created. You should see DNS names like the following:  
      
    When you open the Subnet tab, you can see the mapping to the canonical zone names  
      
      
    On the details page of your Networking Management in the Confluent Cloud, you see the DNS names for the different zones:  
      
    
  + Your job will be to match the DNS names to the correct endpoint in the Route 53 definition. You will need 4 Records, one for the DNS Domain, and 3 for the DNS subdomains.

## Test your cluster

* SSH into your jumphost in the VPC you created and to which you set up the private link endpoint.
* Verify that you can connect to your new cluster from your jumphost.
  + Use openssl s\_client -connect <your-kafka-endpoint> for a first test
  + Create a client configuration for a C++ client including an API Key
  + Install kafkacat on your jumphost. If you used Ubuntu, you can use  
    sudo apt install kafkacat.
  + Test the connection with kafkacat -F <your-property-file> -L.
  + Install Java (11 or 17) on your jumphost.
  + Install the Kafka Java Clients (kafka-topics, kafka-console-producer, etc.) on your jumphost
  + Create a client configuration for Java
  + Create a topic, the produce and consume from it on your Enterprise cluster.

## Proxy configuration

If you navigate in the UI to the topic, you can see that the topics are not accessible via the console. The problem is that the console needs to communicate with the cluster in a different network, and this network is only accessible from your VPC via the private endpoint.

In a production environment, a network administrator could set up a VPN to your private VPC and set up a DNS entry to reroute the traffic. For this bootcamp, we can use a workaround via Proxy.

Either:

* Install a proxy service on your jumphost, for example, haproxy. Configure this proxy to point to your Kafka cluster. The documentation below will help you figure out the required settings.

Or:

* Alternatively, you can use a dynamic (SOCKS v5) proxy.
  + For this, you need to set up a dynamic proxy by adding the option “-D port” to your SSH command when connecting to your jumphost, for example:  
      
    ssh -D 8081 jumphost
  + Install a dynamic proxy in your browser, for example, FoxyProxy (<https://getfoxyproxy.org/>).
  + Configure the Proxy to point to localhost:port (here, port 8081) with a URL pattern that matches your cluster and a regular expression at the beginning and end, for example like this  
      
    [https://.\*lkc-m2yg51.dom1w4k12gl.eu-west-1.aws.confluent.cloud\*](about:blank)  
    (Do not forget to enable your Proxy for your Patterns)
* Test the proxy solution with your cluster
* Leave the cluster running after completing this lab; we will also need it for the next three labs.

# References

<https://docs.confluent.io/cloud/current/clusters/cluster-types.html#enterprise-clusters>

<https://docs.confluent.io/cloud/current/networking/private-links/aws-privatelink.html#use-aws-privatelink-with-ccloud>

<https://docs.confluent.io/platform/current/installation/installing_cp/zip-tar.html#get-the-software>

<https://docs.confluent.io/cloud/current/networking/ccloud-console-access.html#configure-a-proxy>

# Expected Outcomes

Connect a dedicated cluster to your VPC via PrivateLink to an endpoint.

Successfully set up the DNS records to be able to access the cluster from a jumphost within the VPC.

Verify access via OpenSSL and Kafka tools.

# Check your understanding

This colour marks advanced questions.

* How does an Enterprise cluster differ from a dedicated cluster?
  + Are there any similarities?
* Which networking options are available for a dedicated cluster?
* Why would you pick a dedicated cluster over an Enterprise cluster?

# 