**Day 9 Hands on**

**Name : Podutur Lahari**

**Date : 19-11-2024**

To open a cluster in Databricks, follow these steps:

**Step 1: Log in to Databricks**

1. Navigate to the Databricks platform URL for your organization (e.g., https://<region>.azuredatabricks.net for Azure or https://<workspace>.cloud.databricks.com for AWS).
2. Log in using your username and password or through your SSO credentials if Single Sign-On is enabled.

**Step 2: Navigate to the Compute Tab**

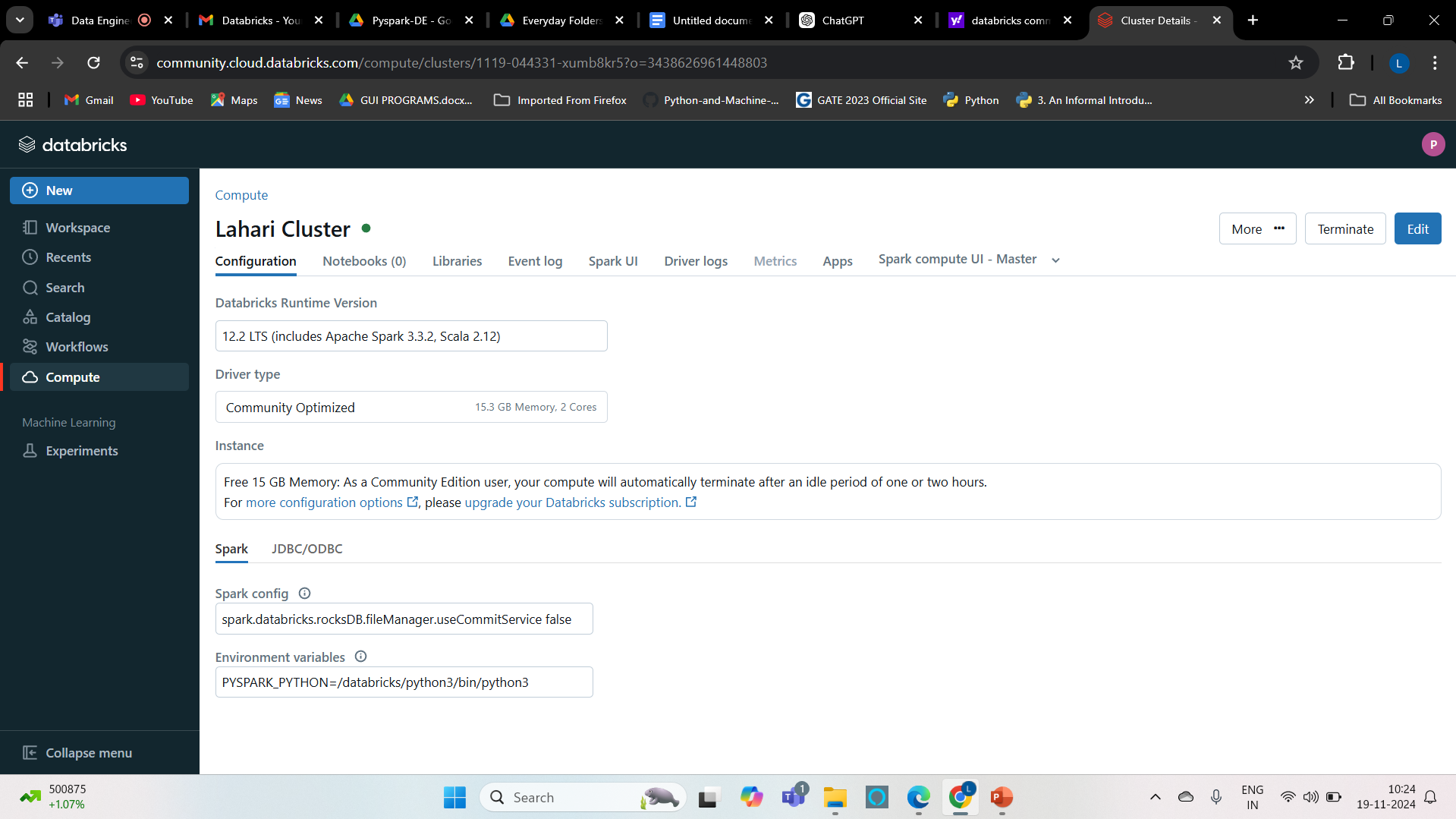
1. Once logged in, you will land on the Databricks workspace.
2. From the left-hand navigation menu, click on "Compute". This tab is where you can manage your clusters.

**Step 3: Create or Select a Cluster**

To Open an Existing Cluster:

1. Under the Compute tab, you will see a list of existing clusters.
2. Locate the cluster you want to use and check its Status.
   * If the cluster is running, you can open it by clicking on its name.
   * If the cluster is stopped, click on the Start button next to its name to initiate it.

To Create a New Cluster:

1. Click on the Create Cluster button at the top of the page.
2. Fill in the following details:
   * Cluster Name: Provide a name for the cluster.
   * Cluster Mode: Select either Single Node, Standard, or High Concurrency.
   * Databricks Runtime Version: Choose the appropriate runtime version for your workload.
   * Autoscaling: Enable or disable autoscaling for worker nodes.
   * Node Type: Select the type of virtual machines or instances to use.
   * Worker and Driver Settings: Specify the number of worker nodes and driver configurations.
3. Click on Create Cluster to start the cluster.

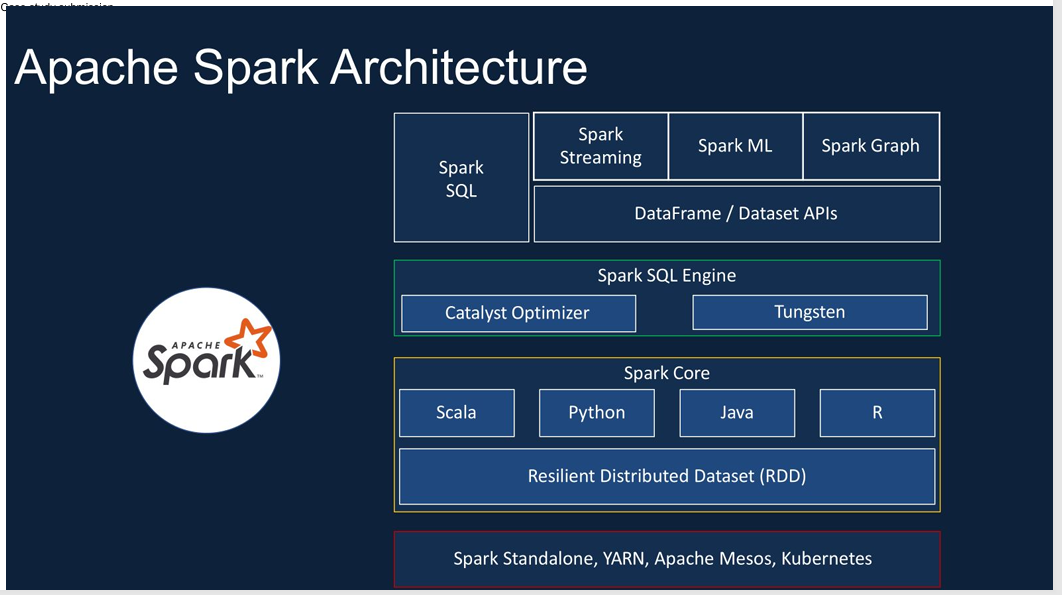
**Step 4: Open the Cluster**

1. Once the cluster is running, click on its name in the list to open the cluster's detail page.
2. On the detail page, you can:
   * View cluster configurations and logs.
   * Access attached libraries and notebooks.
   * Monitor cluster metrics like memory and CPU usage.

**Step 5: Attach a Notebook (Optional)**

1. To work with the cluster, attach a notebook by opening the notebook you want to use.
2. From the notebook, click on the dropdown in the top-right corner and select the cluster to attach to.

**Apache Spark Architecture**



**Spark SQL Architecture**