**Databricks Job And Storage Accounts Creation(Blob and Datalake)**

**Name: Podutur Lahari – DE126**

**Date: 03-12-2024**

**Creating Job in Azure Databricks**

To create a job in Azure Databricks, you can automate tasks such as running notebooks, JARs, Python scripts, or other workloads. Here are the detailed steps:

**Steps to Create a Job in Azure Databricks**

**1. Launch Azure Databricks Workspace**

1. **Log in to the Azure Portal**: Navigate to [Azure Portal](https://portal.azure.com).
2. **Open the Databricks Service**:
   * Go to your resource group or search for the Databricks workspace you created.
   * Click **Launch Workspace** to open the Azure Databricks environment.

**2. Create or Access a Notebook**

1. In the Azure Databricks workspace:
   * Go to the **Workspace** tab on the left-hand menu.
   * Create a new notebook:
     + Right-click on any folder or click **Create** > **Notebook**.
     + Provide a name (e.g., my\_job\_notebook) and choose a language (e.g., Python, SQL, or Scala).
2. Write the code or logic you want the job to execute in this notebook.

**3. Configure a Cluster**

1. Ensure a cluster is available for running your job:
   * Go to the **Compute** section in the left-hand menu.
   * Either create a new cluster or select an existing cluster.
   * Start the cluster if it’s not already running.

**4. Create a Job**

1. Go to the **Workflows** tab in the left-hand menu.
2. Click **Create Job**.
3. Configure the job details:
   * **Name**: Provide a name for your job (e.g., MyDatabricksJob).
   * **Task Type**: Click **Add Task** and configure:
     + **Task name**: Name the task.
     + **Type**: Select the workload type (Notebook, JAR, Python script, etc.).
     + **Notebook path**: Browse and select the notebook you created earlier.
     + **Cluster**: Choose an existing cluster or attach a new one.
   * **Parameters** (optional): Provide parameters if required for the notebook or script.

**A screenshot of a computer

Description automatically generated**

A screenshot of a computer

Description automatically generated

**5. Set the Job Schedule (Optional)**

* To automate job execution:
  + Click on the **Schedule** tab.
  + Select **Add Schedule**.
  + Define the frequency (e.g., daily, weekly) and time for the job to run.

**6. Configure Alerts and Notifications**

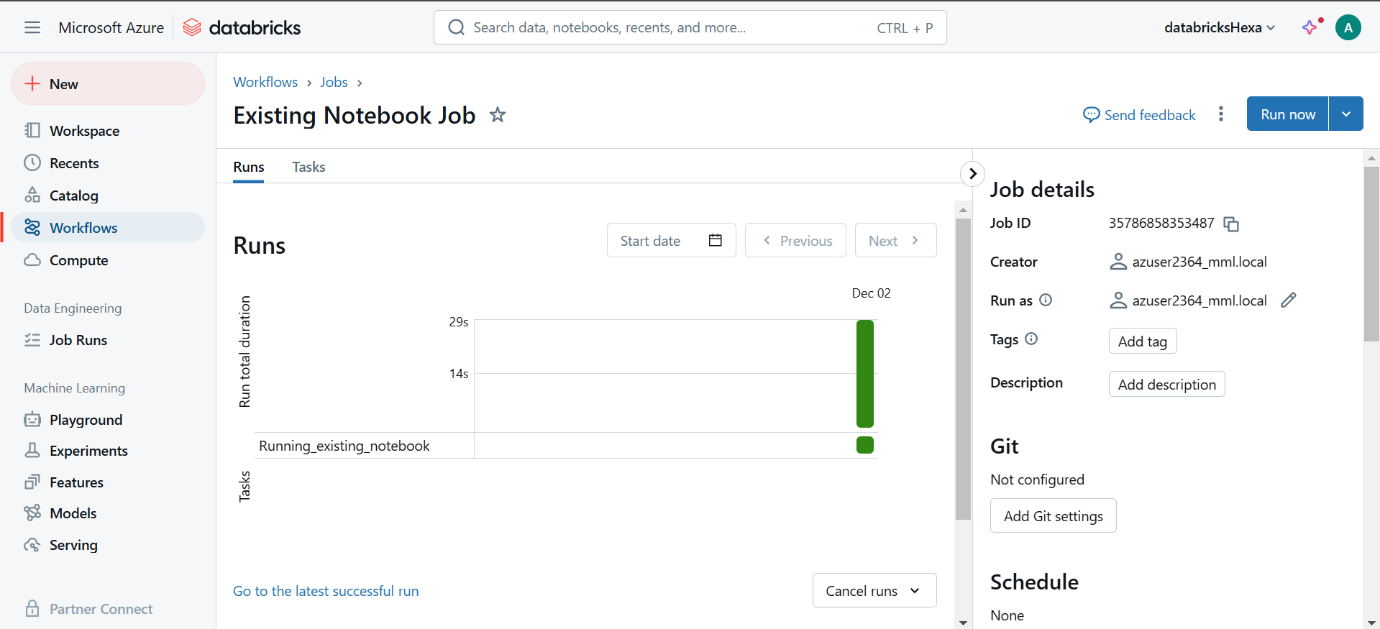
* To monitor job success or failure:
  + Click **Edit Notifications**.
  + Add email addresses or other alert mechanisms to notify recipients about job execution results.

A screenshot of a computer

Description automatically generated

**7. Save and Run the Job**

1. Click **Save** to finalize the job configuration.
2. To test the job, click **Run now**.
   * Monitor the execution in the **Runs** tab.
   * View the job logs and results in the **Run Details** section.



**Execution Results for Job: Sample Tutorial Job**

A screenshot of a computer

Description automatically generated

**Creating Data Storage (Blob Service) In azure Storage Accounts**

**Steps to Create Blob Storage in Azure:**

**1. Create an Azure Storage Account**

* Log in to the [Azure Portal](https://portal.azure.com).
* Select **Create a resource** > **Storage** > **Storage account**.
* Fill in the required details:
  + **Subscription**: Choose your Azure subscription.
  + **Resource group**: Select an existing resource group or create a new one.
  + **Storage account name**: Provide a unique name (e.g., mystorageaccount).
  + **Region**: Select the region closest to your location.
  + **Performance**: Choose **Standard** or **Premium** depending on your needs.
  + **Redundancy**: Select redundancy options (e.g., LRS, GRS).
* Click **Review + Create**, and then **Create**.

**2. Access the Storage Account**

* Once the storage account is created, navigate to it by clicking on its name.

**3. Create a Blob Container**

* In the left menu of the storage account, select **Containers**.
* Click **+ Container** to create a new container.
* Provide a name for the container (e.g., myblobcontainer).
* Choose the **public access level**:
  + **Private**: No public access.
  + **Blob**: Public read access for blobs only.
  + **Container**: Public read access for the entire container.
* Click **Create**.

A screenshot of a computer

Description automatically generated

**4. Upload a Blob**

* Open the container you just created.
* Click **Upload**.
* Browse and select a file to upload.
* Optionally, you can configure the blob tier (Hot, Cool, Archive).
* Click **Upload** to store the file in the blob container.

**5. Access Blob Data**

* To access the blob, use the **URL** provided in the blob properties.
* If the blob is private, you will need a Shared Access Signature (SAS) or credentials to access it.

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

**Creating Data Storage (Datalake Service) In azure Storage Accounts**

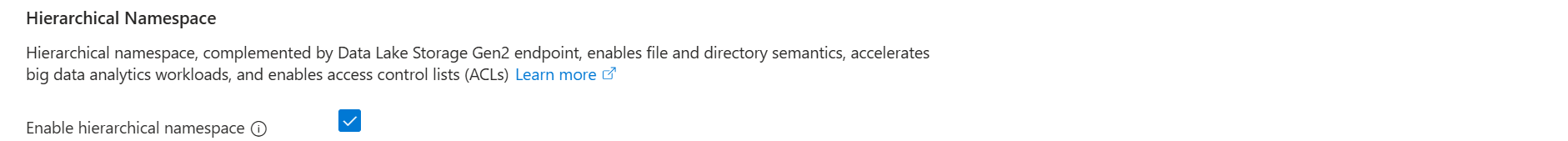
**Steps to Create Data Lake Storage with Access Tier Selection**

**1. Create an Azure Storage Account**

1. Log in to Azure Portal:
   * Go to [Azure Portal](https://portal.azure.com).
2. **Create a Storage Account:**
   * Click Create a resource > Storage > Storage account.
   * Fill in the required details:
     + Subscription: Select your subscription.
     + Resource group: Choose an existing group or create a new one.
     + Storage account name: Enter a globally unique name (e.g., mydatalakestorage).
     + Region: Choose a location.
     + Performance: Select Standard.
     + Redundancy: Choose a redundancy option (e.g., LRS or GRS).
   * Under the Advanced tab:
     + Enable Hierarchical namespace by toggling it to On.
   * Review and click Create.

A screenshot of a computer

Description automatically generated



**2. Access the Storage Account**

1. Navigate to the newly created storage account by selecting it from your resource list.
2. In the left-hand menu, click Data Lake Storage or Containers.

**3. Create a File System (Equivalent to a Blob Container)**

1. Under the Data Lake Storage Gen2 section, click File systems.
2. Click + File system to create a new file system.
3. Provide a name for the file system (e.g., myfilesystem).
4. Click OK to create the file system.

**4. Upload Files and Select Access Tier**

1. Open the newly created file system.
2. Click Upload to add files or folders.
3. Drag and drop files, or browse to select them.
4. Choose the Access Tier during upload:
   * Hot: Default for frequently accessed data.
   * Cool: For less frequently accessed data.
   * Archive: For rarely accessed data.
5. Click Upload to store the files.

A close-up of a text

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated