Spark integration

Integration of Spark with different metastores / catalogs and lakehouses

JERIN SAM

Table of Contents

[What is Spark, Lakehouse, Catalog and Object Storage 3](#_Toc189843534)

[Spark 3](#_Toc189843535)

[Key features 3](#_Toc189843536)

[Lakehouse 3](#_Toc189843537)

[Key features 3](#_Toc189843538)

[Catalog 3](#_Toc189843539)

[Key features 4](#_Toc189843540)

[Object Storage 4](#_Toc189843541)

[Key features 5](#_Toc189843542)

[Integration Architecture 6](#_Toc189843543)

[Wite Operations 6](#_Toc189843544)

[Read Operations 7](#_Toc189843545)

[Integrations 7](#_Toc189843546)

[Spark and Delta Lake Integration 7](#_Toc189843547)

[Services: 7](#_Toc189843548)

[Installation and Configuration 8](#_Toc189843549)

[Spark Session 11](#_Toc189843550)

[Config Testing Jupyter Notebook 14](#_Toc189843551)

[Spark, Delta Lake and Minio Integration 14](#_Toc189843552)

[Services: 14](#_Toc189843553)

[Installation and Configuration 14](#_Toc189843554)

[Spark Session 17](#_Toc189843555)

[Config Testing Jupyter Notebook 19](#_Toc189843556)

[Spark, Delta Lake and Unity Catalog Integration 19](#_Toc189843557)

[Services: 19](#_Toc189843558)

[Installation and Configuration 20](#_Toc189843559)

[Spark Session 24](#_Toc189843560)

[Config Testing Jupyter Notebook 27](#_Toc189843561)

[Spark, Delta Lake, Unity Catalog and Minio Integration 27](#_Toc189843562)

[Services: 27](#_Toc189843563)

[Installation and Configuration 27](#_Toc189843564)

[Integration Issue 30](#_Toc189843565)

[Config Testing Jupyter Notebook 33](#_Toc189843566)

[Spark, Delta Lake, Hive Metastore and Minio Integration 33](#_Toc189843567)

[Services: 33](#_Toc189843568)

[Installation and Configuration 33](#_Toc189843569)

[Spark Session 40](#_Toc189843570)

[Config Testing Jupyter Notebook 47](#_Toc189843571)

[UI for Hive Metastore 47](#_Toc189843572)

[Spark, Iceberg and Minio Integration 47](#_Toc189843573)

[Services: 47](#_Toc189843574)

[Installation and Configuration 47](#_Toc189843575)

[Spark Session 47](#_Toc189843576)

[Config Testing Jupyter Notebook 52](#_Toc189843577)

[Spark, Iceberg, Nessie and Minio Integration 52](#_Toc189843578)

[Services: 52](#_Toc189843579)

[Installation and Configuration 52](#_Toc189843580)

[Spark Session 55](#_Toc189843581)

[Config Testing Jupyter Notebook 61](#_Toc189843582)

[Appendix 61](#_Toc189843583)

[Enable SparkSQL Magic Command in Jupyter Notebook 61](#_Toc189843584)

[Unity Catalog 62](#_Toc189843585)

[Delta Lake Documentation 62](#_Toc189843586)

[Nessie Documentation 62](#_Toc189843587)

[Minio Documentation 62](#_Toc189843588)

[Docker Commands 63](#_Toc189843589)

# What is Spark, Lakehouse, Catalog and Object Storage

## Spark

Apache Spark is an open-source, distributed data processing framework designed for large-scale data analytics and processing.

### Key features

* **Unified Engine:** Supports batch processing, stream processing, machine learning, and graph processing.
* **In-Memory Processing**: Processes data in memory, making it much faster than traditional disk-based processing.
* **Distributed Computing**: Divides tasks across multiple nodes in a cluster for parallel execution.
* **Support for Multiple Languages**: Compatible with Python, Java, Scala, and R.
* **Extensive Libraries**: Includes built-in libraries such as Spark SQL, Spark Streaming, MLlib (for machine learning), and GraphX (for graph processing).

## Lakehouse

Lakehouse is a modern data architecture that combines the best features of a data lake and a data warehouse. It is designed to manage large volumes of structured and unstructured data while enabling analytics and reporting at scale.

### Key features

* **Unified Storage:** Stores raw and processed data in a Data Lake.
* **Support for Transactions**: Ensures ACID (Atomicity, Consistency, Isolation, Durability) compliance for reliable data operations.
* **Schema Enforcement**: Supports schema management and governance, similar to a data warehouse.
* **Data Lake Flexibility**: Retains the ability to handle unstructured and semi-structured data.
* **High Performance**: Optimized for querying and analytical workloads.

**Open-Source Lakehouse techs are –**

1. Delta Lake
2. Iceberg
3. Hudi

## Catalog

A **Catalog** is a service that manages metadata about datasets. It acts as a central repository that stores information about datasets, schemas, and other metadata.

### Key features

* **Metadata Management**: Stores information about tables, columns, data types, and more.
* **Data Governance**: Helps enforce security, access control, and compliance policies.
* **Integration**: Facilitates integration with processing engines like Spark or SQL engines like Trino.

**Open-Source Catalog services are –**

1. Unity Catalog – Primarily supports Delta Table but can provide interoperability for Apache Iceberg as well using UniForm feature.
2. Nessie – Supports Iceberg and Delta Table, provides git like branching feature for data.
3. Hive Metastore – Supports Iceberg and Delta Table.
4. Apache Polaris – Supports Iceberg.

**Following Catalogs support interoperability with all the 3 major lakehouses –**

1. Apache XTable
2. Unity Catalog

## Object Storage

***Object storage*** is a data storage architecture that manages data as objects, rather than in a hierarchical file system or as blocks within sectors.

Each object consists of:

* **Data** – The actual content being stored (e.g., images, videos, documents).
* **Metadata** – Descriptive information about the object (e.g., size, creation date, permissions).
* **Unique Identifier** – A globally unique ID that enables retrieval without a traditional file system.

Object storage is commonly used in cloud environments and is designed for scalability, durability, and accessibility over the internet.

|  |  |  |  |
| --- | --- | --- | --- |
| **Storage Type** | **How Data is Stored** | **How Data is Retrieved** | **Use Case** |
| **Object Storage** | Data is stored as self-contained **objects** (data + metadata + unique identifier) in a **flat namespace**. No folders or directories. | Objects are accessed using a unique identifier (e.g., via REST API). | Cloud storage, backups, big data, AI/ML, and web applications. |
| **File Storage** | Data is stored in a **hierarchical file system** with folders and directories (like in Windows or Linux file systems). | Files are accessed using paths (/folder1/folder2/file.txt). | Shared storage for applications, document management, and network file shares. |
| **Block Storage** | Data is broken into fixed-size **blocks** and stored on disks (like traditional hard drives or SSDs). | The operating system manages blocks and uses them to build file systems. | Databases, high-performance applications, and virtual machines. |

***Note:*** *MinIO supports hierarchical naming, but it does not use a hierarchical file system like traditional file storage*

### Key features

* **Scalability**
  + Can scale horizontally to store petabytes or exabytes of data across multiple distributed systems.
  + Designed for unstructured data like images, videos, and logs.
* **Durability and Redundancy**
  + Data is replicated across multiple locations (e.g., different data centers) to prevent data loss.
  + Uses erasure coding and replication for fault tolerance.
* **Metadata-Rich Storage**
  + Allows customizable metadata (key-value pairs) that can be used for indexing, searching, and managing objects.
  + Enhances data categorization and retrieval.
* **Global Accessibility**
  + Accessible via APIs (e.g., Amazon S3 API, Azure Blob Storage API, Google Cloud Storage API).
  + Supports HTTP/S-based access, making it ideal for cloud applications.
* **Immutable Data Storage**
  + Objects are immutable (cannot be modified after creation).
  + New versions can be created to support versioning and data retention policies.
* **Security and Access Control**
  + Provides encryption (both at rest and in transit).
  + Integrates with Identity and Access Management (IAM) for user authentication and access control.
* **Cost-Effective**
  + Uses tiered storage (e.g., hot, warm, and cold storage) to optimize costs based on access frequency.
  + No need for complex storage management as in block or file storage.
* **Easy Integration with Data Processing Tools**
  + Works well with Big Data (e.g., Hadoop, Spark), AI/ML workloads, and Data Lakes.
  + Can be used with event-driven architectures (e.g., serverless computing, IoT pipelines).

**Open-Source Object Store services are –**

* **MinIO** – Self-hosted S3-compatible object storage
* **Ceph** – Open-source distributed storage

# Integration Architecture

Integration architecture of Spark communicating with Catalog and Data Lake can be seen below.

A diagram of a catalog

Description automatically generated

## Wite Operations

When a ***Write Operation*** is executed in Spark, following processes are executed –

1. ***Interaction with the Catalog***

* **Metadata Registration:**
  + Spark registers the table's metadata (e.g., schema, location, partitioning information) in the configured catalog (e.g., Hive Metastore, Apache Nessie, Unity Catalog).
  + This metadata is stored in the catalog for table discovery, governance, and querying.
* **Storage Path:**
  + The table's storage location (data lake path) is usually specified in the CREATE TABLE statement or determined by the catalog's configuration.

1. ***Interaction with the Data Lake***

* **Data Location**
  + Spark does not store the actual data in the catalog. Instead, the data is written directly to the data lake.
* **Write Operation**
  + When data is inserted into the table (e.g., using INSERT INTO), Spark writes the data directly to the specified path in the data lake.
  + For *managed tables, the catalog manages the location*. For *external tables, the location is user-defined, and the catalog only stores the metadata.*

## Read Operations

When a ***Read Operation*** is executed in Spark, following processes are executed –

1. ***Metadata Retrieval from Catalog***

* Spark queries the catalog/metastore to fetch metadata about the table, such as:
  + **Schema**: Column names and data types.
  + **Storage Location**: The path in the data lake (e.g., s3://bucket/path or /hdfs/path).
  + **Partitioning Information**: Details of partition columns and their locations.
  + **File Format**: The format of the files (e.g., Parquet, ORC, Delta, etc.).
* The catalog/metastore does not contain the data itself, only the metadata required to locate and interpret the data.

1. ***Direct Data Access from Data Lake***
   * Once Spark retrieves the storage location from the catalog/metastore, it bypasses the catalog/metastore for data reading.
   * Spark communicates directly with the underlying data lake (e.g., HDFS, S3, ADLS, GCS) to access and process the data

# Integrations

## Spark and Delta Lake Integration

### Services:

* Spark Master
* Spark Worker
* Jupyter Lab

### Installation and Configuration

* **Installation and Config Folder** ***- /install-and-config/1-spark-delta-integration***

Docker is used to setup the spark and jupyter services. These files need to be accessed in the following order for better understanding–

* + *Dockerfile.spark*
  + *requirement.txt and spark-config*
  + *entrypoint.sh*
  + *docker-compose.yml*

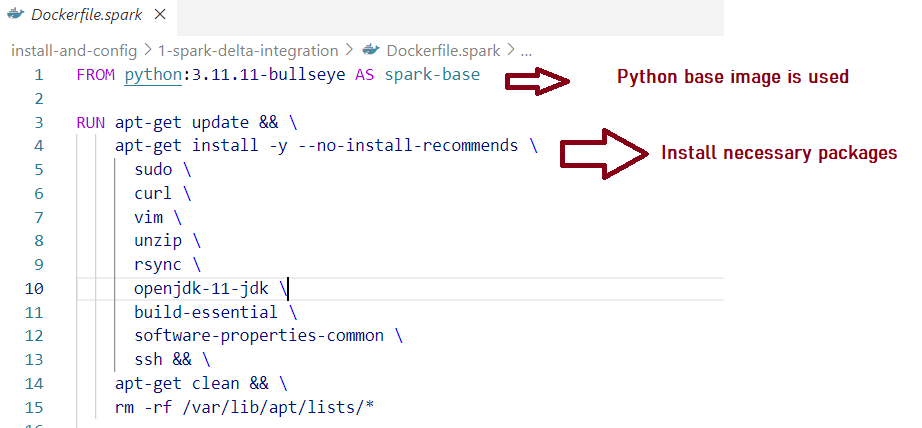
Below snippets shows details of some important scripts present in docker, entrypoint or notebook –

* ***Dockerfile.spark***

In this Dockerfile, Spark is configured with necessary details like Environment Variables, required python packages like jupyter and entrypoint file which needs to execute at the start of docker container.

The same Dockerfile will be used to setup *Spark Master and Worker docker containers.*

Following snippets from Dockerfile of Spark will cover only the important part of the Dockerfile scripts.



A screenshot of a computer program

Description automatically generated

A green arrow pointing at a line

Description automatically generated with medium confidence

A screen shot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

***spark-defaults.conf – In above snippet***

*The spark-defaults.conf file is used to define default configurations for Spark applications. It sets various runtime properties, including memory allocation, execution modes, and metastore connectivity.*

*Spark defaults file is updated with required details, when spark session is created then it will be used by default.*

A close-up of a sign

Description automatically generated

* ***entrypoint.sh***

Entrypoint file is used to execute the bash or shell scripts whenever the container starts.

In this entrypoint file, based on the parameter following scripts will be executed, these parameters are passed from *docker-compose.yml* file to *entrypoint* script. (**Note: Passing parameter from docker-compose file can be seen in the docker-compose snippet** **present below**) –

1. Parameter = ***master***

* Start spark master by executing spark master script.
* Create a pyspark session, which will start jupyter lab server. Token for the jupyter lab can be found by executing below script in spark master container.

|  |
| --- |
| ***jupyter server list*** |

1. Parameter = ***worker***

* Start spark worker by executing spark worker script.

1. Parameter = ***history***

* Start spark history server by executing spark history server script.

A screenshot of a computer error

Description automatically generated

A screenshot of a computer

Description automatically generated

* ***Docker-compose.yml***

***Docker Compose*** is used to define and manage multi-container Docker applications.

This helps to orchestrates multiple containers as part of a single application.

Instead of managing each container individually, manage all of them with one command.

A screenshot of a computer screen

Description automatically generated

In docker container volume section, all the logs generated from entrypoint scripts, data and Spark notebooks are synced between Container and Host Machine.

***Note: Other Services are also created in the similar method shown in the above snippet.***

### Spark Session

Spark Session can be created using following methods –

1. Jupyter Notebook.
2. Executing Pyspark command with necessary configs (as declared in entrypoint file).

In this integration, the PySpark command is used to initialize a Spark session. Since the Dockerfile specifies Jupyter in the PySpark environment variables, executing the PySpark command automatically starts the Jupyter server and creates the Spark session. This session can then be directly accessed in Jupyter Notebook using the ***spark*** variable.

#### Details of Spark Session using Delta Lake

Snippet of Spark Session Script present in entrypoint.sh file–

A close-up of a computer code

Description automatically generated

In the above script, following configs enable Delta Table capability in Spark.

* **appName("jupyter-pyspark-session")**
  + ***Purpose***
    - Sets the name of the Spark application.
  + **Details**
    - Application is named to help identify it in Spark’s UI and logs.
  + **Why It’s Needed**
    - Useful for tracking and monitoring the application in a Spark cluster.
  + **Key Features It Enables**
    - Helps in debugging and logging.
    - Identifies the application in Spark UI and YARN resource manager.
* **.master("spark://spark-master:7077")**
  + **Purpose:**
    - **Defines the** master node for Spark execution.
  + **Details:**
    - The Spark cluster runs in standalone mode, and the master node is accessible at spark://spark-master:7077.
  + **Why It’s Needed:**
    - Ensures the Spark driver knows where to submit jobs for distributed processing.
  + **Key Features It Enables:**
    - Enables Spark’s distributed computing across worker nodes.
    - Ensures job scheduling and resource allocation.
* **--packages** ***io.delta:delta-spark\_2.12:3.2.0***
  + ***Purpose***:
    - Adds the Delta Lake library to Spark session.
  + ***Details***:
    - io.delta:delta-spark\_2.12: Refers to the Delta Lake connector library compatible with Scala 2.12.
    - 3.2.0: Specifies the version of the Delta Lake library.
  + ***Why It’s Needed***:
    - Spark doesn’t natively support Delta Lake; the library is required to enable Delta-specific features.
    - Provides functionality for reading, writing, and managing Delta tables.
  + **Key Features**:
    - Integration with Spark DataFrames and SQL APIs.
    - Delta operations like MERGE, UPDATE, DELETE, and time travel.
* **--conf "*spark.sql.extensions=io.delta.sql.DeltaSparkSessionExtension*"**
  + ***Purpose***:
    - Activates Delta Lake SQL extensions in the Spark session.
  + ***Details***:
    - DeltaSparkSessionExtension: Adds Delta Lake-specific SQL features to the Spark SQL engine.
  + ***Why It’s Needed***:
    - Spark’s native SQL engine doesn’t support Delta-specific commands or features.
    - This extension enhances SQL to support:
      * Time travel queries: Query data at a specific version or timestamp.
      * ACID transactions for data integrity.
      * Schema enforcement and evolution during write operations.
  + **Key Features**:
    - Delta-specific SQL commands:
      * Example*: SELECT \* FROM delta./path/to/delta-table VERSION AS OF 0.*
      * Example: MERGE INTO, UPDATE, DELETE.
* **--conf "*spark.sql.catalog.spark\_catalog=org.apache.spark.sql.delta.catalog.DeltaCatalog"*** 
  + ***Purpose***:
    - Configures the Spark session to use ***DeltaCatalog*** for table management.
  + ***Details***:
    - ***spark\_catalog***: Spark's default catalog for table metadata management.
    - ***DeltaCatalog***: A specialized catalog that manages Delta tables.
  + ***Why It’s Needed***:
    - Spark’s default catalog doesn’t understand Delta table metadata or advanced features like ACID transactions and versioning.
    - Replacing the default catalog ensures Delta tables are handled correctly.
  + **Key Features**:
    - Seamless creation and management of Delta tables.
    - Metadata management for versioning and schema evolution.

### Config Testing Jupyter Notebook

To test the above spark session and integrations, use jupyter notebook present in below location

**Spark Notebook** - ***/spark-working-folder/spark\_apps/1-delta-lake-config.ipynb***

## Spark, Delta Lake and Minio Integration

### Services:

* Spark Master
* Spark Worker
* Jupyter Lab
* Minio – Minio is the S3 Compatible Object Storage

### Installation and Configuration

* **Installation and Config Folder** ***- /install-and-config/*** ***2-spark-delta-minio-integration***

Docker is used to setup the spark, jupyter and Minio services. These files need to be accessed in the following order for better understanding–

* + *Dockerfile.spark*
  + *Dockerfile.minio*
  + *requirement.txt and spark-config*
  + *entrypoint.sh*
  + *entrypoint-minio.sh*
  + *docker-compose.yml*

***Note****: Spark Dockerfile, entrypoint, requirement and Spark Config files are same/ similar to the details mentioned in “Spark and Delta Lake integration” Section*

Below snippets shows details of some important scripts present in Minio docker and its entrypoint.

1. ***Dockerfile.minio***

A screenshot of a computer program

Description automatically generated

1. ***entrypoint.sh***

In this entrypoint file, based on the parameter following scripts will be executed –

1. Parameter = ***master***

* Start spark master by executing spark master script.
* Start Jupyter lab server. Token for the jupyter lab can be found by executing below script in spark master container.

|  |
| --- |
| ***jupyter server list*** |

1. Parameter = ***worker***

* Start spark worker by executing spark worker script.

1. Parameter = ***history***

* Start spark history server by executing spark history server script.

***Note****: Script to start Spark worker and history server is also controlled using this entrypoint file. Below snippet shows only spark master but file also container spark worker and history server.*

A screen shot of a computer

Description automatically generated

1. ***entrypoint-minio.sh***

A screenshot of a computer

Description automatically generated

1. ***docker-compose.yml***

***Note****: Below snippet shows only minio section from docker-compose file but it also container other services like Spark Master, Worker and History server*

A close-up of a white background

Description automatically generated

In docker container volume section, all the logs generated from entrypoint-minio scripts, data and Spark notebooks are synced between Container and Host Machine.

### Spark Session

Spark Session can be created using following methods –

1. Jupyter Notebook.
2. Executing Pyspark command with necessary configs (as declared in entrypoint file).

In this integration, Spark session is initialized in the jupyter notebook itself. Notebook is present at ***/spark-working-folder/spark\_apps/2-delta-lake-minio-config.ipynb.***

A computer screen shot of a computer code

Description automatically generated

***Note****: Delta configurations are same to the details mentioned in Spark session of “Spark and Delta Lake integration” Section.*

***Configuration for Minio***

* **spark.jars.packages**
  + ***Purpose***:

Adds Hadoop S3 and Delta Lake library (jars) to Spark session.

* + ***Details***:
    - ***org.apache.hadoop:hadoop-aws:3.3.4****:* Adds AWS S3 support via Hadoop's s3a connector.
  + ***Why It’s Needed***:
    - Spark doesn’t natively support **S3 storage**. These dependencies enable that functionality.
    - The hadoop-aws library is necessary for Spark to interact with **MinIO (S3-compatible storage)**.
  + ***Key Features:***
    - Enable storage interaction with S3/MinIO.
* **spark.hadoop.fs.s3a.endpoint="http://minio:9000"** 
  + ***Purpose***:

Specifies the S3-compatible endpoint for object storage (MinIO in this case).

* + ***Details***:
    - The s3a endpoint allows Spark to interact with MinIO, which acts as an S3-compatible storage system.
  + ***Why It’s Needed***:
    - Spark needs an endpoint to connect to MinIO.
  + ***Key Features:***
    - Read/write data from MinIO using ***s3a://*** paths.
    - Enables MinIO as a data lake storage backend for Delta Lake.
* **spark.hadoop.fs.s3a.access.key="root"** 
  + ***Purpose***:

Defines the *access key* for authenticating Spark to MinIO

* + ***Details***:
    - Acts like an AWS *Access Key ID*, but for MinIO.
    - *root* is a user credential for MinIO.
  + ***Why It’s Needed***:
    - MinIO requires authentication to allow read/write access.
  + ***Key Features:***
    - Grant access to MinIO storage from Spark.
* **spark.hadoop.fs.s3a.secret.key="jerinminioserver"** 
  + ***Purpose***:
    - Defines the *secret key* for authenticating Spark to MinIO.
  + ***Details***:
    - Works like an AWS *Secret Access Key* for Minio.
    - *jerinminioserver* is the MinIO user password.
  + ***Why It’s Needed***:
    - MinIO requires authentication to allow read/write access.
  + ***Key Features:***
    - Grant access to MinIO storage from Spark.
* **spark.hadoop.fs.s3a.path.style.access="true"** 
  + ***Purpose***:
    - Enables *path-style access* for MinIO instead of *virtual-hosted access*.
  + ***Details***:
    - *true*: Uses path-style (http://minio:9000/*bucket-name*).
    - *false*: Uses virtual-hosted style (http://*bucket-name*.minio:9000).
  + ***Why It’s Needed***:
    - MinIO uses path-style URLs by default.
    - Some S3-compatible systems (including MinIO) don’t support virtual-hosted access.
  + ***Key Features:***
    - Ensures Spark correctly formats MinIO storage URLs.
* **spark.hadoop.fs.s3a.impl="org.apache.hadoop.fs.s3a.S3AFileSystem"** 
  + ***Purpose***:
    - Specifies the *S3A implementation* for Hadoop’s file system.
  + ***Details***:
    - *org.apache.hadoop.fs.s3a.S3AFileSystem:* The Hadoop class that enables Spark to communicate with S3-like storage.
  + ***Why It’s Needed***:
    - Without this, Spark wouldn’t recognize *s3a://* paths.
  + ***Key Features:***
    - Enables MinIO/S3 storage support in Spark.
* **spark.sql.warehouse.dir="s3a://delta-bucket/delta-lake/data"** 
  + ***Purpose***:
    - Defines the *default location* for Spark-managed tables.
  + ***Details***:
    - Specifies that Spark’s warehouse directory is stored in MinIO (s3a://delta-bucket/delta-lake/data).
  + ***Why It’s Needed***:
    - Ensures that Spark stores all tables in MinIO rather than the local file system.
  + ***Key Features:***
    - Allows Delta Lake tables to be stored in MinIO.
    - Ensures Spark persists table metadata in a central storage.

### Config Testing Jupyter Notebook

To test the above spark session and integrations, use jupyter notebook present in below location

**Spark Notebook** - ***/spark-working-folder/spark\_apps/2-delta-lake-minio-config.ipynb***

## Spark, Delta Lake and Unity Catalog Integration

### Services:

* Spark Master
* Spark Worker
* Jupyter Lab
* Unity Catalog

### Installation and Configuration

* **Installation and Config Folder** ***- /install-and-config/3-spark-delta-unity-integration***

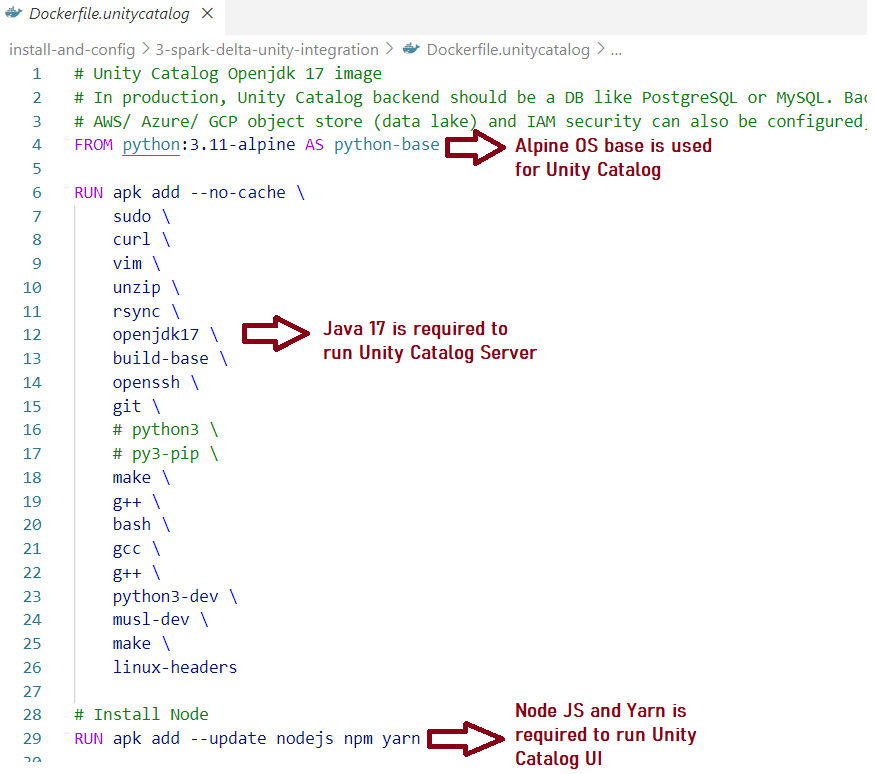
Docker is used to setup the spark, jupyter and Uniity Catalog services. These files need to be accessed in the following order for better understanding–

* + *Dockerfile.spark*
  + *Dockerfile.unitycatalog*
  + *requirement.txt and spark-config*
  + *entrypoint.sh*
  + *entrypoint-unitycatalog.sh*
  + *docker-compose.yml*

***Note****: Spark Dockerfile, entrypoint, requirement and Spark Config files are same/ similar to the details mentioned in “Spark and Delta Lake integration” Section*

Below snippets shows details of some important scripts present in Unity Catalog docker and its entrypoint.

1. ***Dockerfile.unitycatalog***



A screenshot of a computer program

Description automatically generated

A screenshot of a computer program

Description automatically generated

In production, Unity Catalog backend should be a DB like PostgreSQL or MySQL. Backend database can be configured using ***hibernate.properties*** file present at ***/unitycatalog/etc/config*** folder.

AWS/ Azure/ GCP object store (data lake) and IAM security can also be configured, which can be done using ***server.properties*** file present at ***/unitycatalog/etc/config*** folder.

1. ***entrypoint-unitycatalog.sh***

A screenshot of a computer

Description automatically generated

A close-up of a computer code

Description automatically generated

1. ***docker-compose.yml***

***Note****: Below snippet shows only unity catalog section from docker-compose file but it also container other services like Spark Master, Worker and History server which is same as mentioned in “Spark and Delta Integration” section*

A screenshot of a computer

Description automatically generated

### Spark Session

Spark Session can be created using following methods –

1. Jupyter Notebook.
2. Executing Pyspark command with necessary configs (as declared in entrypoint file).

In this integration, the PySpark command is used to initialize a Spark session. Since the Dockerfile specifies Jupyter in the PySpark environment variables, executing the PySpark command automatically starts the Jupyter server and creates the Spark session. This session can then be directly accessed in Jupyter Notebook using the ***spark*** variable.

#### Details of Spark Session using Delta Lake and Unity Catalog

Snippet of Spark Session Script present in entrypoint.sh file –

A close-up of a computer code

Description automatically generated

1. ***-- packages***

* **Purpose**
  + Includes the required libraries for Delta Lake and Unity Catalog.
* **Details**
  + "***io.delta:delta-spark\_2.12:3.2.1***" → Adds Delta Lake support for Spark (Scala 2.12).
  + "***io.unitycatalog:unitycatalog-spark\_2.12:0.2.0***" → Enables Unity Catalog support in Spark.
* **Why It’s Needed**
  + Delta Lake provides ACID transactions and time travel for tables.
  + Unity Catalog helps centrally manage permissions, metadata, and governance.
* **Key Features It Enables**
  + Delta table support (MERGE, DELETE, UPDATE, time travel).
  + Unity Catalog for data governance and access control.
  + Cross-workspace table access in Databricks (if used in Databricks).

1. ***--conf "spark.sql.extensions=io.delta.sql.DeltaSparkSessionExtension"***

* **Purpose**
  + Extends Spark SQL to support Delta Lake commands.
* **Details**
  + Enable Delta-specific SQL functions in Spark.
* **Why It’s Needed**
  + Spark doesn’t natively support Delta tables without this extension.
* **Key Features It Enables**
  + MERGE INTO, DELETE, UPDATE in SQL.
  + Schema enforcement and evolution.
  + Transaction log handling.

1. ***--conf "spark.sql.catalog.spark\_catalog=io.unitycatalog.spark.UCSingleCatalog"***

* **Purpose**
  + Configures Spark's default catalog to use Unity Catalog instead of the default Spark catalog.
* **Details**
  + Uses UCSingleCatalog, which is Unity Catalog’s default catalog implementation.
* **Why It’s Needed**
  + Unity Catalog requires a specific catalog implementation to function.
* **Key Features It Enables**
  + Default catalog becomes Unity Catalog.
  + Allows managing permissions and governance at the catalog level.

1. ***--conf "spark.sql.catalog.unity=io.unitycatalog.spark.UCSingleCatalog"***

* **Purpose**
  + Defines/ Creates a named catalog (unity) that uses Unity Catalog.
* **Details**
  + "unity" is an alias for the Unity Catalog backend.
* **Why It’s Needed**
  + Allows explicit Unity Catalog references (SELECT \* FROM unity.schema.table).
* **Key Features It Enables**
  + Multiple catalog support (e.g., default Spark catalog + Unity Catalog).
  + Enables federated governance across catalogs.

1. ***--conf "spark.sql.catalog.unity.uri=http://unity-catalog:8080"***

* **Purpose**
  + Specifies the Unity Catalog service URL for governance and metadata retrieval.
* **Details**
  + "http://unity-catalog:8080" is the Unity Catalog API endpoint.
* **Why It’s Needed**
  + Unity Catalog runs as a separate service, and Spark needs to communicate with it.
* **Key Features It Enables**
  + Connects Spark to the Unity Catalog governance service.
  + Allows remote metadata management and access control.

1. ***--conf "spark.sql.catalog.unity.token="***

* **Purpose**
  + Defines an authentication token to connect to Unity Catalog.
* **Details**
  + This is empty (""), meaning authentication is either disabled or expected to be passed manually.
* **Why It’s Needed**
  + Unity Catalog requires authentication to enforce permissions.
* **Key Features It Enables**
  + Enables secure authentication for Unity Catalog.
  + Allows fine-grained access control at the table level.

1. ***--conf "spark.sql.defaultCatalog=unity"***

* **Purpose**
  + Sets Unity Catalog as the default catalog for SQL queries.
* **Details**
  + Without this, Spark defaults to spark\_catalog.
* **Why It’s Needed**
  + Ensures all queries use Unity Catalog tables by default.
* **Key Features It Enables**
  + Users don’t need to prefix queries with unity (SELECT \* FROM schema.table instead of SELECT \* FROM unity.schema.table).
  + Makes Unity Catalog the primary catalog for governance.

***Note****: Spark and Delta configurations are same to the details mentioned in Spark session of “Spark and Delta Lake integration” Section.*

### Config Testing Jupyter Notebook

To test the above spark session and integrations, use jupyter notebook present in below location

**Spark Notebook** - ***/spark-working-folder/spark\_apps/3-unity-calatog-config.ipynb***

## Spark, Delta Lake, Unity Catalog and Minio Integration

### Services:

* Spark Master
* Spark Worker
* Jupyter Lab
* Unity Catalog
* Minio – S3 Compatible Object Store

### Installation and Configuration

* **Installation and Config Folder** ***- /install-and-config/*** ***4-issue-spark-delta-unity-minio-integration***

Docker is used to setup the spark, jupyter, Uniity Catalog and Minio services. These files need to be accessed in the following order for better understanding–

* + *Dockerfile.spark*
  + *Dockerfile.unitycatalog*
  + *Dockerfile.minio*
  + *requirement.txt, spark-config and unity-catalog-config*
  + *entrypoint.sh*
  + *entrypoint-unitycatalog.sh*
  + *entrypoint-minio.sh*
  + *docker-compose.yml*

***Note****: Spark, Minio and Unity Catalog’s Dockerfile, entrypoint, requirement and Spark Config files are same/ similar to the details mentioned in “****Spark and Delta Lake integration****”, “****Spark, Delta Lake and Minio integration****” and “****Spark, Delta Lake and Unity Catalog integration****” Section*

Below snippets shows details of some important scripts present in Unity Catalog docker and its entrypoint.

1. ***spark-defaults.conf***

Spark defaults file is updated with Minio details, when spark session is created then it will be used by default.

A screenshot of a computer

Description automatically generated

1. ***Dockerfile.unitycatalog***

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

1. ***entrypoint-unitycatalog.sh***

**Note**: *Unlike in "Spark, Delta Lake, and Unity Catalog integration," where PySpark is used to create a default Spark session and start Jupyter Lab, in this integration, only Jupyter Lab is started, and the Spark session is explicitly created in a Jupyter Notebook*

A screen shot of a computer

Description automatically generated

1. ***Additional Env Variables in Dockerfile.unitycatalog and Dockerfile.spark***

The *AWS\_REGION, AWS\_ACCESS\_KEY\_ID, and AWS\_SECRET\_ACCESS\_KEY* environment variables are utilized by AWS bundle JARs to access MinIO. These variables must be set in the container from which Spark scripts are executed and submitted for processing.

In this setup, Spark Master, Workers, and Jupyter Lab are responsible for executing Spark scripts. The container includes all necessary JARs, including Unity Catalog and AWS S3 SDK JARs.

Since MinIO is AWS S3-compatible and accessed using AWS S3 SDK JARs, the corresponding AWS S3 environment variables are used. Unity Catalog interacts with MinIO through the AWS S3 bundle JAR, which retrieves region and credential information from the environment variables within the Docker container.

A close-up of a person holding an object

Description automatically generated

### Integration Issue

In Unity Catalog to use object store like AWS/ Azure/ GCP, Unity Catalog’s ***server.properties*** file present at ***/unitycatalog/etc/config*** has to be configured.

**Snippet of *server.properties* file**

A screenshot of a computer program

Description automatically generated

Since Minio is AWS S3 compatible storage, so AWS parameters can be used, however Unity Catalog does not provide any parameter for AWS S3 endpoint, it uses default endpoints to communicate with AWS storage therefore Minio endpoint can not be added in the “server.properties” file.

A screenshot of a computer program

Description automatically generated

***Spark Session used to test the integration –***

Following spark session is used to test the integration.

***Spark Notebook Location*** - */spark-working-folder/spark\_apps/4-issue-unity-calatog-and-minio-config.ipynb*



If any table is created with S3 path details then Table gets successfully created in Minio and gets registered in Unity Catalog, however the storage location is incorrectly registered with local file system directory name as prefix in minio S3 location.

A screenshot of a computer code

Description automatically generated

While inserting the data Spark throws “***table not found***” error due to incorrect storage location present in Unity Catalog.

A screenshot of a computer

Description automatically generated

If any table is created without specifying the table location, then Unity Catalog will try to use default location mentioned in the Spark Configuration, and since the location is not accessible by Unity Catalog the create statement will fail with “” error

A screenshot of a computer

Description automatically generated

***Ref of Minio and Unity Catalog Issue:*** <https://community.databricks.com/t5/administration-architecture/setup-unity-catalog-external-location-to-minio/td-p/49895>

A white background with black text

Description automatically generated

### Config Testing Jupyter Notebook

To test the above spark session and integrations, use jupyter notebook present in below location

**Spark Notebook** - ***/spark-working-folder/spark\_apps/4-issue-unity-calatog-and-minio-config.ipynb***

## Spark, Delta Lake, Hive Metastore and Minio Integration

### Services:

* Spark Master
* Spark Worker
* Jupyter Lab
* Hive Metastore
* Postgres – This is used as Hive Metastore backend DB
* Minio

### Installation and Configuration

* **Installation and Config Folder** ***- /install-and-config/*** ***5-spark-delta-hivemetastore-minio-integration***

Docker is used to setup the spark, jupyter, Hive Metastore and Minio services. These files need to be accessed in the following order for better understanding–

* + *Dockerfile.spark*
  + *Dockerfile.hivemetastore*
  + *Dockerfile.minio*
  + *requirement.txt, spark-config and hive-config*
  + *entrypoint.sh*
  + *entrypoint-hivemetastore.sh*
  + *entrypoint-minio.sh*
  + *docker-compose.yml*

***Note****: Spark and Minio Dockerfile, entrypoint, requirement and Spark Config files are same/ similar to the details mentioned in “Spark, Delta Lake and Minio integration” Section*

Below snippets shows details of some important scripts from this integration.

1. ***docker-compose.yml : Postgres Server***

The below screenshot shows the Postgres service used in this integration. Postgres is used as external hive metastore database and is added in hive configuration files.

When the postgres service is created then DB “hivemetastore” will also be created which will be used for storing metadata information.

A screenshot of a computer

Description automatically generated

1. ***hive-config folder***
   * ***hive-env.sh:*** Thisfile is used to set Hive specific Environment Variables.

A red arrow pointing to the right

Description automatically generated

* + ***hive-site.xml:*** This file configures Hive to connect with MinIO for storage and PostgreSQL as an external metastore. It defines various runtime settings, including metadata storage, execution engine, authentication, and optimization parameters.

Hive provides the template to define all the required parameters, this can be found at hive installation directory ***“/conf/hive-default.xml”***

A screenshot of a computer screen

Description automatically generated

A screenshot of a computer

Description automatically generated

1. ***spark-config folder***
   * ***hive-site.xml:*** This configuration is same as mentioned in above step, this will be used by Spark to connect to Hive Metastore. There’s only 1 change in property “***hive.metastore.uris***”. Since Spark needs to connect to hive metastore therefore in this property instead of “***thrift://0.0.0.0:9083***”, docker hive container service name should be passed i.e. “***thrift://hive-metastore:9083***”

A computer code with text and symbols

Description automatically generated with medium confidence

* + ***spark-defaults.conf***

The spark-defaults.conf file is used to define default configurations for Spark applications. It sets various runtime properties, including memory allocation, execution modes, and metastore connectivity.

A screenshot of a computer

Description automatically generated

1. ***Dockerfile.hivemetastore***

This is used to setup hive docker image. Below snippets show important details of Hive Installation and Configuration.

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A group of text boxes

Description automatically generated with medium confidence

A screenshot of a computer program

Description automatically generated

***Note:***

* + ***While installing Hadoop***: Spark 3.5.4 supports Hadoop 3.3.4 by default and the latest version of Hadoop gives error (Number conversion error) during connection with hive external metastore.
  + ***While installing Hive***: Spark 3.5.4 supports Hive 2.3.9 by default and the latest version of Hive gives error during connection with hive external metastore.

1. ***entrypoint-hivemetastore.sh***

Whenever the hive container starts, the scripts in entrypoint file will be executed which will be used to cerate metadata management tables in Postgres DB “*hivemetastore*” and start hive metastore service.

A screenshot of a computer

Description automatically generated

1. ***Dockerfile.spark***

***Note****: Below snippet shows only additional spark configurations used in this Dockefile, other details are same as mentioned in “Spark, Delta Lake and Minio Integration” section.*

A close-up of a computer screen

Description automatically generated

1. ***docker-compose.yml***

***Note****: Below snippet shows only Hive Metastore section from docker-compose file but it also contains other services like Spark Master, Worker, History server and Minio which is same as mentioned in “Spark, Delta Lake and Minio Integration” section.*

A close-up of a computer screen

Description automatically generated

1. ***Additional Env Variables in Dockerfile.spark and Dockerfile.hivemetastore (Optional: Not required in this integration)***

The *AWS\_REGION, AWS\_ACCESS\_KEY\_ID, and AWS\_SECRET\_ACCESS\_KEY* environment variables are utilized by AWS bundle JARs to access MinIO. These variables must be set in the container from which Spark scripts are executed and submitted for processing.

In this setup, Spark Master, Workers, and Jupyter Lab are responsible for executing Spark scripts. The container includes all necessary JARs, including Unity Catalog and AWS S3 SDK JARs.

Since MinIO is AWS S3-compatible and accessed using AWS S3 SDK JARs, the corresponding AWS S3 environment variables are used. Unity Catalog interacts with MinIO through the AWS S3 bundle JAR, which retrieves region and credential information from the environment variables within the Docker container.

A close-up of a person holding an object

Description automatically generated

### Spark Session

Spark Session can be created using following methods –

1. Jupyter Notebook.
2. Executing Pyspark command with necessary configs (as declared in entrypoint file).

In this integration, Spark session is initialized in the jupyter notebook itself. Notebook is present at ***/spark-working-folder/spark\_apps/5-delta-hive-metastore-and-minio-config.ipynb***

A computer code with many small colored text

Description automatically generated with medium confidence

***Note****: Spark. Minio and Delta configurations are same to the details mentioned in Spark session of “Spark Delta Lake and Minio integration” Section.*

1. **.appName("add-delta-lake\_1")**

* **Purpose:**
  + Defines the name of the Spark application for tracking in the Spark UI.
* **Details:**
  + This helps identify the specific job in Spark's monitoring tools.
* **Why It’s Needed:**
  + Allows easy tracking and debugging of jobs running in a Spark cluster.
* **Key Features It Enables:**
  + Visibility in Spark UI.
  + Easier job management.

1. **master("spark://spark-master:7077")**

* **Purpose:**
  + Specifies the Spark master URL for running in a standalone cluster.
* **Details:**
  + Defines the cluster mode by specifying the master node's address.
* **Why It’s Needed:**
  + Essential for deploying Spark applications in a distributed environment.
* **Key Features It Enables:**
  + Distributed computing.
  + Efficient resource allocation.

1. **config("spark.jars.packages", "io.delta:delta-spark\_2.12:3.3.0, org.apache.hadoop:hadoop-aws:3.3.4, org.postgresql:postgresql:42.6.0")**

* **Purpose:**
  + **Automatically downloads and loads required dependencies.**
* **Details:**
  + ***io.delta:delta-spark\_2.12:3.3.0***: Enables Delta Lake support.
  + ***org.apache.hadoop:hadoop-aws:3.3.4***: Adds AWS S3 connectivity.
  + ***org.postgresql:postgresql:42.6.0***: Required for PostgreSQL connectivity.
* **Why It’s Needed:**
  + Ensures the necessary libraries are available for Delta Lake, MinIO (S3), and Hive Metastore.
* **Key Features It Enables:**
  + Delta Lake support.
  + MinIO (S3) integration.
  + PostgreSQL-based Hive Metastore connectivity.

1. **config("spark.sql.extensions", "io.delta.sql.DeltaSparkSessionExtension")**

* **Purpose:**
  + Allows Spark to recognize and work with Delta Lake.
* **Details:**
  + Extends Spark’s SQL functionality to support Delta table operations.
* **Why It’s Needed:**
  + Essential for enabling ACID transactions, schema enforcement, and time travel in Delta Lake.
* **Key Features It Enables:**
  + Delta Lake support in SQL queries.
  + Schema evolution and enforcement.
  + Time travel capabilities.

1. **.config("spark.sql.catalog.spark\_catalog","org.apache.spark.sql.delta.catalog.DeltaCatalog")**

* **Purpose:**
  + Configures Spark's default catalog to use Delta Lake Catalog.
* **Details:**
  + Replaces the default catalog (spark\_catalog) with DeltaCatalog to manage Delta tables.
* **Why It’s Needed:**
  + Ensures Spark interacts with Delta tables via Delta Catalog.
* **Key Features It Enables:**
  + Centralized Delta table management.
  + ACID-compliant transactions.

1. **Configuring MinIO (S3-Compatible Storage)**

* **Commands:**
  + *.config("spark.hadoop.fs.s3a.endpoint", "http://minio:9000")*
  + *.config("spark.hadoop.fs.s3a.access.key", "root")*
  + *.config("spark.hadoop.fs.s3a.secret.key", "jerinminioserver")*
  + *.config("spark.hadoop.fs.s3a.path.style.access", "true")*
  + *.config("spark.hadoop.fs.s3a.impl", "org.apache.hadoop.fs.s3a.S3AFileSystem")*
* **Purpose:**
  + Enables Spark to read/write data from MinIO, an S3-compatible object storage.
* **Details:**
  + Sets MinIO endpoint (http://minio:9000).
  + Provides authentication credentials.
  + Enables path-style access (true).
  + Uses Hadoop S3A implementation.
* **Why It’s Needed:**
  + Necessary for integrating MinIO as a storage backend for Spark and Delta Lake.
* **Key Features It Enables:**
  + Reading/writing Delta tables to MinIO.
  + S3-compatible object storage.
  + Secure access using credentials.

1. **config("spark.sql.warehouse.dir", "s3a://hivebucket/delta-lake/data")**

* **Purpose:**
  + Defines the default location for managed Delta tables.
* **Details:**
  + Uses MinIO (S3A) to store Delta tables.
* **Why It’s Needed:**
  + Ensures all Delta tables are stored in a structured location for easy retrieval.
* **Key Features It Enables:**
  + Centralized data storage.
  + Efficient data retrieval.

1. **Enabling Hive Metastore**

* **Commands:**
  + *.config("spark.sql.catalogImplementation", "hive")*
  + *.config("spark.hadoop.hive.metastore.uris", "thrift://hive-metastore:9083")*
* **Purpose:**
  + Enables Hive Metastore for managing Delta table metadata.
* **Details:**
  + Uses Hive Metastore (***thrift://hive-metastore:9083***).
  + Allows external table metadata management via Hive.
* **Why It’s Needed:**
  + Enables structured metadata management, making it easier to access tables from different systems.
* **Key Features It Enables:**
  + Centralized catalog for Delta tables.
  + HiveQL-based queries on Delta tables.

1. **Connecting Hive External Metastore to PostgreSQL**

* **Commands:**
  + *.config("spark.hadoop.datanucleus.connection.driver", "org.postgresql.Driver")*
  + *.config("spark.hadoop.datanucleus.ConnectionURL", "jdbc:postgresql://postgresdb:5432/hivemetastore")*
  + *.config("spark.hadoop.datanucleus.ConnectionUserName", "jerin")*
  + *.config("spark.hadoop.datanucleus.ConnectionPassword", "jerinpostgresql")*
* **Purpose:**
  + Configures Hive Metastore to use PostgreSQL as the backend database.
* **Details:**
  + Uses PostgreSQL JDBC driver.
  + Sets connection URL, username, and password.
* **Why It’s Needed:**
  + Ensures that table metadata is stored persistently and can be queried efficiently.
* **Key Features It Enables:**
  + Persistent Hive Metastore.
  + Secure database-backed metadata storage.

1. **Enabling Hive Support and Creating Spark Session**

* **Commands:**
  + *.enableHiveSupport()*
  + *.getOrCreate()*
* **Purpose:**
  + Enables full Hive support within Spark and initializes the session.
* **Details:**
  + ***.enableHiveSupport()* 🡪** allows interaction with Hive tables.
  + ***.getOrCreate()* 🡪** ensures the session is created or retrieved.
* **Why It’s Needed:**
  + Allows Hive-compatible queries and seamless integration of Spark with external catalogs.
* **Key Features It Enables:**
  + Querying Hive tables from Spark.
  + Reusing the same session across multiple applications.

#### spark.sql.catalog.spark\_catalog=org.apache.spark.sql.delta.catalog.DeltaCatalog and spark.sql.catalogImplementation=hive

##### Configuration Behaviour:

1. **spark.sql.catalog.spark\_catalog=org.apache.spark.sql.delta.catalog.DeltaCatalog**
   * This setting configures **spark\_catalog** to use **Delta Lake's catalog** instead of the default Hive catalog.
   * When you create or query tables using spark\_catalog, Spark will assume they are **Delta tables** by default.
   * Commands like CREATE TABLE and DESCRIBE TABLE will operate assuming **Delta Lake tables**.
2. **spark.sql.catalogImplementation=hive**
   * This setting tells Spark to use **Hive Metastore** as the metadata store.
   * Even though ***spark\_catalog is configured for Delta***, the table metadata (location, schema, etc.) will still be managed by **Hive Metastore**.
   * If a table is non-Delta (e.g., a standard Hive table), it can still be queried through Hive Metastore.

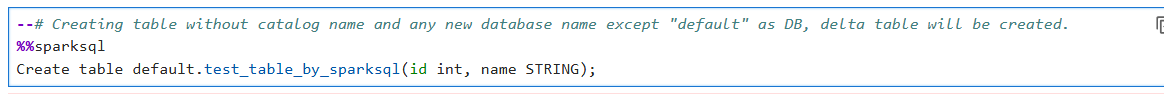
##### Outcome of This Configuration:

* + Delta tables registered under *spark\_catalog* will use Hive Metastore for metadata storage.
    - Delta will store its **transaction logs** under the table's data location.
    - Metadata such as schema, partitions, and table definitions will be managed in **Hive Metastore**.
  + **Hive-managed tables can still be used separately** (but they must be accessed explicitly if needed).

##### Table creation behaviour:

* A table will ***be created as a Delta table*** if any of the following conditions are met:
  + The table is created ***without*** specifying the catalog name or database name.
  + The table is created ***without*** specifying a new database name (i.e., using the default database).
  + The table is created ***explicitly*** inside the default database (as shown in below snippet).

***Example Code:***

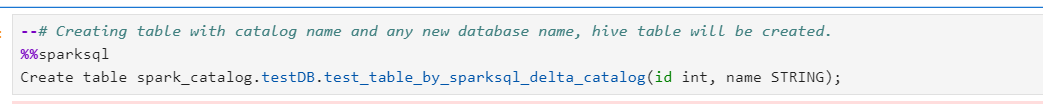


***Code Explanation:***

Since ***default*** db is part of ***spark\_catalog***, and in Spark Config, ***spark\_catalog*** is configured to use ***DeltaCatalog***, Spark automatically creates the table as a ***Delta table***.

* A table will ***be created as a Hive table*** if any of the following conditions are met:
  + The table is created ***with*** specifying the catalog name or new database name.

***Example Code:***



***Code Explanation:***

Explicitly mentioned ***spark\_catalog.testDB*** while creating the table and since Spark doesn’t ***automatically assume*** ***Delta*** for explicitly mentioned catalogs therefore ***Hive*** table will be created.

* + The table is created ***with*** specifying the catalog name or even “default” database name.

***Example Code:***A close-up of a computer screen

Description automatically generated

***Code Explanation:***

Explicitly mentioned ***spark\_catalog.default*** while creating the table and since Spark doesn’t ***automatically assume*** ***Delta*** for explicitly mentioned catalogs therefore ***Hive*** table will be created ***even though the default db is registered as Delta***.

* + - Spark assumes ***spark\_catalog is Delta***, but only when it is used ***implicitly***.
    - Explicitly mentioning ***spark\_catalog.default bypasses this behaviour, and Spark treats it as a Hive table.***

### Config Testing Jupyter Notebook

To test the above spark session and integrations, use jupyter notebook present in below location

**Spark Notebook** - ***/spark-working-folder/spark\_apps/5-delta-hive-metastore-and-minio-config.ipynb***

### UI for Hive Metastore

*Apache Hue can be used as Open Source Hive Metastore UI, which can also be integrated with Jupyter as well.*

## Spark, Iceberg and Minio Integration

### Services:

* Spark Master
* Spark Worker
* Jupyter Lab
* Minio – Minio is the S3 Compatible Object Storage

### Installation and Configuration

* **Installation and Config Folder** ***- /install-and-config/6-spark-iceberg-minio-integration***

Docker is used to setup the spark, jupyter and Minio services. These files need to be accessed in the following order for better understanding–

* + *Dockerfile.spark*
  + *Dockerfile.minio*
  + *requirement.txt and spark-config*
  + *entrypoint.sh*
  + *entrypoint-minio.sh*
  + *docker-compose.yml*

***Note****: Spark and Minio’s Dockerfile, entrypoint, requirement and Spark Config files are same/ similar to the details mentioned in “Spark, Delta Lake and Minio integration” Section*

### Spark Session

Spark Session can be created using following methods –

1. Jupyter Notebook.
2. Executing Pyspark command with necessary configs (as declared in entrypoint file).

In this integration, Spark session is initialized in the jupyter notebook itself. Notebook is present at ***/spark-working-folder/spark\_apps/6-iceberg-minio-config.ipynb***

A close-up of a computer screen

Description automatically generated

***Note****: Minio configurations are same to the details mentioned in Spark session of “Spark, Delta Lake and Minio integration” Section.*

***Iceberg Configuration in Spark Session –***

1. ***spark.jars.packages***

* **Purpose**
  + Loads the necessary JAR dependencies for Iceberg, Hive, and AWS S3 (MinIO).
* **Details**
  + ***"org.apache.iceberg:iceberg-spark-runtime-3.5\_2.12:1.7.1"*** 🡪 Iceberg support for Spark 3.5 (Scala 2.12).
  + ***"org.apache.hive:hive-exec:3.1.3"*** 🡪 Hive execution engine for metadata management.
  + ***"org.apache.hadoop:hadoop-aws:3.3.4"*** 🡪 Hadoop AWS connector for S3/MinIO access.
* **Why It’s Needed**
  + Iceberg requires a runtime library to function in Spark.
  + Hive execution is needed for metadata and schema management.
  + Hadoop AWS library allows Spark to interact with S3-compatible storage.
* **Key Features It Enables**
  + Iceberg table format support in Spark.
  + Hive Metastore integration for schema evolution.
  + S3/MinIO connectivity for storing Iceberg data.

1. ***spark.sql.extensions="org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions"***

* **Purpose**
  + Enables Iceberg-specific SQL commands in Spark.
* **Details**
  + Extends Spark with Iceberg SQL capabilities.
* **Why It’s Needed**
  + Spark does not natively support Iceberg table commands like MERGE, REPLACE, and ROLLBACK.
* **Key Features It Enables**
  + MERGE INTO, COPY ON WRITE, and TIME TRAVEL.
  + Iceberg-specific optimizations like hidden partitioning.

1. ***spark.sql.catalog.spark\_catalog="org.apache.iceberg.spark.SparkSessionCatalog”***

* **Purpose**
  + Defines Iceberg as the default catalog in Spark.
* **Details**
  + SparkSessionCatalog makes Iceberg the primary catalog for table operations.
* **Why It’s Needed**
  + Without this, Spark will use its default catalog, which does not support Iceberg.
* **Key Features It Enables**
  + Allows Spark to directly interact with Iceberg tables.
  + Enables querying Iceberg tables without specifying a custom catalog.

1. ***spark.sql.catalog.local = "org.apache.iceberg.spark.SparkCatalog"***

* **Purpose**
  + Defines an alternative catalog (local) that also uses Iceberg.
* **Details**
  + local is an additional catalog that can store tables in Hadoop/S3.
* **Why It’s Needed**
  + Allows users to store tables in different locations (e.g., MinIO instead of Hive Metastore).
* **Key Features It Enables**
  + Supports multiple catalogs (e.g., spark\_catalog and local).
  + Enables custom storage options for different datasets.

1. ***spark.sql.catalog.spark\_catalog.type = "hive"***

* **Purpose**
  + Configures the default Iceberg catalog to use Hive Metastore.
* **Details**
  + Uses Hive Metastore to store table metadata.
* **Why It’s Needed**
  + Hive Metastore allows centralized metadata management for Iceberg tables.
* **Key Features It Enables**
  + Iceberg metadata persistence.
  + Enables cross-session table access.

1. ***spark.sql.catalog.local.type = "hadoop"***

* **Purpose**
  + Configures the local catalog to store metadata directly in a Hadoop-compatible storage (MinIO/S3).
* **Details**
  + Instead of using Hive Metastore, this catalog stores metadata in S3.
* **Why It’s Needed**
  + Useful for setups without Hive Metastore.
* **Key Features It Enables**
  + Iceberg metadata storage in S3 (MinIO).
  + Simplifies deployment by removing the need for Hive Metastore.

1. ***MinIO (S3) Configurations***

These settings allow Spark to read and write data to MinIO (an S3-compatible object store).

* **Purpose**
  + Configures Spark to interact with MinIO via Hadoop AWS connectors.
* **Details**
  + ***spark.hadoop.fs.s3a.endpoint = "http://minio:9000"*** 🡪 MinIO service URL.
  + ***spark.hadoop.fs.s3a.access.key = "root"*** 🡪 MinIO access key.
  + ***spark.hadoop.fs.s3a.secret.key = "jerinminioserver"*** 🡪MinIO secret key.
  + ***spark.hadoop.fs.s3a.path.style.access = "true"*** 🡪 Enables path-style URLs.
  + ***spark.hadoop.fs.s3a.impl = "org.apache.hadoop.fs.s3a.S3AFileSystem"*** 🡪 Uses Hadoop’s S3A connector.
* **Why It’s Needed**
  + Spark needs authentication and endpoint details to interact with MinIO/S3.
* **Key Features It Enables**
  + S3-compatible storage in Spark.

1. ***spark.sql.catalog.local.warehouse = "s3a://iceberg-bucket/delta-lake/data"***

* **Purpose**
  + Defines the Iceberg table storage location in MinIO.
* **Details**
  + ***s3a://iceberg-bucket/delta-lake/data*** 🡪 All tables under local catalog are stored here.
* **Why It’s Needed**
  + Specifies where Iceberg tables physically store their data.
* **Key Features It Enables**
  + Iceberg table data is stored in MinIO instead of HDFS.
  + Supports scalable cloud storage.

1. ***spark.sql.defaultCatalog = "local"***

* **Purpose**
  + Sets the default catalog to local instead of spark\_catalog.
* **Details**
  + All queries use local catalog by default.
* **Why It’s Needed**
  + Ensures Iceberg tables are created in the MinIO-backed catalog.
* **Key Features It Enables**
  + Users don’t need to specify local. prefix in SQL queries.

1. ***Hive Metastore Configurations***

These settings disable schema verification in Hive Metastore.

* **Purpose**
  + Avoids issues when using Hive Metastore with Iceberg.
* **Details**
  + ***spark.hadoop.datanucleus.autoCreateSchema = "true"*** 🡪 Automatically creates missing schemas.
  + ***spark.hadoop.datanucleus.fixedDatastore = "false"*** 🡪 Allows dynamic schema changes.
  + ***hive.metastore.schema.verification = "false"*** 🡪 Disables strict schema checking.
  + ***hive.metastore.schema.verification.record.version = "false"*** 🡪 Ignores version mismatches.
* **Why It’s Needed**
  + Iceberg and Hive Metastore handle schema differently, so these settings prevent conflicts.
* **Key Features It Enables**
  + Allows Iceberg tables to work without strict schema enforcement.
  + Prevents Hive Metastore errors related to schema changes.

### Config Testing Jupyter Notebook

To test the above spark session and integrations, use jupyter notebook present in below location

**Spark Notebook** - ***/spark-working-folder/spark\_apps/6-iceberg-minio-config.ipynb***

## Spark, Iceberg, Nessie and Minio Integration

### Services:

* Spark Master
* Spark Worker
* Jupyter Lab
* Nessie
* Minio

### Installation and Configuration

* **Installation and Config Folder** ***- /install-and-config/7-spark-iceberg-nessie-minio-integration***

Docker is used to setup the spark, jupyter and Nessie services. These files need to be accessed in the following order for better understanding–

* + *Dockerfile.spark*
  + *Dockerfile.minio*
  + *Dockerfile.pysparkjupyter*
  + *requirement.txt and spark-config*
  + *entrypoint.sh*
  + *entrypoint-pysparkjupyter.sh*
  + *entrypoint-minio.sh*
  + *docker-compose.yml* ***– Nessie is directly created in Docker Compose file***

***Note****: Spark, Minio Dockerfile, entrypoint, requirement and Spark Config files are same/ similar to the details mentioned in* *“Spark and Delta Lake integration” , “Spark, Delta Lake and Minio integration” Section*

Below snippets shows details of some important scripts present in Nessie and Pyspark Jupyter docker and its entrypoint files.

1. ***Dockerfile.pysparkjupyter***

This is similar to Spark Dockerfile, it is created to have a separate docker container to run jupyter lab server.

Details for Spark Docker container can be found in “Spark and Delta Lake integration”.

A screenshot of a computer program

Description automatically generated

1. ***entrypoint-pysparkjupyter.sh***

In this entrypoint file, script to start jupyter lab server exists.

A close-up of a computer screen

Description automatically generated

1. ***docker-compose.yml***

***Note****: Below snippet shows only Nessie section from docker-compose file but it also container other services like Spark Master, Worker, History server, Minio and jupyter which is same as mentioned in “Spark and Delta Integration” , “Spark, Delta and Minio Integration” sections*

A screenshot of a computer

Description automatically generated

***Note****:*

* 1. *Starting Nessie using Dockerfile and entrypoint file was causing issue to start the Nessie server, therefore its used directly in the docker-compose file.*
  2. *It is important to add the network details in all the services (mention in above snippet) present in the docker compose file so that all server are connect through same network.*

1. ***Additional Env Variables in Dockerfile.pysparkjupyter and Dockerfile.spark***

The *AWS\_REGION, AWS\_ACCESS\_KEY\_ID, and AWS\_SECRET\_ACCESS\_KEY* environment variables are utilized by AWS bundle JARs to access MinIO. These variables must be set in the container from which Spark scripts are executed and submitted for processing.

In this setup, Spark Master, Workers, and Jupyter Lab are responsible for executing Spark scripts. The container includes all necessary JARs, including Unity Catalog and AWS S3 SDK JARs.

Since MinIO is AWS S3-compatible and accessed using AWS S3 SDK JARs, the corresponding AWS S3 environment variables are used. Unity Catalog interacts with MinIO through the AWS S3 bundle JAR, which retrieves region and credential information from the environment variables within the Docker container.

A close-up of a person holding an object

Description automatically generated

### Spark Session

Spark Session can be created using following methods –

1. Jupyter Notebook.
2. Executing Pyspark command with necessary configs (as declared in entrypoint file).

In this integration, Spark session is initialized in the jupyter notebook itself. Notebook is present at ***/spark-working-folder/spark\_apps/7-nessie-minio-config.ipynb***

A screenshot of a computer

Description automatically generated

A screenshot of a computer code

Description automatically generated

1. ***setAppName('nessie-app')***

* **Purpose**
  + Sets the name of the Spark application.
* **Details**
  + Application is named to help identify it in Spark’s UI and logs.
* **Why It’s Needed**
  + Useful for tracking and monitoring the application in a Spark cluster.
* **Key Features It Enables**
  + Helps in debugging and logging.
  + Identifies the application in Spark UI and YARN resource manager.

1. ***setMaster("spark://spark-master:7077")***

* **Purpose:**
  + Defines the master node for Spark execution.
* **Details:**
  + The Spark cluster runs in standalone mode, and the master node is accessible at spark://spark-master:7077.
* **Why It’s Needed:**
  + Ensures the Spark driver knows where to submit jobs for distributed processing.
* **Key Features It Enables:**
  + Enables Spark’s distributed computing across worker nodes.
  + Ensures job scheduling and resource allocation.

1. ***set('spark.jars.packages', 'org.postgresql:postgresql:42.7.3,***

***org.apache.iceberg:iceberg-spark-runtime-3.5\_2.12:1.5.0,***

***org.projectnessie.nessie-integrations:nessie-spark-extensions-3.5\_2.12:0.77.1,***

***software.amazon.awssdk:bundle:2.20.147,***

***software.amazon.awssdk:url-connection-client:2.20.147')***

* **Purpose:**
  + Specifies external JAR dependencies for Spark.
* **Details:**
  + Includes required JARs for Iceberg, Nessie, AWS SDK (for MinIO), and PostgreSQL.
* **Why It’s Needed:**
  + ***org.postgresql:postgresql:42.7.3***🡪PostgreSQL JAR: Required if Nessie uses PostgreSQL as a backend.
  + ***org.apache.iceberg:iceberg-spark-runtime-3.5\_2.12:1.5.0***🡪Iceberg JAR: Provides Spark runtime support for Iceberg tables.
  + ***org.projectnessie.nessie-integrations:nessie-spark-extensions-3.5\_2.12:0.77.1***🡪Nessie JAR: Enables Nessie as a catalog implementation for Iceberg.
  + ***software.amazon.awssdk:bundle:2.20.147* & *software.amazon.awssdk:url-connection-client:2.20.147***🡪AWS SDK JAR: Required for MinIO storage operations.
* **Key Features It Enables:**
  + Enables Iceberg table operations in Spark.
  + Allows Spark to interact with Nessie as a catalog.
  + Supports MinIO as an S3-compatible storage backend.

1. **set('spark.sql.extensions', 'org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions,org.projectnessie.spark.extensions.NessieSparkSessionExtensions')**

* **Purpose:**
  + Enables Spark extensions for Iceberg and Nessie.
* **Details:**
  + ***org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions***: Enables Iceberg table operations in Spark.
  + ***projectnessie.spark.extensions.NessieSparkSessionExtensions***:Adds support for Nessie transactions and version control.
* **Why It’s Needed:**
  + These extensions allow Spark SQL to support Iceberg table management and Nessie as a version-controlled catalog.
* **Key Features It Enables:**
  + Iceberg table versioning, schema evolution, and partitioning.
  + Nessie’s time-travel capabilities for Iceberg tables.
  + Support for multi-branch transactions in Spark.

1. **set('spark.sql.catalog.nessie', 'org.apache.iceberg.spark.SparkCatalog')**

* **Purpose:**
  + Configures/ creates ***nessie*** as a catalog in Spark.
* **Details:**
  + Defines the nessie catalog using the SparkCatalog class from Iceberg.
* **Why It’s Needed:**
  + Allows Spark to interact with Nessie as a catalog for Iceberg tables.
* **Key Features It Enables:**
  + Spark can read/write Iceberg tables managed by Nessie.
  + Enables transaction control and table versioning using Nessie.

1. **set('spark.sql.catalog.nessie.uri', v\_nessieCatalogUri)**

* **Purpose:**
  + Defines the Nessie server’s API endpoint.
* **Details:**
  + Uses ***http://nessie:19120/api/v1*** as the default URI for connecting to Nessie.
* **Why It’s Needed:**
  + Ensures Spark can communicate with the Nessie catalog via its REST API.
* **Key Features It Enables:**
  + Allows Spark to perform catalog operations (e.g., creating branches, commits).
  + Enables versioned table management in Nessie.

1. **set('spark.sql.catalog.nessie.ref', 'main')**

* **Purpose:**
  + Sets the default branch/reference in Nessie.
* **Details:**
  + Points to the "main" branch of the Nessie catalog.
* **Why It’s Needed:**
  + Ensures Spark operations use the correct branch for version control.
* **Key Features It Enables:**
  + Supports branching and merging in Nessie for Iceberg tables.
  + Enables time-travel queries using branch-based references.

1. **set('spark.sql.catalog.nessie.authentication.type', 'NONE')**

* **Purpose:**
  + Disables authentication for Nessie.
* **Details:**
  + NONE means no authentication mechanism is used.
* **Why It’s Needed:**
  + Useful in local or development environments without authentication.
* **Key Features It Enables:**
  + Simplifies Nessie setup for testing purposes.
  + Can be modified later for secure authentication.

1. **set('spark.sql.catalog.nessie.catalog-impl', 'org.apache.iceberg.nessie.NessieCatalog')**

* **Purpose:**
  + Sets the catalog implementation class for the Nessie catalog.
* **Details:**
  + Uses ***org.apache.iceberg.nessie.NessieCatalog*** to manage Iceberg tables within Nessie.
* **Why It’s Needed:**
  + Ensures that Iceberg tables are registered in the Nessie catalog.
* **Key Features It Enables:**
  + Iceberg table lifecycle management (creation, updates, deletion) within Nessie.

1. **set('spark.sql.catalog.nessie.s3.endpoint', v\_minioStorageUri)**

* **Purpose:**
  + Defines the MinIO endpoint for object storage.
* **Details:**
  + Uses ***http://172.19.0.2:9000*** as the MinIO service address.
* **Why It’s Needed:**
  + Allows Spark to interact with MinIO as an S3-compatible storage backend.
* **Key Features It Enables:**
  + Enables reading and writing Iceberg table data to MinIO.
  + Provides cloud-like object storage for Iceberg tables.

1. **set('spark.sql.catalog.nessie.warehouse', v\_minioWarehouse)**

* **Purpose:**
  + Specifies the Iceberg warehouse location in MinIO.
* **Details:**
  + Uses s3://nessiebucket/ as the root warehouse directory.
* **Why It’s Needed:**
  + Ensures Iceberg tables are stored in MinIO under the specified path.
* **Key Features It Enables:**
  + Centralized storage location for all Iceberg-managed tables.
  + Supports partitioning and schema evolution in Iceberg.

1. **set('spark.sql.catalog.nessie.io-impl', 'org.apache.iceberg.aws.s3.S3FileIO')**

* **Purpose:**
  + Sets the file I/O implementation for Iceberg.
* **Details:**
  + Uses S3FileIO, which enables Iceberg to read/write data to MinIO using S3 APIs.
* **Why It’s Needed:**
  + Ensures Spark can properly store Iceberg table data in MinIO.
* **Key Features It Enables:**
  + Efficient data storage and retrieval for Iceberg tables.
  + Compatibility with S3-based object storage.

**\*\* Difference between property “*catalog-impl*” and “*io-impl*”:**

|  |  |  |
| --- | --- | --- |
|  | **Spark Config Property** | |
| **Feature** | **catalog-impl (Metadata)** | **io-impl (File Storage)** |
| **Purpose** | Defines how metadata (schemas, tables, versions) is stored and managed. | Defines how table data (files) is stored and accessed. |
| **Manages** | Manages the metadata catalog like Tables, schemas, versions, commits | Manages file I/O operations like location of Table data files (Parquet, ORC, Avro) |
| **Stored In** | Nessie, Hive, AWS Glue | MinIO (S3), HDFS, Local disk |
| **Main Purpose** | Track schema changes, branching, and table versions | Efficiently read/write table data |

**\*\* Important points to remember:**

1. *Nessie to interact with Minio, use IP address with port like* [*http://172.19.0.2:9000*](http://172.19.0.2:9000) *instead of* [*http://minio:9000*](http://minio:9000) *, using docker service name minio will throw error “****connection closed quietly****”. To get the minio docker container IP address use docker command “****docker inspect <container-id or name>****”*
2. *Nessie uses jars “software.amazon.awssdk:bundle” and “software.amazon.awssdk:url-connection-client” to interact with Minio therefore “AWS\_REGION”, “AWS\_ACCESS\_KEY\_ID” and “AWS\_SECRET\_ACCESS\_KEY” Environment Variables need to be added in the Spark and Jupyter containers. These amazon jars always need these Environment Variable to authenticate with Minio.*

### Config Testing Jupyter Notebook

To test the above spark session and integrations, use jupyter notebook present in below location

**Spark Notebook** - ***/spark-working-folder/spark\_apps/7-nessie-minio-config.ipynb***

# Appendix

## Enable SparkSQL Magic Command in Jupyter Notebook

To enable SparkSQL magic command in Jupyter Notebook, install python package ***“sparksql-magic”.***

This package is already added in ***requirement.txt*** file present in all the integration’s folder. This file will be used by *Dockerfile* to install the python packages.

A screenshot of a computer

Description automatically generated

To enable the magic command in the notebook, execute python script ***“%load\_ext sparksql\_magic***” in the notebook after the installation of ***“sparksql-magic”*** python package.

A white background with black text

Description automatically generated

Once the Spark SQL magic command is enabled and to execute Spark SQL command as a normal SQL command without using ***spark.sql()*** function, add ***%%sparksql*** before the query,

A white rectangular object with text

Description automatically generated with medium confidence

## Unity Catalog

* Website: <https://www.unitycatalog.io/>
* Github: <https://github.com/unitycatalog/unitycatalog>
* Documentation: <https://docs.unitycatalog.io/>

## Delta Lake Documentation

* Website: <https://delta.io/>
* Documentation: <https://docs.delta.io/latest/delta-spark.html>
* Spark and Delta Documentation: <https://docs.delta.io/latest/quick-start.html#set-up-apache-spark-with-delta-lake>

## Nessie Documentation

* Website: <https://projectnessie.org/>
* Configuration: <https://projectnessie.org/nessie-0-102-4/configuration/>
* Integration Blog: <https://dev.to/alexmercedcoder/hands-on-with-apache-iceberg-on-your-laptop-deep-dive-with-apache-spark-nessie-minio-dremio-polars-and-seaborn-2hgk>

## Minio Documentation

* Website: <https://min.io/>
* Documentation: <https://min.io/docs/minio/container/index.html>
* Spark and Minio Integration: <https://medium.com/@abdullahdurrani/working-with-minio-and-spark-8b4729daec6e>

## Docker Commands

Following are some important commands, to learn more about docker commands refer GitHub repo: ***docker-learn-hands-on***

|  |
| --- |
| *# Display help information for the `docker ps` command.*  *docker ps --help*  *# List all currently running Docker containers.*  *docker ps*  *# Start all services defined in the `docker-compose.yml` file in detached mode.*  *docker compose up -d*  *# Start all services in detached mode and scale the `spark-worker` service to 1 instance.*  *docker compose up -d --scale spark-worker=1*  *# Stop a specific running container by name or ID.*  *docker stop <container-name-or-id>*  *# Stop all running containers using their container IDs.*  *docker stop $(docker ps -aq)*  *# Display help information for the `docker image ls` command.*  *docker image ls --help*  *# List all locally available Docker images.*  *docker image ls*  *# Remove the specified Docker image.*  *docker image rm dbt-learn-hands-on*  *# List all currently running containers.*  *docker container ls*  *# Display help information for the `docker container ls` command.*  *docker container ls --help*  *# List all containers, including stopped ones.*  *docker container ls -a*  *# Start a stopped container by name or ID.*  *docker container start <container-name-or-id>*  *# Stop a running container by name or ID.*  *docker container stop <container-name-or-id>*  *# Remove all stopped containers.*  *docker container rm $(docker ps -aq)*  *# Forcefully remove all containers, including running ones.*  *docker container rm -f $(docker container ls -aq)*  *# Forcefully remove all Docker images.*  *docker image rm -f $(docker image ls -aq)*  *# Open an interactive shell session inside the specified running container.*  *docker exec -it <container-name-or-id> /bin/bash*  *# Rebuild Docker images without using cache and start services in detached mode, forcefully recreating them.*  *docker compose build --no-cache && docker compose up -d --force-recreate*  *# Copy a file or directory from the host machine to a running container.*  *docker cp <host-path> <container-name-or-id>:<container-destination-path>*  *# Copy a file or directory from a running container to the host machine.*  *docker cp <container-name-or-id>:<container-source-path> <host-destination-path>* |