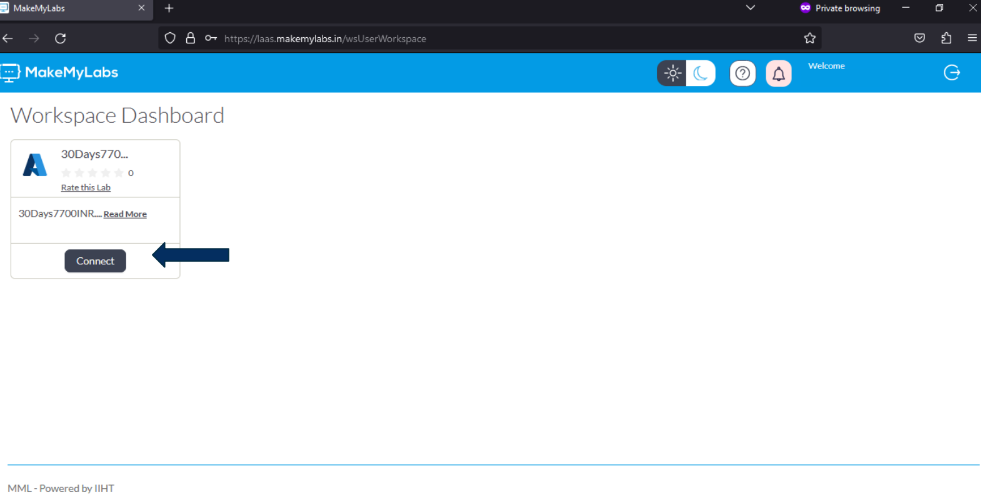
28/12/2023

Chakradhar Bhogapurapu, BIBA

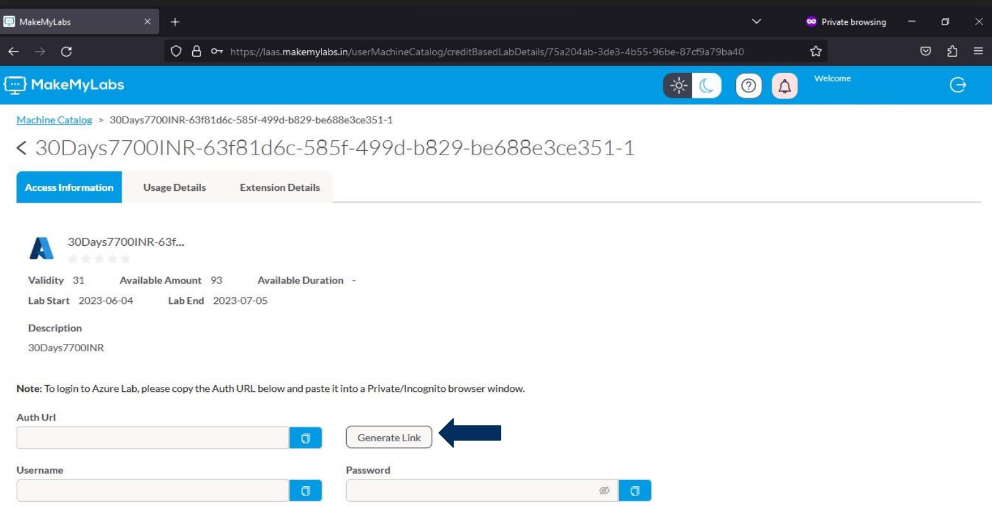
Azure Databricks



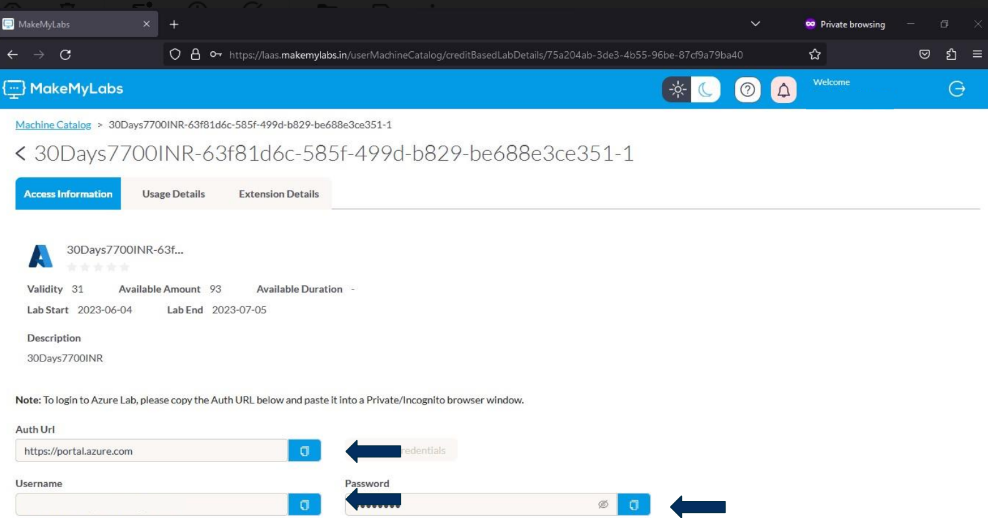
To login into azure databricks, we need to enter the login credentials in the given input fileds.



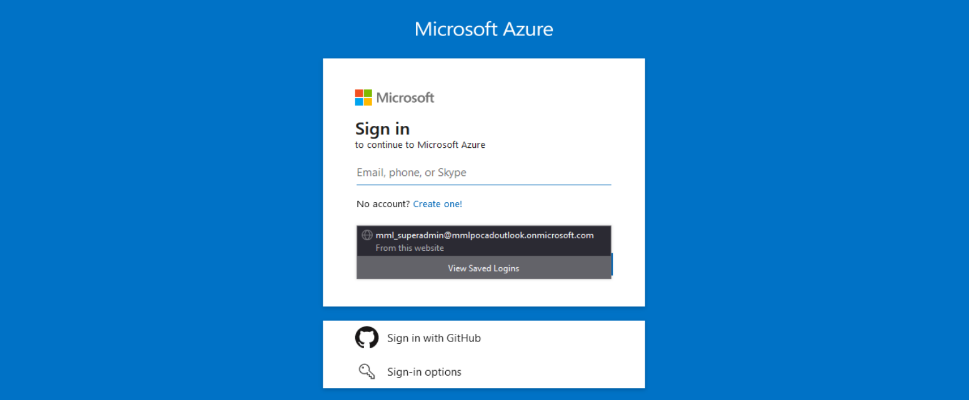
After logging in, we need to select the “connect” button.



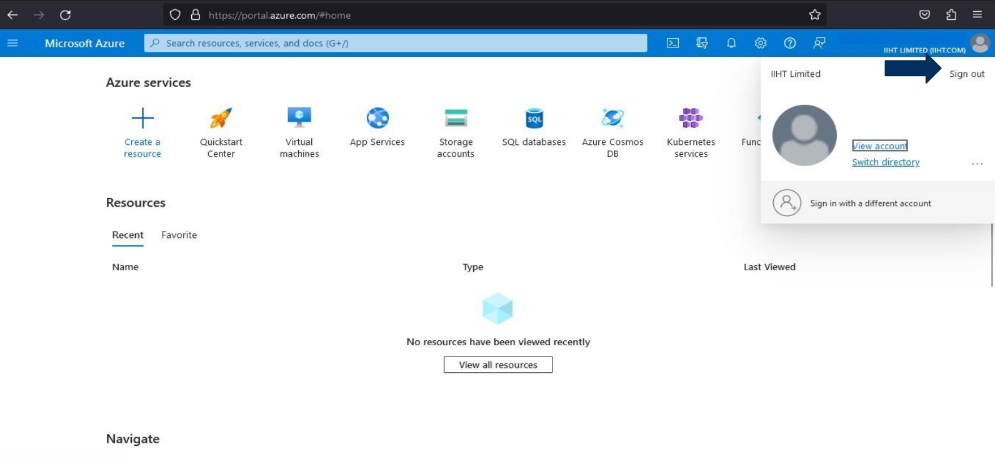
After hitting connect button, the above shown page is displayed. Now click on generate link.



After clicking the generate link button, we will get the link that will redirected to Microsoft Azure login page. Also the username and password required to login to the Microsoft azure account, will be displayed.

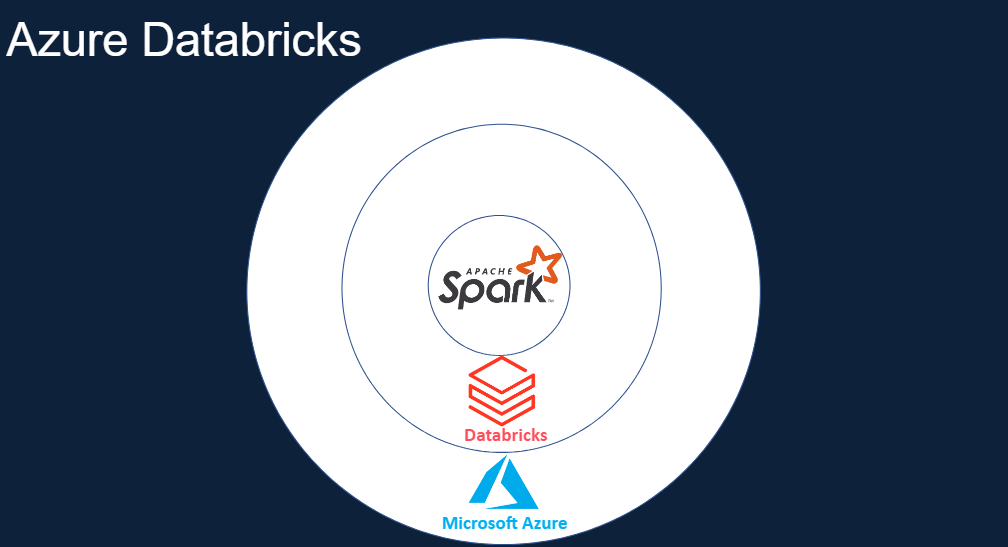


The above shown image is the login page for Microsoft azure account.



After giving the login credentials, we will get logged into the azure account. To sign out, click on the profile icon and we get a sign out button. By clicking it, we will get signed out.

🡪Introduction to Azure Databricks

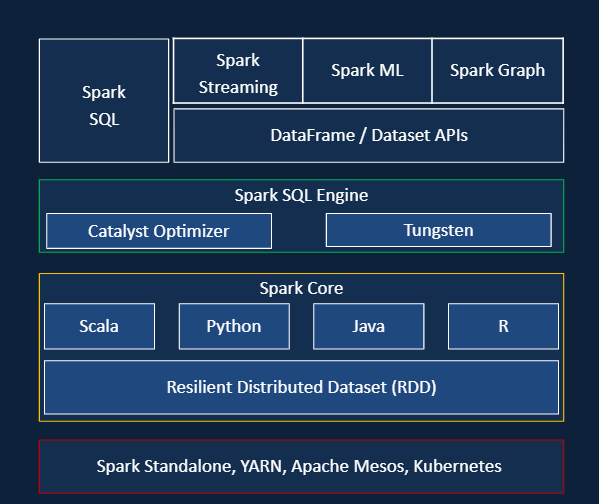


Apache Spark is integrated with in Databricks. Databricks is integrated with in Microsoft Azure.

Apache Spark is a lightning-fast unified analytics engine for big data processing and machine learning.

It is simple and easy to use APIs

It has In-memory processing engine

🡪Apache Spark Architecture

🡪Spark has clusters, Notebook, Administration on controls, Databases, SQL Analytics, ML Flow etc.

🡪Azure Dataabricks has Messaging services, Power BI, Azure ML, Azure Data Factor, Azure Dev Ops, etc.

🡪Databricks workspace components

* Notebook
* Jobs
* Clusters
* Data
* Models

🡪Cluster configuration

* Single/Multi node
* Access Mode : Like single user(only one user access), shared(multiple user access), customer(legacy configuration), etc.
* Databricks Runtime
* Auto Termination : The cluster terminates after x minutes of inactivity
* Auto scaling : User specifies the min and max work nodes. Not recommended for streaming workloads.
* Cluster VM Type/Size
* Cluster Policy

🡪There will be pay as you go type that will be only paid for what we have used. We need to stop the resources when ever there is no use with the cloud.

🡪Unnecessary cloud usage will result in loss of huge monry.