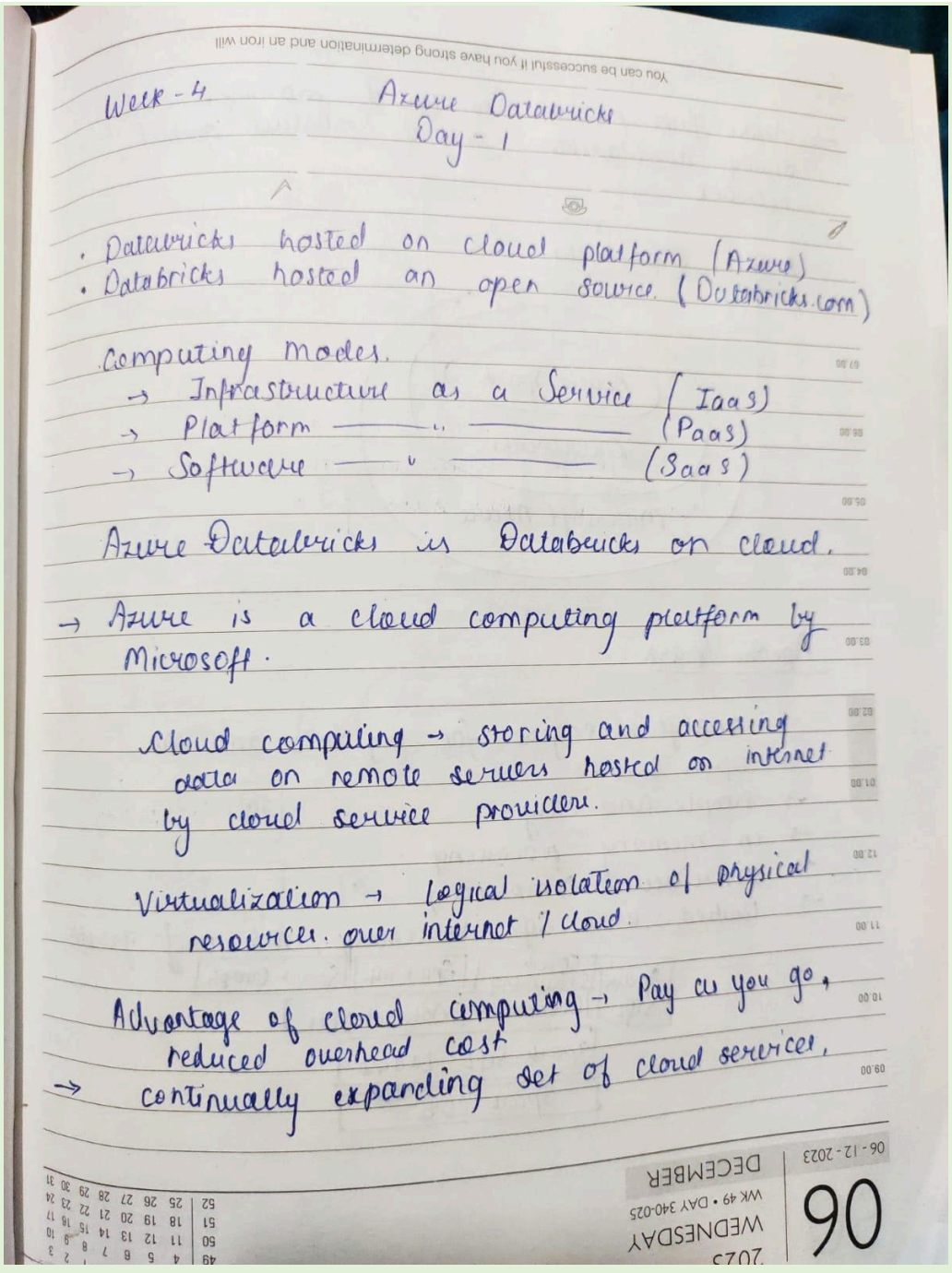
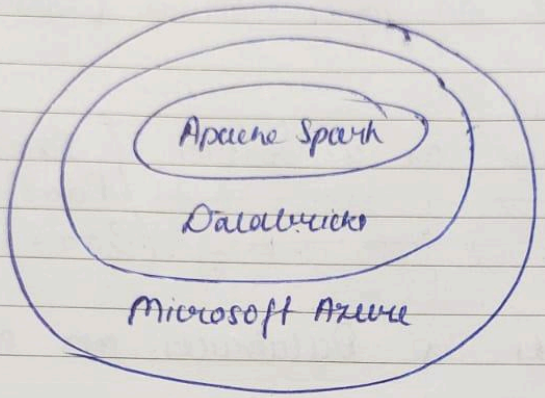
Azure Databricks Assignment - 1
Mitushi Vishwakarma

- **Introduction to databricks**
- **Notes :**



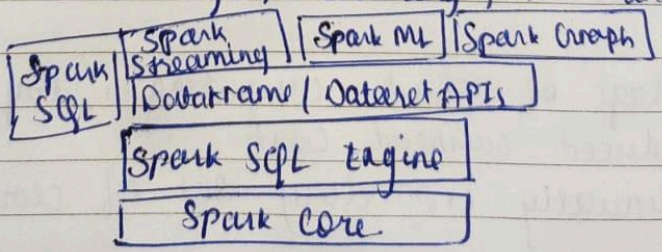
→ Azure gives freedom to build, manage and deploy applications on a massive global network

Azure Databricks



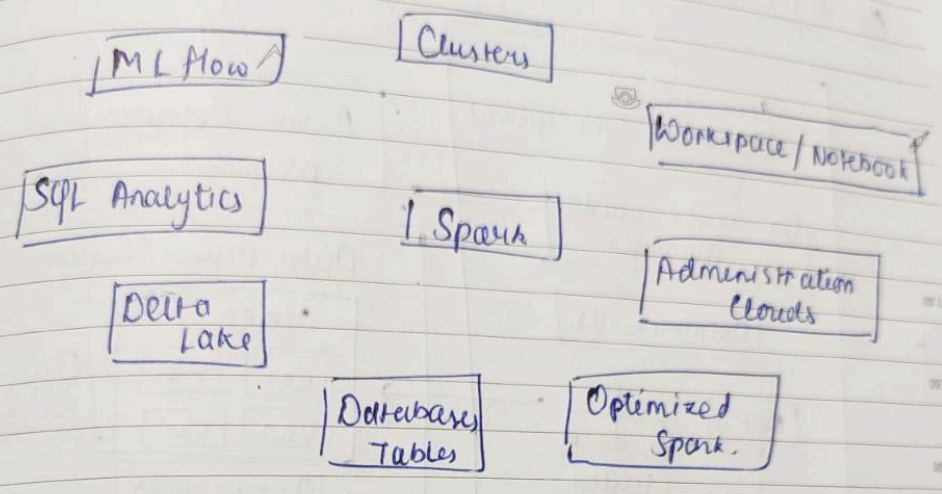
Apache Spark :-

→ analytical engine for big data and ML
→ Open source
→ simple and easy
→ in-memory processing
→ distributed computing
→ unified with SQL, streaming, ML and graphx

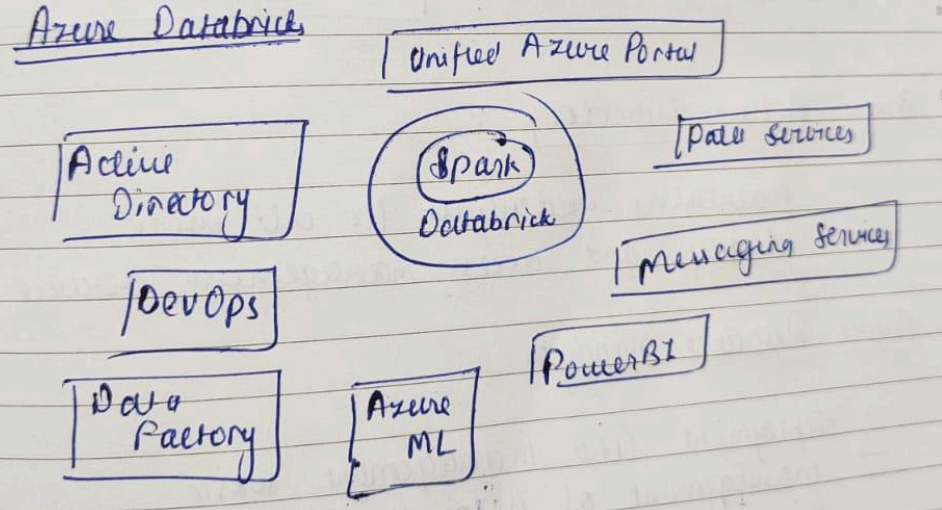


Strengthening our identity is one way of reinforcing people's confidence

Databricks.



Azure Databricks

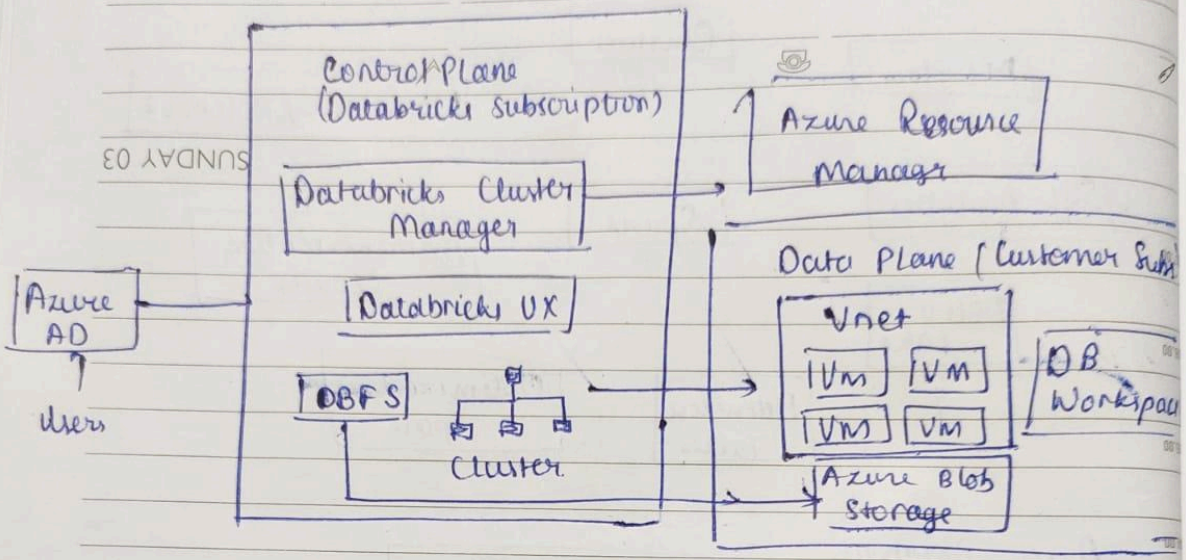


04
2023
MONDAY
WK 49 • DAY 338-027
DECEMBER
04-12-2023

AD → Active Directory

Individual commitment to a group effort is what makes a company work

Architecture



Azure Active Directory :-

→ Maintains credentials for all users
→ Identity and access management service

Azure Resource Manager :-

→ deployment and management service
→ management of resources in Azure account

Azure Blob Storage :-

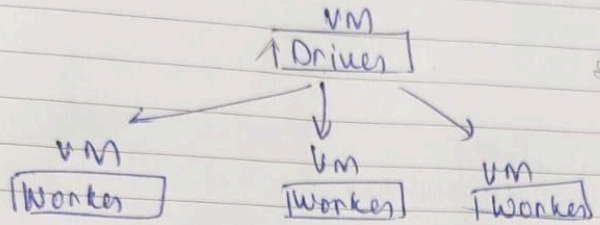
→ Scalable storage in cloud.

SUNDAY 03

02 | DECEMBER | 02-12-2023
SATURDAY | WK 48 • DAY 336-029
2023

Any one can be an ACE. Attitude + Commitment = Excellence

Databricks Clusters



Driver sends data to Workers.
We run query on databricks notebook it connect to driver then to workers and gives o/p.

Cluster Types

All purpose

Created manually
expensive
shared
persistent
for interactive workloads

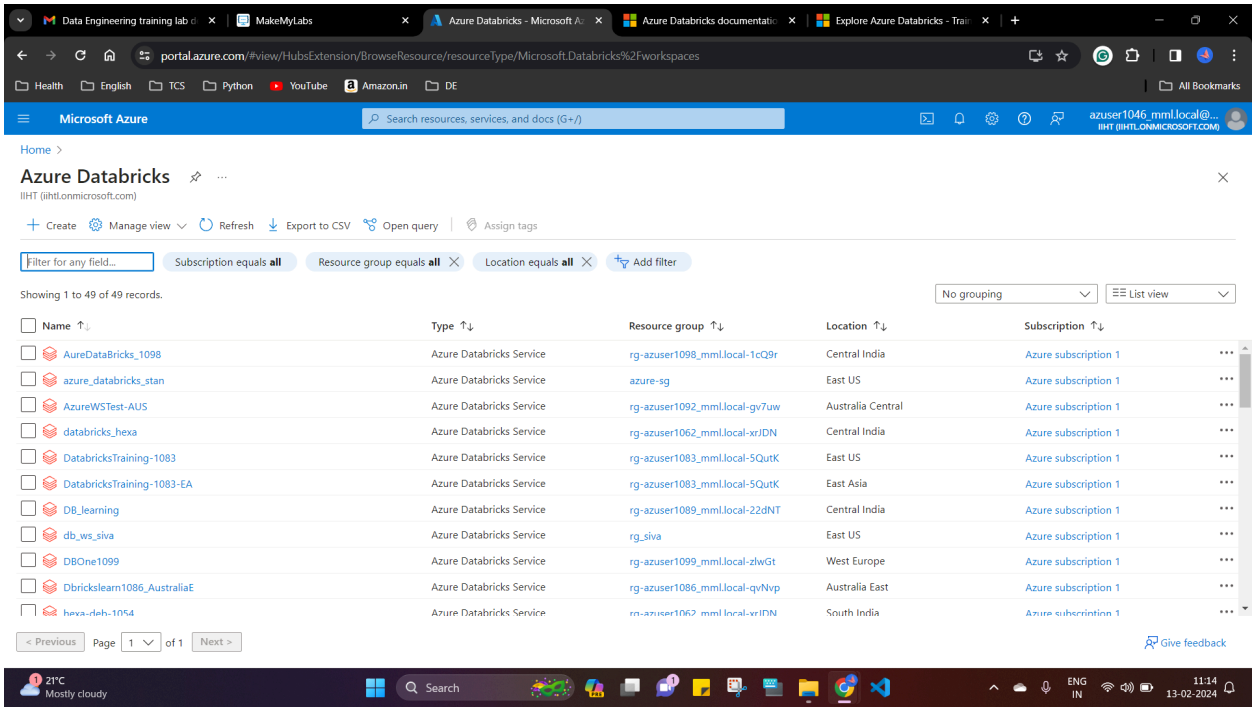
Job Cluster

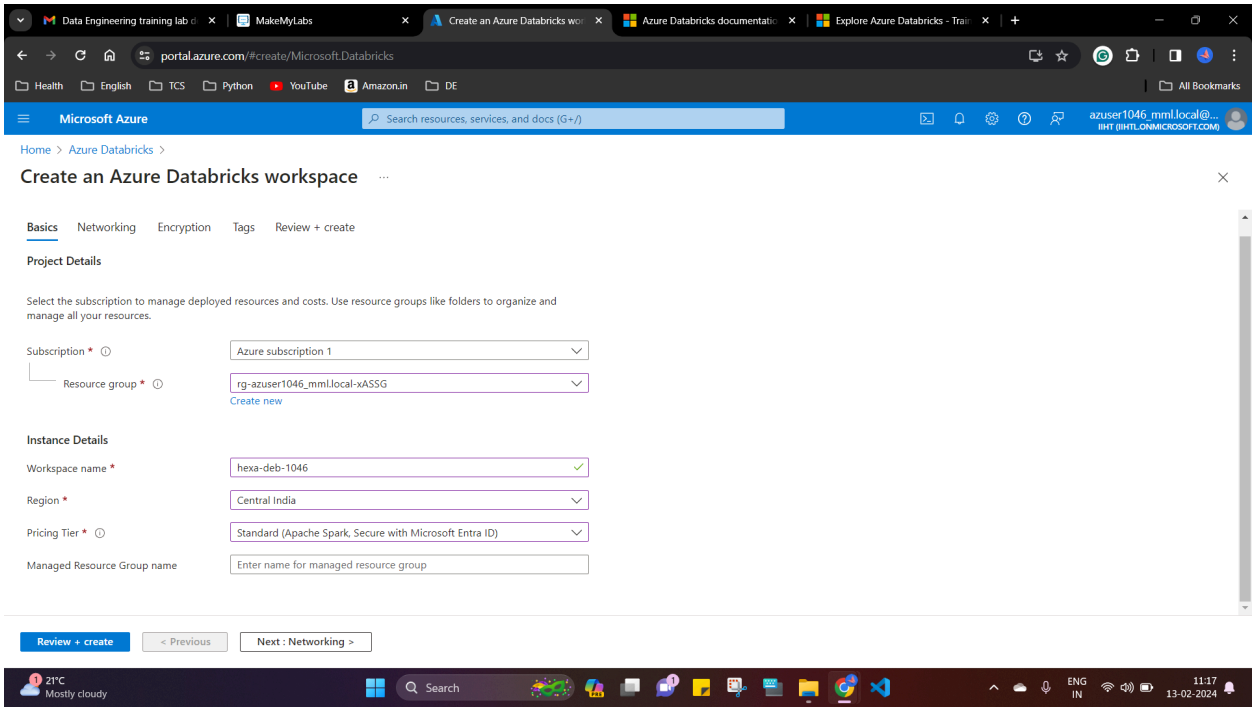
created by job
cheaper.
isolated for job
terminate at end of job
for automation

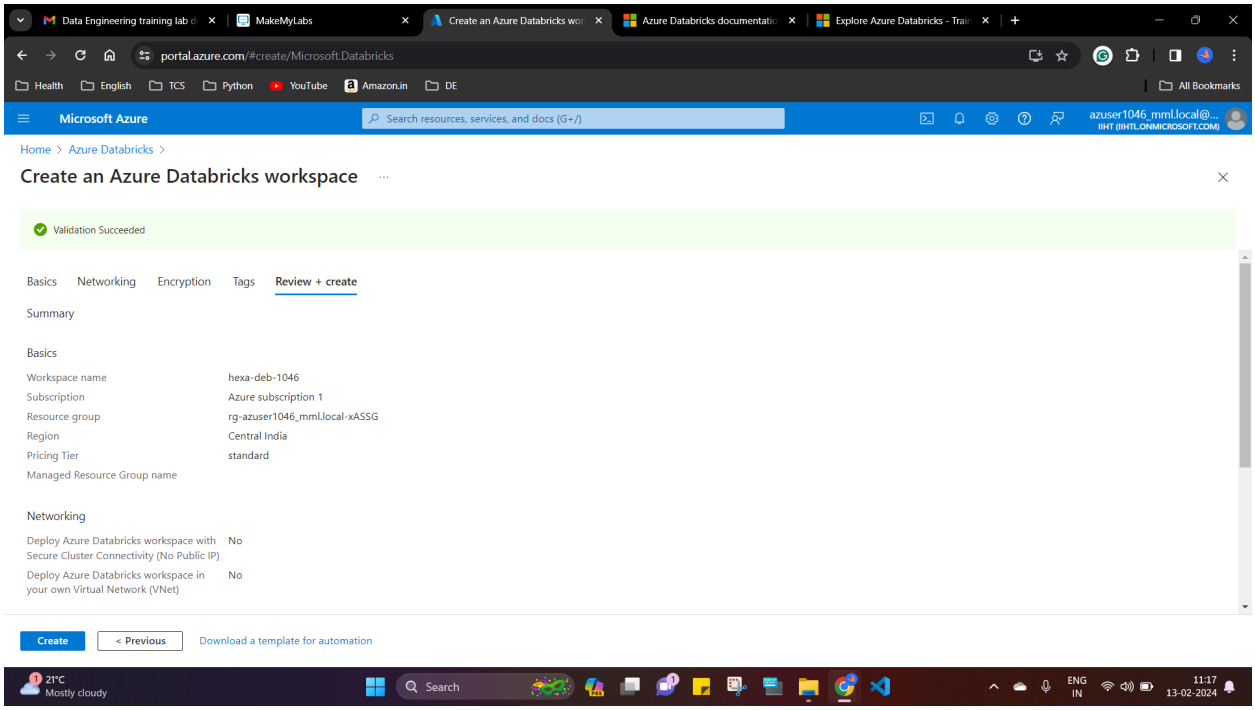
Configuration of cluster decides cost.

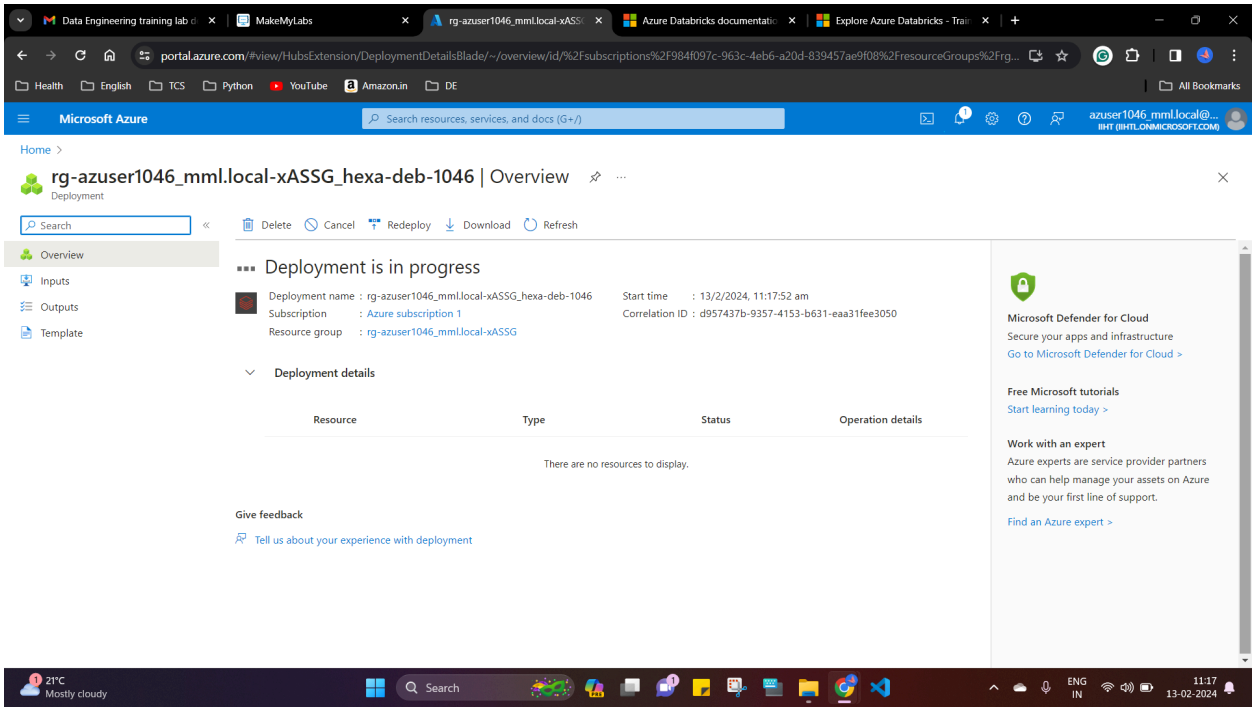
Single / Multi Node. → multiple virtual machines
↓
less cost
single VMs

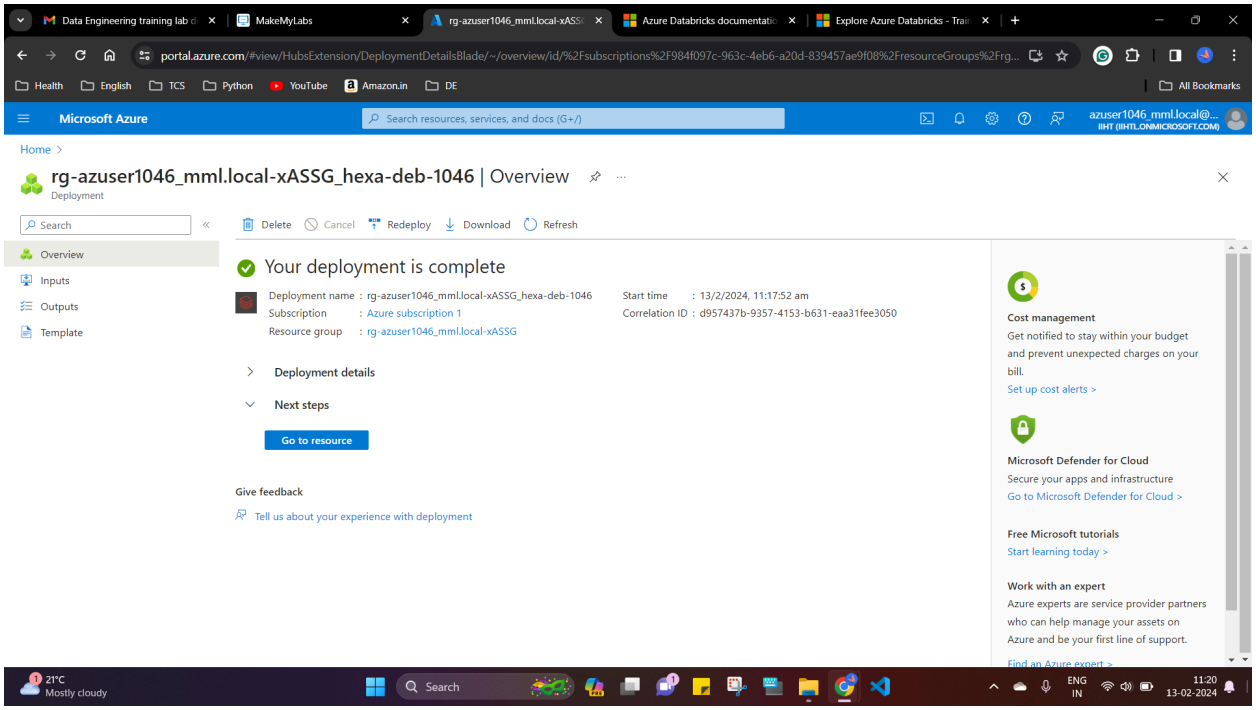
- # Setting up Azure Databricks Workspace

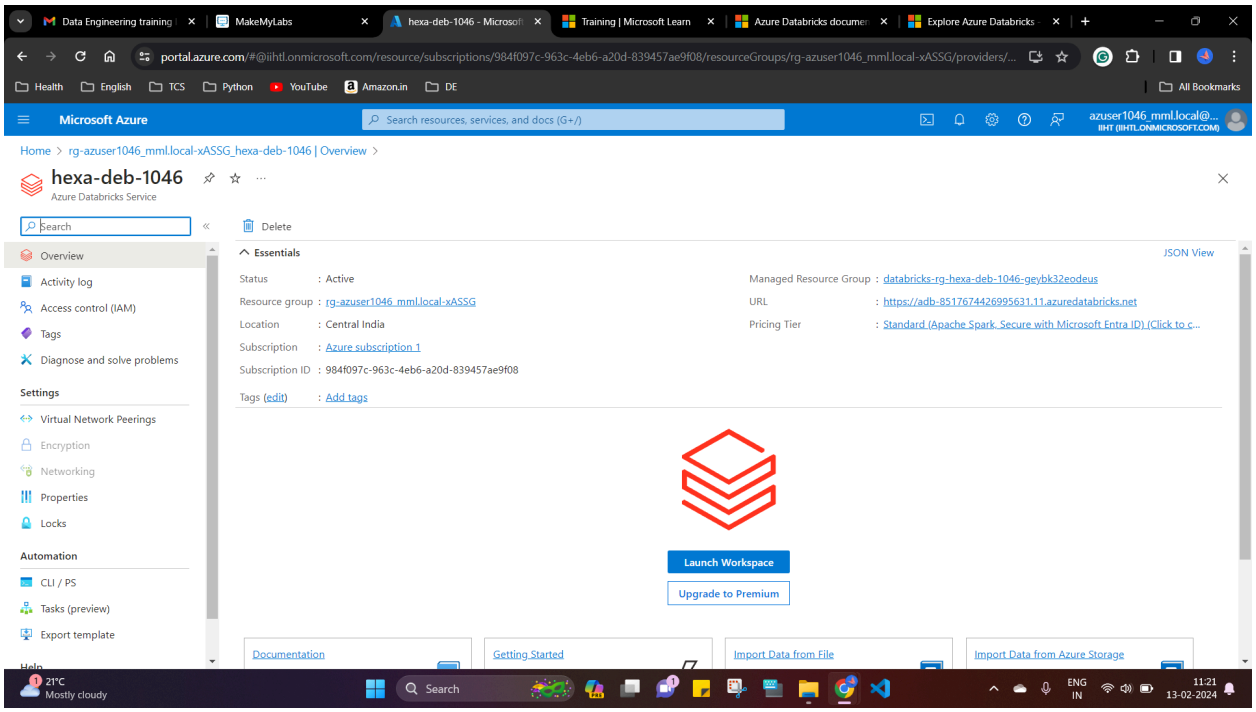




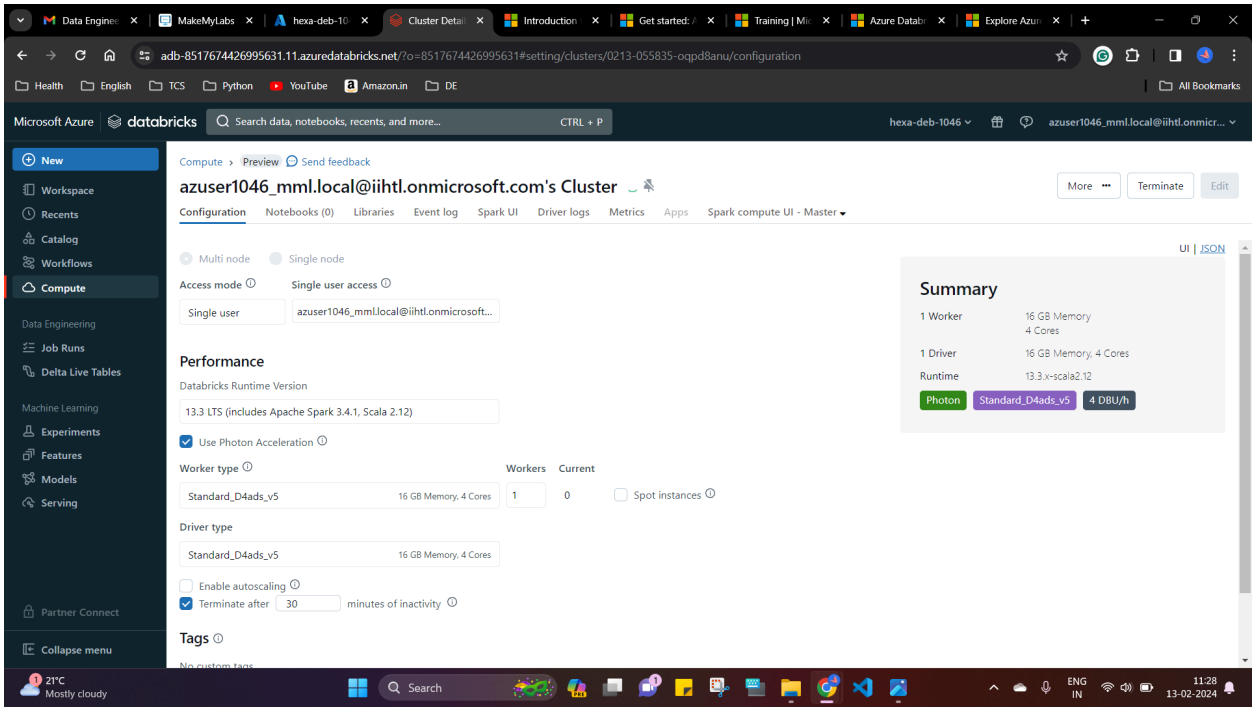

Data Engineering training lab d
MakeMyLabs
Create an Azure Databricks wor
Azure Databricks documentatio
Explore Azure Databricks - Trai
portal.azure.com/#create/Microsoft.Databricks
Health
English
TCS
Python
YouTube
Amazon.in
DE
All Bookmarks
Microsoft Azure
Search resources, services, and docs (G+/)
azuser1046_mml.local@...
IIHT (IIHTL.ONMICROSOFT.COM)
Home > Azure Databricks >
Create an Azure Databricks workspace
Validation Succeeded
Basics
Networking
Encryption
Tags
Review + create
Summary
Basics
Workspace name hexa-deb-1046
Subscription Azure subscription 1
Resource group rg-azuser1046_mml.local-xASSG
Region Central India
Pricing Tier standard
Managed Resource Group name
Networking
Deploy Azure Databricks workspace with Secure Cluster Connectivity (No Public IP) No
Deploy Azure Databricks workspace in your own Virtual Network (VNet) No
Create
< Previous
Download a template for automation
21°C
Mostly cloudy
Search
ENG
IN
11:17
13-02-2024

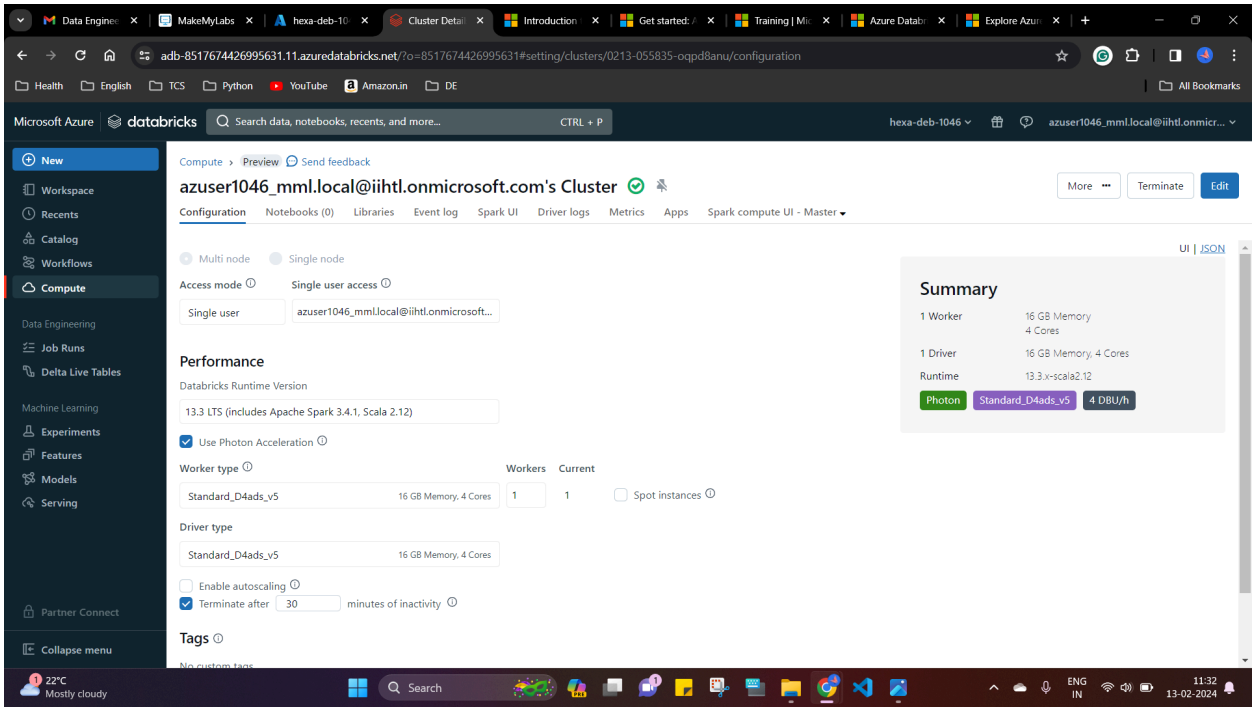

Data Engineering training lab d
MakeMyLabs
rg-azuser1046_mml.local-xASS
Azure Databricks documentatio
Explore Azure Databricks - Trai
portal.azure.com/#view/HubsExtension/DeploymentDetailsBlade/~/overview/id/%2Fsubscriptions%2F984f097c-963c-4eb6-a20d-839457ae9f08%2FresourceGroups%2Frg...
Health
English
TCS
Python
YouTube
Amazon.in
DE
All Bookmarks
Microsoft Azure
Search resources, services, and docs (G+/)
azuser1046_mml.local@...
IIHT (IIHTL.ONMICROSOFT.COM)
Home >
rg-azuser1046_mml.local-xASSG_hexa-deb-1046 | Overview
Deployment
Search
Overview
Inputs
Outputs
Template
Delete
Cancel
Redeploy
Download
Refresh
Deployment is in progress
Deployment name : rg-azuser1046_mml.local-xASSG_hexa-deb-1046
Subscription : Azure subscription 1
Resource group : rg-azuser1046_mml.local-xASSG
Start time : 13/2/2024, 11:17:52 am
Correlation ID : d957437b-9357-4153-b631-eaa31fee3050
Deployment details
Resource
Type
Status
Operation details
There are no resources to display.
Give feedback
Tell us about your experience with deployment
Microsoft Defender for Cloud
Secure your apps and infrastructure
Go to Microsoft Defender for Cloud >
Free Microsoft tutorials
Start learning today >
Work with an expert
Azure experts are service provider partners who can help manage your assets on Azure and be your first line of support.
Find an Azure expert >
21°C
Mostly cloudy
Search
ENG
IN
11:17
13-02-2024


Microsoft Azure
Search resources, services, and docs (G+/)
Home
rg-azuser1046_mml.local-xASSG_hexa-deb-1046 | Overview
Deployment
Search
Overview
Inputs
Outputs
Template
Delete
Cancel
Redeploy
Download
Refresh
Your deployment is complete
Deployment name : rg-azuser1046_mml.local-xASSG_hexa-deb-1046
Subscription : Azure subscription 1
Resource group : rg-azuser1046_mml.local-xASSG
Start time : 13/2/2024, 11:17:52 am
Correlation ID : d957437b-9357-4153-b631-eaa31fee3050
Deployment details
Next steps
Go to resource
Give feedback
Tell us about your experience with deployment
Cost management
Get notified to stay within your budget and prevent unexpected charges on your bill.
Set up cost alerts >
Microsoft Defender for Cloud
Secure your apps and infrastructure
Go to Microsoft Defender for Cloud >
Free Microsoft tutorials
Start learning today >
Work with an expert
Azure experts are service provider partners who can help manage your assets on Azure and be your first line of support.
Find an Azure expert >


Microsoft Azure
Search resources, services, and docs (G+/)
Home
rg-azuser1046_mml.local-xASSG_hexa-deb-1046 | Overview
hexa-deb-1046
Azure Databricks Service
Search
Overview
Activity log
Access control (IAM)
Tags
Diagnose and solve problems
Settings
Virtual Network Peerings
Encryption
Networking
Properties
Locks
Automation
CLI / PS
Tasks (preview)
Export template
Delete
Essentials
JSON View
Status : Active
Resource group : rg-azuser1046_mml.local-xASSG
Location : Central India
Subscription : Azure subscription 1
Subscription ID : 984f097c-963c-4eb6-a20d-839457ae9f08
Tags (edit) : Add tags
Managed Resource Group : databricks-rg-hexa-deb-1046-geybk32eodeus
URL : https://adb-8517674426995631.11.azuredatabricks.net
Pricing Tier : Standard (Apache Spark, Secure with Microsoft Entra ID) (Click to c...
Launch Workspace
Upgrade to Premium
Documentation
Getting Started
Import Data from File
Import Data from Azure Storage

- **Setting up Azure Databricks workspace and configuring clusters**





- **Creating a Databricks notebook and Implementing Databricks for real-time data processing**