🔥 **Notes**
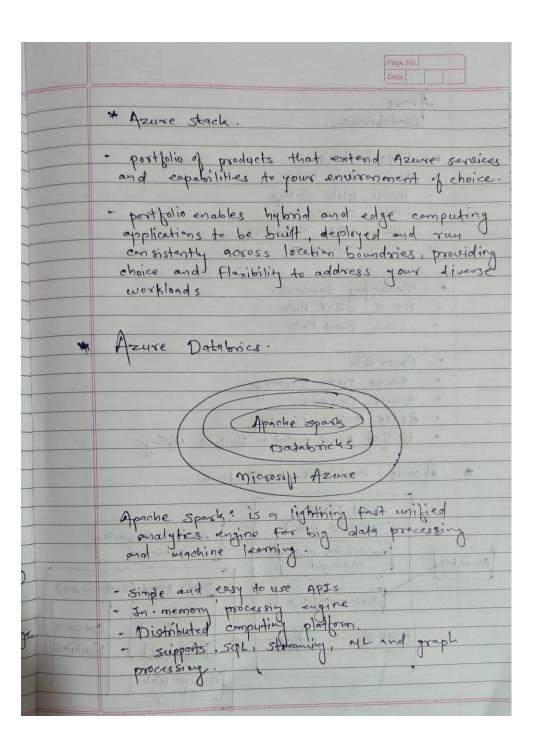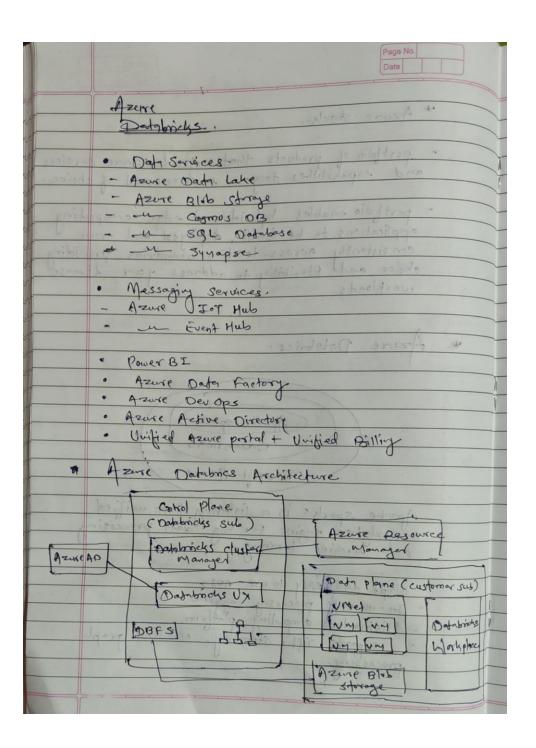
Microsoft Azure

- Azure is microsoft's cloud computing platform.
- Azure gives freedom to build, manage and deploy applications on massive global network using tools and frameworks.

✷ Cloud Computing.

IaaS	PaaS	SaaS
Infrastructure as a service	platform as a service	Software as a service
(Host)	(Build)	(Consume)

- Rapidly setup environments to drive business priorities

- scale to meet peak demands.

- Increase daily activities, efficiency & reduced cost.

✷ Core cloud Services.

Azure compute options.
(VM, Containers, Azure app services, serverless computing)

Azure data storage options.
(Azure sql database, Azure cosmos DB, Azure Blob storage, Azure Data Lake storage)

Azure networking options.
   Azure IoT.

* Azure stack.

- portfolio of products that extend Azure services and capabilities to your environment of choice.

- portfolio enables hybrid and edge computing applications to be built, deployed and run consistently across location boundaries, providing choice and flexibility to address your diverse workloads.

* Azure Databrics.

Apache spark
Databricks
Microsoft Azure

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

- Simple and easy to use APIs
- In-memory processing engine
- Distributed computing platform.
- supports SQL, streaming, ML and graph processing.

Azure
Databricks.

- Data Services
  - Azure Data Lake
  - Azure Blob Storage
  - " Cosmos DB
  - " SQL Database
  - " Synapse

- Messaging Services.
  - Azure IoT Hub
  - " Event Hub

- Power BI
- Azure Data Factory
- Azure DevOps
- Azure Active Directory
- Unified Azure portal + Unified Billing

✱ Azure Databricks Architecture

Control Plane
(Databricks sub)

Databricks cluster
Manager

Azure AD

Databricks UX

DBFS

Azure Resource
Manager

Data plane (customer sub)

VNet

[VM] [VM]

[VM] [VM]

Databricks
Workplace

Azure Blob
Storage

* Databricks Cluster Types

All purpose	Job cluster
- Created manually	Created by jobs
- persistent	Terminated at end of job
- suitable for interactive Workloads	for automated workloads.
- shared among many users	Isolated just for job
- Expensive to run	cheaper to run