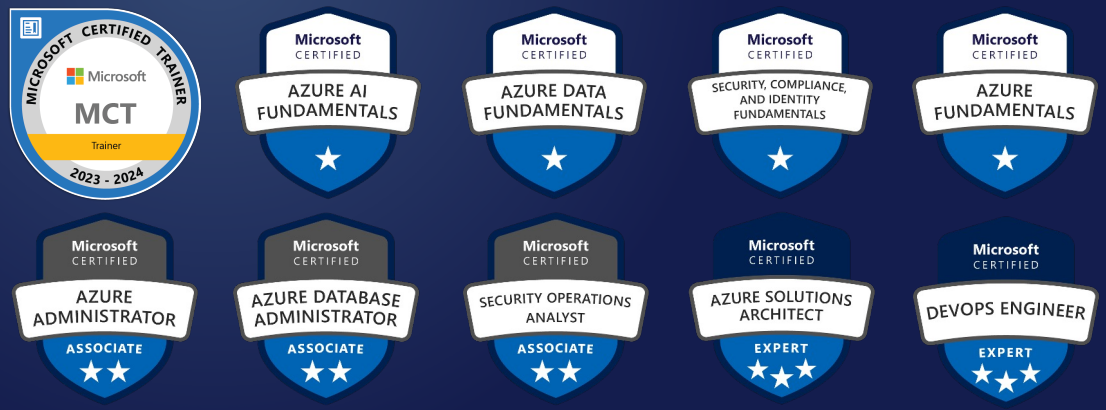
How to copy data from Azure Data Factory to MongoDB Atlas



Piti (Fyi) Champeethong

Sr. Consulting Engineer @ MongoDB Inc

FB: MongoDB Thailand User Group (THMUG)

14-Apr-2023

Agenda

- What is Azure Data Factory?
- What is Azure Storage Account?
- What is MongoDB Atlas?
- Demo
- Q&A

What is Azure Data Factory?

- Cloud-based ETL service
- Data Pipeline / Data Flow
- Integration Runtimes (Azure, Self-Hosted)

Code-Free ETL as a service

Ingest

- Multi-cloud and on-premise hybrid copy data
- 100+ native connectors
- Serverless and auto-scale
- Use wizard for quick copy jobs

Control Flow

- Design code-free data pipelines
- Generate pipelines via SDK
- Utilize workflow constructs: loops, branches, conditional execution, variables, parameters, ...

Data Flow

- Code-free data transformations that execute in Spark
- Scale-out with Azure Integration Runtimes
- Generate data flows via SDK
- Designers for data engineers and data analysts

Schedule

- Build and maintain operational schedules for your data pipelines
- Wall clock, event-based, tumbling windows, chained

Monitor

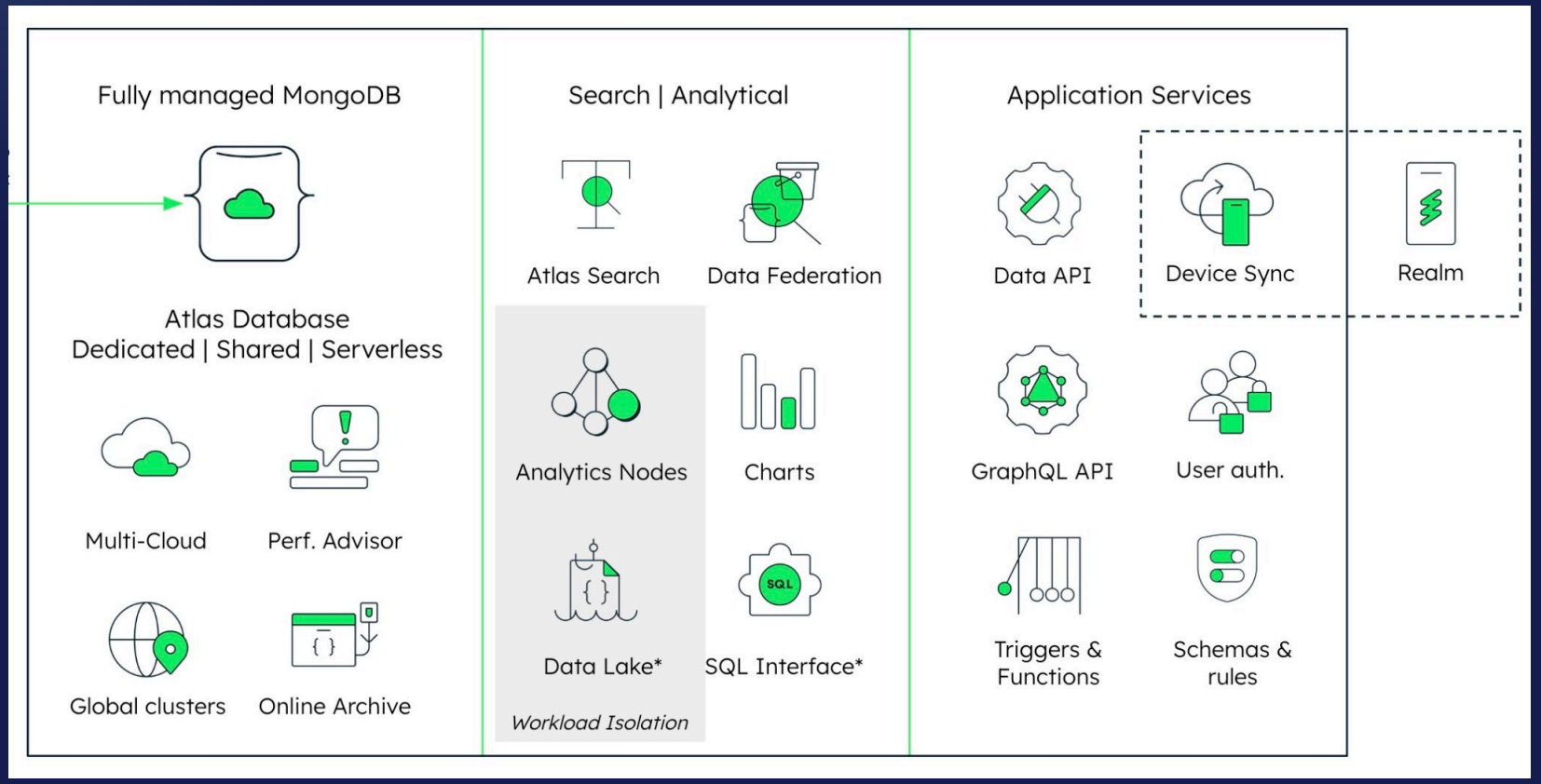
- View active executions and pipeline history
- Detail activity and data flow executions
- Establish alerts and notifications

What is Azure Storage Account?

- Microsoft’s object storage service.
- Redundancy options: LRS, GRS, RA-GRS, ZRS GZRS
- Designed for distributed access.
- Supports REST API, PowerShell, Azure CLI, and client libraries (.NET, Java, Node.js, Python, Go)

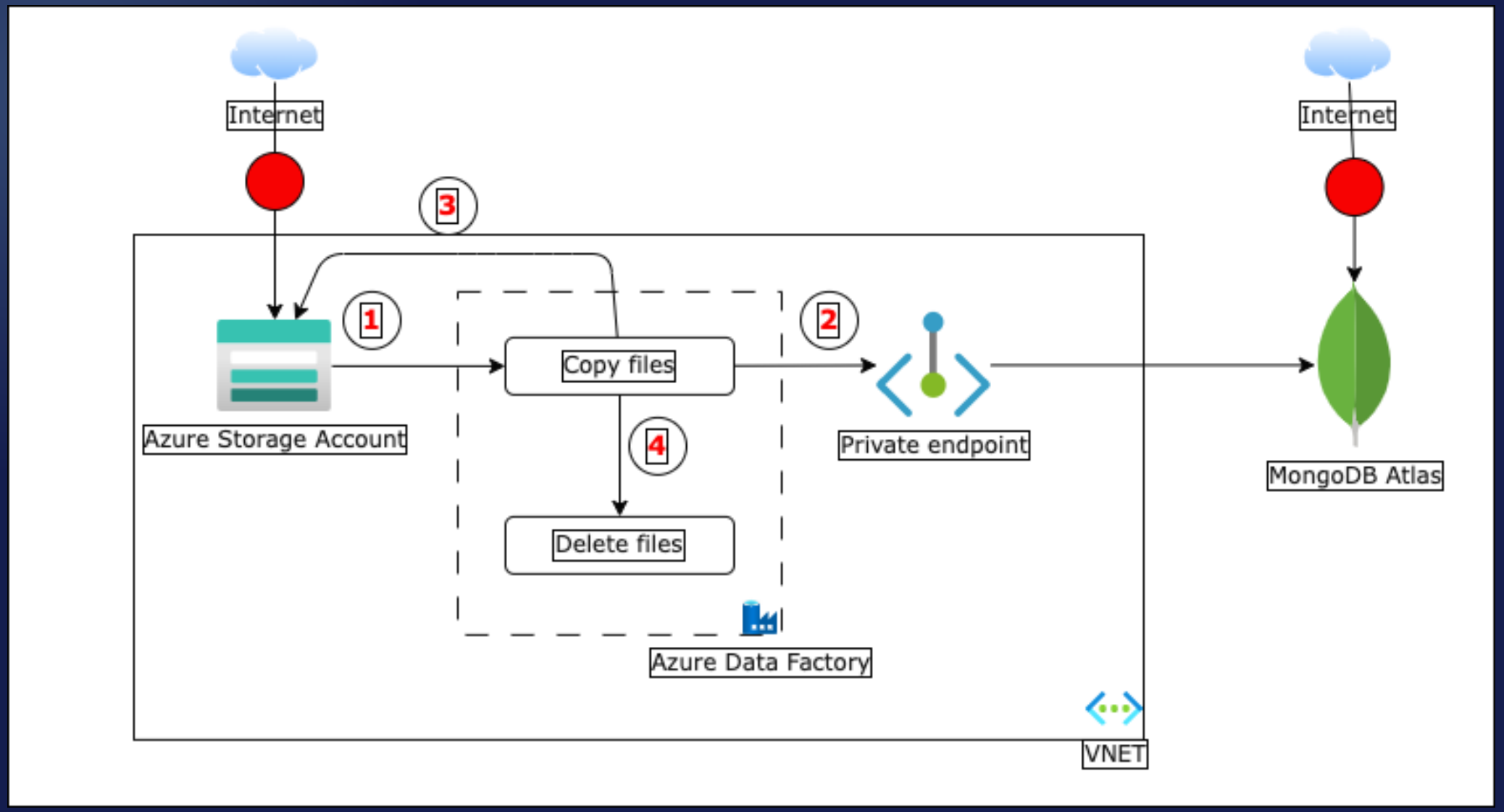
- What is MongoDB Atlas?

- Developer Data Platform (Frontend, API, MongoDB)
- Always free plan: M0, Size: 512MB
- Use Infrastructure from Azure, AWS, GCP



MongoDB

- Demo Time


Internet
3
1
2
Copy files
4
Delete files
Azure Storage Account
Private endpoint
Azure Data Factory
VNET
Internet
MongoDB Atlas

References

Create a self-hosted integration runtime - Azure Data Factory & Azure Synapse | Microsoft Learn

Managed virtual network and managed private endpoints - Azure Data Factory | Microsoft Learn

https://learn.microsoft.com/en-us/azure/data-factory/solution-template-copy-files-multiple-containers

Pipelines - Create Run - REST API (Azure Data Factory) | Microsoft Learn

ninefyi/230414-adfmdb: Copy data from Azure Blob Storage to MongoDB Atlas Using Azure Data Factory (github.com)

THANK YOU!

Q&A