Azure Synapse AnalyticsDuration: 40 hours

Pre-requisites and Important Notes:

- Participants should be comfortable to work with
    - Azure portal
    - Data warehousing concepts
    - ETL Concepts
    - SQL Scripting
- All the participants need to pre-read the study material and PPTs before coming to the session for next day.
- Every participant needs to invest at least 2 hours for offline study as well. However, 4 hours is recommended.

Azure Fundamentals

- What is Azure Data Lake Storage Gen2
- How to create Data Lake Storage Gen2?

Manage a Data Lake Gen2

- Explore Storage Explore (Object & hierarchical file storage)
- Access control & data transfer options
- Hands-on: Shared Access Signature (SAS)
- Encryption
- Hands-on: Monitoring
- Hands-on: Security & Firewall
- Performance

Overview of ETL tools

- Azure SQL
- Azure Data factory
- Azure Data bricks

Azure Synapse Analytics

- What is Azure Synapse Analytics (formerly SQL DW)?
- Azure Synapse Benefits

Azure Synapse Analytics (formerly SQL DW)

- Architecture Azure Synapse MPP Architecture
- Storage and Sharding patterns
- Data Distribution and Distributing Keys Partitioning

Provisioning Dedicated SQL pool

- Hands-on: Create a SQL pool
- Hands-on: Create a server-level firewall rule
- Hands-on: Get the fully qualified server name
- Hands-on: Connect to the server using SSMS as server admin
- Hands-on: Run some queries
- Hands-on: Pause compute
- Hands-on: Resume compute

Synapse Analytics Workspace

- SQL Pool/SQL On Premise/Apache Spark/Synapse Studio
- Design a modern DWH
- SQL pool capabilities - Create Metadata objects and Petabyte scale ingestion
- Understand big data engineering with Apache Spark in Azure Synapse Analytics

Data ingestion with Apache spark notebooks

- Hands-on: Transform data with dataframes in Apache spark pools
- Hands-on: Integrate SQL and Apache spark pools
- Best practices in Azure Synapse Analytics
- Hands-on: Data integration with Azure Synapse pipelines
- Secure and compliance control for sensitive data (Data Security)
- Design hybrid- transactional and Analytical processing using Azure Synapse Analytics

Polybase?

- Why Polybase?
- Performance Consideration in Polybase
- Hands-on: Azure key vaults

Copy and transform data in Azure Synapse Analytics

- Create an Azure Synapse Analytics linked service
- Create Linked service
- Create Dataset
- Copy Activity
- Parallel copy from Azure Synapse Analytics
- Use PolyBase to load data into Azure Synapse Analytics

Azure Synapse Analytics features

- Limitless scale
- Powerful insights
- Unified experience

Scale Dedicated SQL pool

- Hands-on: Connect to the server as server admin
- Hands-on: View service objective
- Hands-on: Scale compute
- Hands-on: Monitor scale change request
- Hands-on: Check dedicated SQL pool (formerly SQL DW) state
- Hands-on: Check operation status

Datawarehouse Concepts and Polybase

- Storage and Sharding Patterns
- Data Distribution and Distributing Keys
- Data Types and Table Types
- Partitioning
- Best Practices for Fact and Dimensiontables
- Hands-on: : Analyze Data distribution before migration Azure
- Different loading methods
- Loading with SSIS vs PolyBase
- Hands-on: Loading with Polybase
- Azure Database vs Azure Datawarehouse
- Hands-on: Monitor Data Storage

Backup and Restore Dedicated SQL pool

- Hands-on: Create user-defined restore points
- Hands-on: Restore an existing dedicated SQL pool
- Troubleshooting

Dedicated SQL pool Workload Management (Mostly conceptual)

- What is workload management?
- Workload management concepts
- Workload classification
- Workload importance
- Workload isolation
- Workload management monitoring
- Memory and concurrency limits
- Resource classes

Secure a dedicated SQL pool

- Connection security
- Advance data security
- Hands-on: Network Security
- Hands-on: Transparent data encryption
- Hands-on: Dynamic Data Masking
- Hands-on: Access Management
- Hands-on: Authentication
- Hands-on: Authorization
- Encryption

Monitor your Dedicated SQL pool workload

- Hands-on: Monitor connections
- Hands-on: Monitor query execution
- Hands-on: Monitor waiting queries
- Hands-on: Monitor tempdb
- Hands-on: Monitor memory
- Hands-on: Monitor transaction log size
- Hands-on: Monitor transaction log rollback
- Hands-on: Monitor PolyBase load

- Hands-on: Monitor query blockings
- Hands-on: Retrieve query text from waiting and blocking queries

Securing

- What is Data Discovery & Classification?
- Hands-on: Discover, classify, and label sensitive columns
- Hands-on: Audit access to sensitive data
- Hands-on: Permissions
- SQL vulnerability assessment
- Hands-on: Enable vulnerability assessment
- Find and remediate vulnerabilities in your Azure SQL databases

Best practices for dedicated SQL pools (Only conceptual)

- Reduce cost with pause and scale
- Maintain statistics
- Tune query performance
- Group INSERT statements into batches
- Use PolyBase to load and export data quickly
- Load then query external tables
- Hash distribute large tables
- Do not over-partition
- Minimize transaction sizes
- Reduce query result sizes
- Use the smallest possible column size
- Use larger resource class to improve query performance
- Use smaller resource class to increase concurrency
- Use DMVs to monitor and optimize your queries

Troubleshoot dedicated SQL pool (Only conceptual)

- Connection related issues
- Tools to troubleshoot issues
- Data ingestion and preparation related issues
- Performance related issues
- System management related issues
- Issues due to Differences from SQL Database