

## Course duration

- 1 day

## Course Benefits

- Learn to gain a deeper knowledge and understanding of the Azure SQL Data Warehouse Architecture and how to write it.

## Course Outline

1. Introduction to the Azure SQL Data Warehouse
  1. Introduction to the Family of SQL Server Products
  2. Introduction to the Family Continued
  3. Microsoft Azure SQL Data Warehouse
  4. Symmetric Multi-Processing (SMP)
  5. What is Parallel Processing?
  6. The Basics of a Single Computer
  7. Data in Memory is fast as Lightning
  8. Parallel Processing of Data
  9. A Table has Columns and Rows
  10. The Azure SQL Data Warehouse has Linear Scalability
  11. The Architecture of the Azure SQL Data Warehouse
  12. Nexus is now available on the Microsoft Azure Cloud
  13. The MPP Engine is the Optimizer
  14. The Azure SQL Data Warehouse System
  15. The Azure SQL Data Warehouse System is Scalable
  16. The Control Node
  17. The Data Rack
  18. The Landing Zone
  19. The Backup Node
  20. Software as a Service (SaaS) and the Elastic Database
  21. Azure Data Lake
  22. Azure Disaster Recovery
  23. Security and Compliance
  24. How to Get an EXPLAIN Plan
2. The Azure SQL Data Warehouse Table Structures
  1. The 5 Concepts of Azure SQL Data Warehouse Tables
  2. Tables are Either Distributed by Hash or Replicated (1 of 5)
  3. Table Rows are Either Sorted or Unsorted (2 of 5)
  4. Tables are Stored in Either Row or Columnar Format (3 of 5)
  5. Tables can be Partitioned (4 of 5)

6. There are Permanent, Temporary and External Tables (5 of 5)
7. Creating a Table with a Distribution Key
8. Creating a Table that is replicated
9. Distributed by Hash vs. Replication
10. The Concept is all about the Joins
11. Creation of a Hash Distributed Table with a Clustered Index
12. A Clustered Index Sorts the Data Stored on Disk
13. Each Node Has 8 Distributions
14. How Hashed Tables are Stored among a Single Node
15. Hashed Tables Will Be Distributed Among All Distributions
16. Creation of a Replicated Table
17. How Replicated Tables are Stored among a Single Node
18. Replicated Table will be duplicated among Each Node
19. Distributed by Replication
20. How Hashed and Replicated Tables Work Together
21. Tables are stored as Row-based or Column-based
22. Creation of a Columnar Table that is hashed
23. How Hashed Columnar Tables are Stored on a Single Node
24. How Hashed Columnar Tables are Stored on All Distributions
25. Comparing Normal Table vs. Columnar Tables
26. Columnar can move just One Segment to Memory
27. Segments on Distributions are aligned to rebuild a Row
28. Why Columnar?
29. Columnar Tables Store Each Column in Separate Pages
30. Visualize the Data – Rows vs. Columns
31. Creation of a Columnar Table that is replicated
32. Creating a Partitioned Table per Month
33. A Visual of One Year of Data with Range per Month
34. Another Create Example of a Partitioned Table
35. Creating a Partitioned Table per Month That is a Columnstore
36. Visual of Row Partitioning and Columnar Storage
37. CREATE TABLE AS (CTAS) Example
38. Creating a Temporary Table
39. Facts about Tables
3. Hashing and Data Distribution
  1. Distribution Keys Hashed on Unique Values Spread Evenly
  2. Distribution Keys with Non-Unique Values Spread Unevenly
  3. Best Practices for Choosing a Distribution Key
  4. The Hash Map determines which Distribution owns the Row
  5. The Hash Map determines which Node will own the Row
  6. A Review of the Hashing Process
  7. Non-Unique Distribution Keys have Skewed Data
4. The Technical Details
  1. Every Node has the Exact Same Tables
  2. Hashed Tables are spread across All Distributions
  3. The Table Header and the Data Rows are Stored Separately
  4. A Distribution Stores the Rows of a Table inside a Data Block

5. To Read a Data Block a Node Moves the Block into Memory
6. A Full Table Scan Means All Nodes Must Read All Rows
7. Rows are organized inside a Page
8. Moving Data Blocks is Like Checking in Luggage
9. As Row-Based Tables Get Bigger, the Page Splits
10. Data Pages are Processed One at a Time per Unit
11. Creating a Table that is a Heap
12. Heap Page
13. Extents
14. Creating a Table that has a Clustered Index
15. Clustered Index Page
16. The Row Offset Array is the Guidance System for Every Row
17. The Row Offset Array Provides Two Search Options (1 of 2)
18. The Row Offset Array Provides Two Search Options (2 of 2)
19. The Row Offset Array Helps with Inserts
20. B-Trees
21. The Building of a B-Tree for a Clustered Index (1 of 3)
22. The Building of a B-Tree for a Clustered Index (2 of 3)
23. The Building of a B-Tree for a Clustered Index (3 of 3)
24. When Do I Create a Clustered Index?
25. When Do I Create a Non Clustered Index?
26. B-Tree for Non Clustered Index on a Clustered Table (1 of 2)
27. B-Tree for Non Clustered Index on a Clustered Table (2 of 2)
28. Adding a Non Clustered Index to A Heap
29. B-Tree for Non Clustered Index on a Heap Table (1 of 2)
30. B-Tree for Non Clustered Index on a Heap Table (2 of 2)
31. Max Levels on the Azure SQL Data Warehouse
32. Azure SQL Data Warehouse Data Types
33. Character Data Types for SQL Server
34. Numeric Data Types for SQL Server
35. Date and Time Data Types for SQL Server
36. Additional Data Types for SQL Server
5. CREATE Statistics
  1. CREATE Statistics Syntax
  2. CREATE Statistics on a Percentage of a Table
  3. CREATE Statistics on a Sample by Using the System Default
  4. CREATE Statistics on a Multi-Column Join Key
  5. What to Column(s) to CREATE Statistics On
  6. CREATE Statistics Using a WHERE Clause
  7. Updating All Statistics on a Table
  8. Updating Only Certain Statistics on a Table
  9. Dropping Statistics on Certain Statistics on a Table
  10. Showing the Statistics
  11. DBCC SHOW\_STATISTICS
  12. DBCC SHOW\_STATISTICS WITH HISTOGRAM

## Class Materials

Each student will receive a comprehensive set of materials, including course notes and all the class examples.