

## Course duration

- 1 day

## Course Benefits

- Learn to gain a deeper knowledge and understanding of the Amazon Redshift Architecture and how to write it.

## Course Outline

1. What is Columnar?
  1. What is Parallel Processing?
  2. The Basics of a Single Computer
  3. Data in Memory is Fast as Lightning
  4. Parallel Processing of Data
  5. A Table has Columns and Rows
  6. Each Parallel Process Organizes the Rows inside a Data Block
  7. Moving Data Blocks is Like Checking in Luggage
  8. Facts That Are Disturbing
  9. Why Columnar?
10. Row Based Blocks vs. Columnar Based Blocks
11. As Row-Based Tables Get Bigger, the Blocks Split
12. Data Blocks Are Processed One at a Time Per Unit
13. Columnar Tables Store Each Column in Separate Blocks
14. Visualize the Data – Rows vs. Columns
15. Row Based Blocks Can Waste Memory Space and Resources
16. The Architecture of Redshift
17. Redshift has Linear Scalability
18. Distribution Styles
19. Distribution Key Where the Data is Unique
20. Another Way to Create A Table
21. Distribution Key Where the Data is Non-Unique
22. Distribution Key is ALL
23. Even Distribution Key
24. Matching Distribution Keys for Co-Location of Joins
25. Big Table / Small Table Joins
26. Fact and Dimension Table Distribution Key Designs
27. Improving Performance By Defining a Sort Key
28. Sort Keys Help Group By, Order By and Window Functions
29. Each Block Comes With Metadata
30. How Data Might Look On A Slice

31. Creating Three Tables with Different Sort Key Strategies
32. A Table with a Distribution Key and a Single-Sortkey
33. A Normal Sort Key Example
34. Creating a Table with an Interleaved Sort Key
35. Interleaved Vs. a Normal Sort Key
36. The ANALYZE Command Collects Statistics
37. Redshift Automatically ANALYZES Some Create Statements
38. What is a Vacuum?
39. When is a Good Time to Vacuum?
40. The VACUUM Command Grooms a Table
41. Database Limits
42. Creating a Database
43. Creating a User
44. Dropping a User
45. Inserting into a Table
46. Renaming a Table or a Column
47. Adding and Dropping a Column to a Table
2. Best Practices for Table Design
  1. Converting Table Structures to Redshift
  2. Converting Table Structures to Redshift Finale
  3. Best Practices for Designing Tables
  4. Choose the Best Sort Key
  5. Each Block Comes with Metadata
  6. Creating a Sort Key
  7. Sort Keys Help Group By, Order By and Window Functions
  8. Choose a Great Distribution Key
  9. Distribution Key Where the Data is Unique
  10. Matching Distribution Keys for Co-Location of Joins
  11. Big Table / Small Table Joins
  12. Define Primary Key and Foreign Key Constraints
  13. Primary Key and Foreign Key Examples
  14. Use the Smallest Column Size When Creating Tables
  15. Use Date/Time Data Types for Date Columns
  16. Specify Redundant Predicates on the Sort Column
  17. Setting the Statement\_Timeout to Abort Long Queries
3. System Tables
  1. Redshift Has System Tables that Log to Disk (Prefix STL)
  2. Redshift Has System Tables that are Virtual (STV Prefix)
  3. Redshift Has System Catalog Tables Visible to Users
  4. Amazon Redshift System Tables
  5. Trouble Shooting Catalog Table pg\_table\_def
  6. Seeing the System Tables in your Nexus Tree
  7. Catalog Table pg\_table\_def
  8. Checking Tables for Skew (Poor Distribution)
  9. Checking All Statements That Used the Analyze Command
  10. Checking Tables for Skew (Poor Distribution)
  11. Checking for Details About the Last Copy Operation

12. Checking When a Table Has Last Been Analyzed
13. Checking for Column Information on a Table
14. System tables for troubleshooting data loads
15. Determining Whether a Query is Writing to Disk
16. Showing Alert events
17. Showing the Last Queries Run on the System
18. Showing Queries that Last More than One Second
19. Listing Queries From Longest to Shortest for a Particular Day
20. Reporting Queries with High CPU Time
21. Reporting Queries of Nested Loops Returning Many Rows
22. Finding Queries Aborted Because of a Monitoring Rule
23. The Number of MB blocks used by each column in a Table
24. Checking if a Table is Distributed Over All Slices
25. List Schemas and Tables in a Database from the PG Catalog
26. A View to See the State of the system Queues for Workloads
27. SELECT From the WLM\_QUEUE\_STATE\_VW View
28. WLM\_QUEUE\_STATE\_VW View Definitions
29. A View Showing the State of Current Queries and Queues
30. WLM\_QUERY\_STATE\_VW View Definitions

#### 4. Compression

1. Compression Types
2. Byte Dictionary Compression
3. Delta Encoding
4. LZO Encoding
5. Mostly Encoding
6. Runlength encoding
7. Text255 and Text 32k Encodings
8. ANALYZE COMPRESSION
9. Copy

#### 5. Temporary Tables

1. Create Table Syntax
2. Basic Temporary Table Examples
3. More Advanced Temporary Table Examples
4. Advanced Temporary Table Examples
5. Table Limits and CTAS
6. Performing a Deep Copy
7. Deep Copy Using the Original DDL
8. Deep Copy Using a CTAS
9. Deep Copy Using a Create Table LIKE
10. Deep Copy by Creating a Temp Table and Truncating Original
11. CREATING A Derived Table
12. The Three Components of a Derived Table
13. Naming the Derived Table
14. Aliasing the Column Names in The Derived Table
15. Visualize This Derived Table
16. Most Derived Tables are Used To Join To Other Tables
17. Multiple Ways to Alias the Columns in a Derived Table

18. Our Join Example with a Different Column Aliasing Style
19. Column Aliasing Can Default For Normal Columns
20. CREATING A Derived Table using the WITH Command
21. Our Join Example With The WITH Syntax
22. WITH Statement That Uses a SELECT \*
23. A WITH Clause That Produces Two Tables
24. The Same Derived Query shown Three Different Ways
25. Clever Tricks on Aliasing Columns in a Derived Table
26. A Derived Table lives only for the lifetime of a single query
27. An Example of Two Derived Tables in a Single Query
28. Connecting To Redshift Via Nexus

## 6. Explain

1. Three Ways to Run an EXPLAIN
2. EXPLAIN – Steps, Segments and Streams
3. EXPLAIN Terms For Scans and Joins
4. EXPLAIN Terms For Aggregation and Sorts
5. EXPLAIN Terms For Set Operators and Miscellaneous Terms
6. EXPLAIN Terms For Set Operators and Miscellaneous Terms
7. EXPLAIN Example and the Cost
8. EXPLAIN Example and the Rows
9. EXPLAIN Example and the Width
10. Simple EXPLAIN Example and the Costs
11. Look for These Keywords to Track Data Movement
12. EXPLAIN Join Example Using DS\_BCAST\_INNER
13. EXPLAIN Join Example Using DS\_DIST\_NONE
14. EXPLAIN Showing DS\_DIST\_NONE Visually
15. EXPLAIN With a Warning
16. EXPLAIN For Ordered Analytics Such as CSUM
17. EXPLAIN For Scalar Aggregate Functions
18. EXPLAIN For Hash Aggregate Functions
19. EXPLAIN Using Limit, Merge and Sort
20. EXPLAIN Using a WHERE Clause Filter
21. EXPLAIN Using the Keyword Distinct
22. EXPLAIN for Subqueries

## 7. User Defined Functions

1. Creating a User Defined Scalar Function
2. Function Syntax
3. Creating a Simple Function
4. Creating a Function That Shows the Sunday Date of the Week
5. Create a Flight\_Table that Holds Longitude and Latitude
6. A Function Example for Measuring Distance in Miles
7. A Function Example for Measuring Distance
8. Create a Flight\_Table that Holds Longitude and Latitude
9. A Function Example for Comparing Two Numbers
10. A Function Example Using Multiple Tables
11. SQL that Utilizes Two User Defined Functions (UDFs)
12. Function Volatility

13. Amazon Redshift Vs. Python Data Types
14. Privileges
8. Workload Management
  1. Create the WLM\_QUEUE\_STATE\_VW View
  2. SELECT From the WLM\_QUEUE\_STATE\_VW View
  3. WLM\_QUEUE\_STATE\_VW View Definitions
  4. Create the WLM\_QUERY\_STATE\_VW View
  5. WLM\_QUERY\_STATE\_VW View Definitions
  6. Open Up Two Sessions in your Nexus
  7. SELECT From our WLM\_QUERY\_STATE\_VW View
  8. Run a Long-Running Query in Tab
  9. In Tab Run These Two Queries
  10. After Setup of Four Queues
  11. How to use the SET command to Place a Query in a Queue
  12. Checking which Queue the Query is Executing In?
  13. How to Reset the Query Group
  14. Creating and Altering a Group
  15. Admin User Can Still SET to a Different Queue if they Want
  16. Overriding the Concurrency Level

## Class Materials

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