

- COURSE CODE:** DWHMDC
- COURSE TITLE:** Data Warehousing – MDCs (Multidimensional Clustering Tables) & MQTs (Materialized Query Tables) Demystified
- PLATFORM:** DB2 UDB for Linux/UNIX/Windows
- AUDIENCE:** IT developers, database administrators or anyone responsible for a data warehouse or related discipline.
- PREREQUISITES:** Completion of our course 'Data Warehouse Fundamentals' (DWHFUN) or equivalent knowledge.
- DURATION:** 3 days
- SUMMARY:** This course focuses on the use of IBM's Multidimensional Clustering (MDC) and the summary tables (e.g., Materialized Query Tables (MQT)) to design and implement data warehouse and other massive database applications.
Performance issues are central to all discussions.
- OBJECTIVES:** Upon completion of this presentation, the participant should be competent to use MDC and MQT disciplines to design data warehouses in particular and large performance-sensitive databases in general.
Additionally, the course participant should be qualified to select and tune queries to exploit the MDC and MQT implementations.
- HANDS-ON:** About 60%

1. REVIEW OF THE BASICS

- Data warehouse vs. data mart
- Data modelling options
- Data mining
- Metadata
- OLAP vs. OLTP
- ETL (Extract, Transform, Load)
- Operational vs. historical data

2. MULTIDIMENSIONAL CLUSTERING (MDC) TERMINOLOGY

- Definition and scope
- What is clustering?
- What are the benefits of clustering?
- Record based clustering
- Block based clustering (MDC)
- Dimension
- Block
- Slice
- Cell
- Block index
- Dimension block index

3. MDC DESIGN CONSIDERATIONS

- How to identify dimension candidates
- How to estimate storage requirements
- How to estimate the number of cells
- How to control space utilisation
- How to choose extent size and cell density
- How to design the block index for performance
- What queries will benefit from MDC
- MDC and multiple dimensions
- MDC and Multiple indexed columns
- MDC and monotonic columns
- Workshop

4. MDC MAINTENANCE

- During INSERTs
- During UPDATEs
- During DELETEs
- During LOADs
- Logging considerations

5. MDC TABLE IMPLEMENTATION

- Existing table → MDC table
- MDC and SMS table spaces
- MDC Advisor feature of DB2 Design Advisor
- MDC tables and partitioning
- How to use the Create Table wizard to define MDC tables
- Workshop

6. MDC PERFORMANCE CONSIDERATIONS

- Extent size
- Monitoring space utilisation
- Monitoring extent size efficiency
- Monitoring cell efficiency
- Index organisation
- Block vs. RID index
- Range predicates and dimensions
- Index ANDing predicates
- Index ORing predicates
- Nested loop joins
- Multiple table joins
- Workshop

7. MATERIALIZED QUERY TABLES (MQT) OVERVIEW

- Definition and scope
- Some typical uses
- Terms and concepts
- MQT vs. AST (Automatic Summary Table)

8. MQT PLANNING AND IMPLEMENTATION

- MQT table creation – Optimization Advisor
- Refresh option
- Referential integrity and informational constraints
- Collecting statistics on base tables
- Use of DB2 Explain tool
- Ascertain that query is using MQT
- Ascertain most of queries in workload/workload compatible
- Index usage
- Workshop

9. MQT PERFORMANCE DESIGN

- Query allocation and prioritisation
- Generalise local predicates to GROUP BY
- Create materialized view
- Estimate MQT view size
- Query rewrite
- Verify rewrite of query to MQT
- Consolidate MQTs
- Memory requirements
- Assess performance gains
- Load MQT with procedure data
- Refresh option
- Optimization Advisor
- Table replication
- DB2 Explain
- Workshop (SQL & Explain)