

INSTRUCTOR-LED TRAINING COURSE

TERADATA PHYSICAL DATABASE TUNING

Lecture/Lab

ILT 25969

4 Days

COURSE DESCRIPTION

This course builds on concepts learned in the Teradata Physical Database Design course and provides an in-depth examination of the processes and procedures to follow once a Teradata database has been implemented and is in production. Performance and tuning topics involving advanced indexing strategies and querying topics are investigated. The EXPLAIN facility is utilized to analyze querying techniques and optimizer strategies.

AUDIENCE

- ~ Database Administrators
- ~ Architects/Designers
- ~ Application Developers

PREREQUISITES

To get the most out of this training, you should have the following knowledge or experience.

- ~ **Introduction to Teradata Database** (ILT #25964 or WBT #26438)
- ~ **Teradata SQL** (Instructor-led (#25965) or WBT (#54458))
- ~ **Teradata Physical Database Design** (#25968)

COURSE OBJECTIVES

After successfully completing this course, you will be able to:

- ~ Identify various deadlocking issues and related resolutions
- ~ Use various join strategies, including Product Joins, Merge Joins, Exclusion Merge Joins, and Hash Joins
- ~ Use various indexing strategies including Partitioned Primary Indexes, Join Indexes, Hash Indexes, and Aggregate Join Indexes
- ~ Explain the reasons for using Statistics including how they are collected, where they are stored, and related strategies
- ~ Discuss the basic concepts of the EXPLAIN terminology
- ~ Discuss various strategies for tuning queries, including CASE, Large-Table/Small Table Joins and IN-List processing

COURSE OUTLINE*

DAY 1	DAY 2	DAY 3	DAY 4
~ Introduction	~ PPI Strategies	~ Date Internals	~ Multiple Table Join Indexes
~ SQL Parser	~ Multilevel PPI	~ EXPLAIN	~ Single Table Join Indexes
~ Transactions, Requests, and Statements	~ Teradata Columnar	~ SQL Tuning	~ Hash Indexes
~ Database Locks	~ Aggregation and DISTINCT		~ Aggregate Indexes
~ Join Processing Analysis			~ Advanced Statistics

* Timing and topics covered by day may vary

COURSE CONTENT

Module 0 – Introduction

- Recommended Prerequisite Knowledge
- Course Objectives
- Course Modules
- Data Warehouse Usage Evolution

Module 1 - SQL Parser

- ~ Internal, Channel, and LAN Parcels
- ~ Request Parcel
- ~ The Data Parcel
- ~ SQL Parser Overview
- ~ Software Cache
 - ~ Request-to-Steps Cache
 - ~ Request-to-Steps Cache Check
 - ~ Request-to-Steps Cache Logic
- ~ Syntaxer
- ~ Resolver
- ~ Dictionary Cache
- ~ Statistics Cache
- ~ Security
- ~ Optimizer
- ~ Generator
- ~ Apply and Dispatcher
- ~ SQL Parser Review
- ~ Summary
- ~ Review Questions

Module 2 - Transactions, Requests, and Statements

- ~ Transaction Processing Discussion
- ~ Processing SQL
- ~ Statement Processing
- ~ Request Processing
- ~ Teradata Database Mode Implicit Transaction Processing
- ~ Teradata Database Mode Explicit Transaction Processing
- ~ More on Processing Modes
 - ~ BTET – Implicit Transaction Processing
 - ~ BTET – Explicit Transaction Processing
 - ~ BTET – Failed Transaction Processing
- ~ The Request
- ~ BTET – DDL and Transaction Processing
- ~ BTET Mode and Delete Table
- ~ ANSI Mode Transaction Processing
- ~ ANSI – DDL and Transaction Processing
- ~ ANSI Mode and Delete Table
- ~ Comparison of ANSI and Teradata Modes
- ~ BTEQ and Request Processing
- ~ Setting the Transaction Mode in BTEQ
- ~ “SQL Assistant” and Request Processing
- ~ Setting the Mode in “SQL Assistant”
- ~ Teradata Studio and Request Processing

- ~ Changing the Transaction Mode in Teradata Studio
- ~ ANSI Mode and “SQL Assistant”
- ~ BTET Mode and “SQL Assistant”
- ~ Module 2: Summary
- ~ Module 2: Review Questions
- ~ Module 2: Lab Exercises

Module 3 - Database Locks

- ~ Locks
- ~ The Lock Queue
- ~ Locking Modifier
- ~ The NOWAIT Option
- ~ ANSI Isolation Levels
- ~ Isolation Levels and STARTUP Strings
- ~ What Is a Deadlock?
- ~ Distinct Pseudo Tables and the Global Deadlock
- ~ Deadlock Handling
- ~ Locking Inside Views
- ~ Multi-Statement Requests
- ~ Locking Inside Views with no Locking Modifier
- ~ Locking Inside Views
- ~ Local Deadlocks
- ~ Conflicting Locks and Multi-Statement Requests
- ~ Accessing Lock Contentions
- ~ Deadlock Guidelines
- ~ Review Questions
- ~ Lab Exercises

Module 4 - Join Processing Analysis

- ~ Join Processing
- ~ Product Join
- ~ Product Join Explain
- ~ Merge Join
- ~ Merge Join with Duplication and Sorting
- ~ Merge Join Strategy
- ~ Merge Join with Row Redistribution
- ~ Merge Join with Redistribution Explain
- ~ Merge Join with Matching Primary Indexes
- ~ Merge Join on Matching Primary Indexes Explain
- ~ Alternative to Using OR Join Conditions
- ~ Exclusion Joins
- ~ Exclusion Merge Join Explain
- ~ Inclusion Merge Join Explain
- ~ Hash Joins
- ~ Three-Table Join Explain
- ~ Partial Redistribution and Partial Duplication (PRPD)
- ~ PRPD Join Steps Simplified
- ~ Three-Way PRPD Explanation
- ~ Teradata PRPD Join Example
- ~ PRPD One Relation Skewed on a Non-PI Column
- ~ Nested Joins

- ~ Nested Join Explain
- ~ Join Processing Observations
- ~ Review Questions

Module 5 - PPI Strategies

- ~ Logical Example of NPPI versus PPI
- ~ File System Changes
- ~ General Row Layout
- ~ Secondary Index Changes (for PPI)
- ~ NUSIs and Partition Elimination
- ~ Miscellaneous SQL – Row Identifiers
- ~ Index Constraints View
- ~ Utility Considerations
- ~ Join Considerations with PPI
- ~ NPPI to PPI Join – Sliding Window
- ~ NPPI to PPI Join – Hash Ordered Spool File Join
- ~ PPI to PPI Join – Rowkey-Based Join
- ~ EXPLAIN Terminology
- ~ 14.10 EXPLAIN Terminology & Modifiers
- ~ Product Join Enhanced by DPE
- ~ Row Hash Match Scan Merge Join – Enhanced by DPE
- ~ Partitioning Strategies
- ~ Guidelines
- ~ PPI Enhancements
- ~ Teradata 14.0 PPI Enhancements
- ~ Summary
- ~ Review Questions
- ~ Lab Exercises

Module 6 - Multilevel PPI

- ~ Multilevel PPI Concepts
- ~ Multilevel Partitioning Example
- ~ Three-Level PPI Example
- ~ Qualifying All Partition Columns
 - ~ Qualifying #L1 and #L2
 - ~ Qualifying #L1 and #L3
 - ~ Qualifying #L2 and #L3
 - ~ Qualifying #L1
 - ~ Qualifying #L2
- ~ Different PI and Partition Columns
- ~ Evaluating the Partitioning Expression
- ~ Evaluating a Three-Level Partition
- ~ Contrasting Styles
- ~ Things to Consider
- ~ Module 6: Summary
- ~ Module 6: Review Questions
- ~ Module 6: Lab Exercises

Module 7 - Teradata Columnar

- ~ Teradata Columnar Introduction

- ~ Teradata Columnar Benefits
- ~ Columnar vs. Row Format Comparison
- ~ Columnar I/O Savings Illustration
- ~ Columnar Drawback
- ~ No Primary Index Table DDL
- ~ The No Primary Index Table
- ~ Column Partition Table DDL (with Auto-Compression)
- ~ Column Partition Table DDL (without Auto-Compression)
- ~ Column Partition Container (NO AUTO COMPRESS)
- ~ The Column Partition Table (NO AUTO COMPRESS)
- ~ CP Table Query #1 (NO AUTO COMPRESS)
- ~ Auto-Compression for CP Tables
- ~ Auto-Compression Techniques for CP Tables
- ~ User-Defined Compression Techniques
- ~ Column Partition Container (Automatic Compression)
- ~ The Column Partition Table (with Auto-Compression)
- ~ Table Query #2 (with Auto-Compression)
- ~ CP Table with Row Partitioning DDL
- ~ The Column Partition Table (with Row Partitioning)
- ~ CP Table with Multi-Column Container DDL
- ~ The CP Table with Multi-Column Container
- ~ CP Table Hybrid Row & Column Store DDL
- ~ The CP Table (with Hybrid Row & Column Store)
- ~ Populating a CP Table
- ~ DELETE Considerations
- ~ UPDATE and USI/NUSI Considerations
- ~ CP Table Restrictions
- ~ Partitioning Options
- ~ Column-Partitioned Primary Index
- ~ NoPI Table Options
- ~ TD14.10 Available Syntax
- ~ Teradata 14.10 Columnar Enhancements
- ~ TD15.10 – Columnar Table with Primary AMP
- ~ TD15.10 – Columnar Table with Primary Index
- ~ Sample Statistics on PPI Tables
- ~ Summary
- ~ Review Questions
- ~ Lab Exercises

Module 8 - Aggregation and DISTINCT

- ~ Aggregations
- ~ ARSA Algorithm
- ~ A Query and Its Data
 - ~ Local Aggregation
 - ~ Local Aggregation – Cache Overflow
 - ~ Redistributing the Aggregate Cache
 - ~ Sort
 - ~ Global Aggregation
- ~ More on Local vs. Global Aggregation

- ~ DISTINCT Processing
- ~ Step 1: Redistribution
- ~ Step 2: Eliminate Duplicate Rows
- ~ Optimization: Aggregate Processing for DISTINCT
- ~ Multiple COUNT DISTINCT
- ~ Aggregation and Distinct Summary
- ~ Module 8: Review Questions

Module 9 - Date Internals

- ~ Integer vs. Date Arithmetic
- ~ Checking for Valid Dates
- ~ Determining Integers for Dates
- ~ Dates Prior to 1900
- ~ Determining Date vs. Integer Arithmetic
- ~ Using Concatenation to Get Year and Month
- ~ Automating Date Computations
- ~ Using Date Functions
- ~ Using FORMAT to Get Year and Month
- ~ Be Careful When Doing Computations!
- ~ Other FORMAT Options
- ~ Review Questions
- ~ Lab Exercises

Module 10 - EXPLAIN

- ~ EXPLAIN Facility
- ~ EXPLAIN Terminology
 - ~ 15.10 EXPLAIN Terminology
 - ~ 16.00 EXPLAIN Terminology
- ~ Row and Time Estimates
- ~ What is Peeking?
- ~ Peeking at Parameterized Values
- ~ Collected Statistics?
- ~ No Collected Statistics
- ~ With Collected Statistics
- ~ Also with Collected Statistics
- ~ Joins and Incremental Planning and Execution
- ~ Incremental Planning and Execution (IPE)
- ~ EXPLAINing Aggregation
- ~ Parallel Steps
- ~ EXPLAIN with Parallel Steps
 - ~ EXPLAIN of a SELECT (BMSMS)
 - ~ EXPLAIN of a CREATE TABLE
 - ~ EXPLAIN of ORed NUSI's
 - ~ EXPLAINing NOT IN on NULLable Columns
- ~ Single Sender Redistribution (SSR)
- ~ Summary
- ~ Review Questions

Module 11 - SQL Tuning: Part 1

- ~ Non-Correlated Subquery Processing
- ~ Correlated Subquery Processing

- ~ Comparison
- ~ NOT IN vs. Not Equal
- ~ Examining LT/ST Joins
- ~ In-List Processing – Introduction
- ~ In-List Processing
- ~ Enhanced IN-List Star Join Strategy
- ~ IN-List Access Path
- ~ Review Questions

Module 12 - SQL Tuning: Part 2

- ~ Outer Join Defined
- ~ ON/AND (Condition on Inner Table)
- ~ WHERE (Condition on Inner Table)
- ~ ON/AND (Condition on Outer Table)
- ~ WHERE (Condition on Outer Table)
- ~ NOT IN on Nullable Columns
- ~ Beginning with the IN List
- ~ Changing the IN to NOT IN
- ~ Applying Truth Table Logic
- ~ Changing the WHERE Condition
- ~ Outer Join Alternative to Nullable NOT IN
- ~ Writing the Outer Join Alternative
- ~ Analyzing the Outer Join Alternative
- ~ Correlated Subquery for Nullable NOT IN
- ~ Views with Inner Joins
- ~ Views with Outer Joins
- ~ Establishing Uniqueness for the Optimizer
- ~ Partial GROUP BY (PGB) – “Early-Aggregation”
- ~ “Early Aggregation” – Limitation
- ~ Partial GROUP BY (PGB) and “Partial Sum”
- ~ Review Questions
- ~ Lab Exercises

Module 13 - Multiple Table Join Indexes

- ~ Join Indexes
- ~ What Is a Multi-Table Join Index?
- ~ Compressed Join Index
- ~ How Are Compressed Join Indexes Stored?
- ~ An Inner Join Example
- ~ EXPLAINing the Join
- ~ An Outer Join
- ~ Join Index – Outer Join
- ~ The Same Inner Join
- ~ Will the Optimizer Use the Join Index?
- ~ Naming a Column
- ~ The EXPLAIN
- ~ A Three-Way with Outer Join
- ~ The Same Inner Join
- ~ A Two-Table Join Index and a Three-Table Join
- ~ The Three-Way Join
- ~ Avoiding Redistribution

- ~ Fully Covering One Join
- ~ Using a Value-Ordered NUSI
- ~ Value Ordered NUSI and a Range Constraint
- ~ The EXPLAIN
- ~ Value-Ordered Primary Index
- ~ Revisiting the Range Constraint
- ~ The New EXPLAIN
- ~ A Sparse Join Index
- ~ Sparse Join Index Strategies
- ~ Multi-Table Join Index with Multilevel PPI
- ~ Multi-Table Join Index with Multilevel PPI Explain
- ~ Preserve Join Index Column Compression
- ~ More Information on Join Indexes
- ~ Review Questions
- ~ Lab Exercises
- ~ Optional Lab Exercises

Module 14 - Single Table Join Indexes

- ~ A Single-Table Join Index
- ~ A Single-Table Non-Compressed Join Index
- ~ Another Single-Table Join Index
- ~ A Single-Table Join Index Strategy
- ~ The Rewrite
- ~ Column Partitioned Join Index
- ~ Using the Rowid for Partial Covering
- ~ The Strategy
- ~ The Tables
- ~ Using the Rowid
- ~ The Join-Back
- ~ The Strategy
- ~ The USI Alternative
- ~ The Global Index
- ~ Global Index Usage
- ~ Global Usage
- ~ EXPLAIN
- ~ NUSI EXPLAIN
- ~ Review Questions

Module 15 - Hash Indexes

- ~ What is a Hash Index?
- ~ Why Would You Use a Hash Index?
- ~ Comparison to Join Indexes
- ~ Why the Dramatic Difference?
- ~ About Compression and Structure
- ~ A Hash Ordered Strategy
- ~ Join-Back Strategy
- ~ Join-Back EXPLAIN
- ~ Reviewing the Locks
- ~ Value-Ordered, Compressed, Single-Table Join Index
- ~ Value Ordered NUSI Strategy
- ~ Join Index Strategy

- ~ Contrast These Indexes
- ~ The Classic Large-Table/Small-Table Join
- ~ What We'd Like to See
- ~ The EXPLAIN
- ~ Review Questions

Module 16 - Aggregate Indexes

- ~ Aggregate Join Index
- ~ Simple Aggregate Join Index
- ~ Show of Aggregate Join Index
- ~ Computing an Average
- ~ A Simple Aggregate Join Index
- ~ DELETE and the Aggregate Join Index
- ~ Simple Multi-Table Aggregate Index
- ~ Sparse Aggregate Index
- ~ Using EXTRACT with Aggregate Index
- ~ Non-Compressed Join Indexes and Aggregation
- ~ NUSI Aggregation
- ~ Multi-Table Compressed Join Indexes and Aggregation
- ~ An Extreme Multi-Table Compressed Join Index
- ~ Aggregate Join Index with PPI
- ~ Notes on JOIN/HASH Indexes
- ~ Review Questions
- ~ Lab Exercises

Module 17 - Advanced Statistics

- ~ Statistics Review
- ~ Optimizer's Search for Statistics
- ~ Statistics Extrapolation
- ~ Dedicated Statistics Dictionary Cache
- ~ Refresh or Re-Collect Statistics
- ~ COLLECT STATISTICS Command (Index Format with 14.0 Options)
- ~ Statistics Data – What is Collected?
- ~ Table Level Statistics Data
- ~ Table Level Summary Statistics
- ~ Collect Statistics PARTITION Option
- ~ Statistic Histogram Keeps History
- ~ Show Summary Statistics Values on Table
- ~ SUMMARY Statistics after Three Collections
- ~ Rollup Optimization
- ~ ROLLUP Stats Performance Optimization
- ~ Pre-Aggregate Optimization
- ~ SHOW STATISTICS Options
- ~ SHOW STATISTICS Sample Output
- ~ SHOW STATISTICS VALUES Sample Output
- ~ Transfer of Statistics
- ~ HELP STATISTICS
- ~ HELP CURRENT STATISTICS
- ~ Privileges
- ~ Statistics Collection Customization

- ~ Collect Statistics MAXINTERVALS Option
- ~ Collect Statistics MAXVALUELENGTH Option
- ~ User-Specified Column Ordering and Naming
- ~ Single Table Expression Statistics
- ~ Statistics Recollection Optimizations
- ~ Threshold Functionality – Reduce Unnecessary Collections
- ~ Setting Thresholds
- ~ Collect Statistics SAMPLE Option
- ~ Identifying Sampled Percentage Values
- ~ Identifying Unused Statistics
- ~ Finding Missing Statistics
- ~ Teradata 14.10 AutoStats Features
- ~ Teradata Stats Manager Portlet
- ~ Summary
- ~ Review Questions
- ~ Lab Exercises