

Database Design for DB2™

Course No.	1075
Description	This course reveals how decisions made during data modeling can impact the performance of a system once implemented in DB2 UDB®. In addition to instructor led discussion and presentations students will be given a chance to identify potential problems and suggest solutions during several group-oriented workshops and a case study held throughout the course.
Audience	Application developers, system administrators, and database administrators who will be creating and implementing a data model in DB2 UDB® for Z/OS or OS/390.
Prerequisites	DB2 UDB® Application Programming and a basic understanding of RDBMS terminology
Objectives	<ul style="list-style-type: none"> • Understand basic data modeling techniques • Understand when to choose a non-meaningful primary key versus a meaningful primary key • Understand the characteristics of foreign keys that impact performance and how to implement them appropriately in a RDBMS • Discuss the concept of referential integrity (RI) and when to consider both RDBMS-enforced RI and application-enforced RI • Understand how RI affects performance • Describe the characteristics of relationships and their impact on performance • Discuss and apply techniques for verifying a data design • Understand how to create DB2 UDB® objects and the impact a physical implementation has on performance • List the different types of denormalization and be able to apply each in the appropriate situation • Create and utilize a design matrix to make denormalization decisions • Understand the basics of distributed data design
Major Topics	<ul style="list-style-type: none"> • Data Modeling • Transition from Logical to Physical Objects • Primary Keys and Indexes • Foreign Keys & Referential Integrity • Relationships & Referential Integrity • Verifying Your Design • DB2 UDB® Tablespaces, Database, & Storage Groups • Physical design considerations • Distributed data design
Format	Lecture/discussion, individual and group workshops
Duration	3 days



Course Contents

1. Introduction

- Application Development Life Cycle
- Development Activity
- Rapid Application Development
- Context Diagram
- Data Modeling Review
- Entity-Relationship Diagramming
- Course Schedule

2. Data Modeling

- Components of a Data Model
- Entities
- Entity Types
- Relationships
- Partnership Sets
- Attributes
- Keys
- Video Store Example Diagram
- Video Store Example - Supporting Documentation
- Normalization
- First Normal Form - Violation
- Second Normal Form - Violation
- Why Third Normal Forms?
- Chapter Summary
- Workshop
- Final Project Tracking Entity List
- Project Tracking Data Model

3. Getting Physical

- Database Managers
- Relationship Database Managers
- Characteristics of Relational Data
- Characteristics of SQL
- Transforming Data Model Objects Into Physical Objects
- Accessing Data in a RDBMS
- How Does a RDBMS Access Our Data
- How Does a RDBMS Access Your Data
- Multiple Table Access
- Modeling Decisions and Their Affect on Performance

4. Primary Keys and Indexes

- Primary Keys
- Candidate Keys
- Other Factors in Choosing a Primary Key

- Updating Primary Keys
- Primary Key (Keylength)
- Index Size Example for DB2 UDB
- Multiple-Column Primary Keys
- Retrieval Considerations
- Cardinality of a Composite Index
- Primary Key Sequencing Characteristics
- Index Structure
- Inserts and Sequential Primary Keys
- Meaningful and Non-meaningful Primary Keys
- Indexes
- Reasons Not to Index a Column
- Other Examples
- Clustering

5. Foreign Keys and Referential Integrity

- Foreign Keys
- Foreign Key Column
- Foreign Key Indexes
- Understanding Foreign Key Index Structures
- Updating the Foreign Key
- Foreign Key Cardinality
- Minimizing Row-ID Chains
- Referential Integrity
- Referential Integrity Terminology
- Referential Integrity Rules
- Insert Rule
- Delete Rule
- Update Rule
- Enforcing Referential Integrity
- Choosing an RI Enforcement Option
- Application Characteristics
- Performance
- RI Performance

6. Relationships and Referential Integrity

- Relationships
- Optional vs. Mandatory Relationships
- Mandatory Relationships
- Optional Relationships
- One-to-One Relationships
- Considerations for Combining One-to-One Relationships
- Many-to-Many Relationships
- Associative Entities



- Associative Entity Primary Keys
- Multiple Parents of Associative Entities
- Recursive Relationships
- Why CASCADE?
- Miscellaneous Referential Integrity Rules
- Super-type and Sub-type
- Creating Physical Tables for Super-type and Sub-type Entities

7. Verifying Your Design

- Why Verify the Database Design
- Database Design Verification
- Normalization
- First Normal Form - Violation
- Second Normal Form - Violation
- Why Third Normal Forms?
- SQL Prototyping
- The Acid Test

8. DB2 Tablespaces, Database and Storage Groups

- Definitions
- Storage Groups
- Databases
- Tablespaces
- Segmented Tablespaces
- Partitioned Tablespaces
- Estimating Space for Tables and Tablespaces
- Altering Tablespaces
- Physical Structures
- Managing VSAM Files

9. Physical Design Considerations

- Database Design Steps
- Managing the Physical Design
- Evaluating Performance
- Solving Performance Problems
- Denormalization

- Denormalization Decision Tree
- Denormalization Types
- Prejoined Tables
- Report Tables
- Mirror Tables
- Split Tables
- Combined Tables
- Redundant Data
- Repeating Groups
- Derivable Data
- Repeating Groups
- Speed Tables
- Using a Usage Matrix

10. Distributed Data Design

- Understanding Distributed Data
- Data Distribution Choices
- Extracted Tables
- Basic Distributed Tables
- Snapshot Tables
- Replication Tables
- Distributed Tables
- Vertical Distribution
- Horizontal Distribution
- Distributed Data Tables
- DB2 UDB Reality

11. DB2 Locking and Performance

- Locking
- SQL Performance Factors
- Access Paths and EXPLAIN
- Catalog Tables Updated by RUNSTATS
- Access Paths and EXPLAIN

Appendices

- A. DB2 UDB "World" Database

