Design Myths

Dejan Sarka

Introduction

- Dejan Sarka (dsarka@solidq.com, dsarka@siol.net, @DejanSarka)
  - 30 years of experience
  - SQL Server MVP, MCT,…
  - 15 books
  - 15+ courses
  - Focus:
    - Data modeling
    - Data mining
    - Data quality

Storage Myth

- Relational database is just a storage system
- WRONG!
- Relational database is the heart of IT systems
  - Relational model is about data integrity
  - A RDBMS enforces data integrity

Physical Myth

- Knowing details of a RDBMS is very important
  - Design for performance, not for theory
- WRONG!
- Logical problems are more important that physical ones
  - Physical problems pass over
  - Relational model is a logical model
  - A good logical design should perform well; the opposite is not always true

Repeating Group Myth

- A table with repeating group of columns needs normalization (1st NF)
  - EMPLOYEES(EmpId, EmpName, Child1, Child2, Child3,...)
- WRONG!
- Repeating columns with similar names is not a repeating group!
  - Just a constraint built in the model
  - ORDERS(OrdId, CustName, OrderDate, DueDate, ShipDate,...)

5

Atomicity Myth

- Every collection in a single column of a single row is a repeating group, is not atomic
- WRONG!
- We need to know the correct definition of atomicity
  - A value is atomic as long as you operate with it only with methods defined on the data type of the value (e.g. Datetime data)
  - XML and UDTs with collections are atomic!

6

Keys Myth

- There are Natural and Surrogate key
- WRONG!
- There are only keys
  - Who defines what is natural, and what is not natural?
  - Slightly better definition: "A key is natural if it is used outside the database and application, in real life" (definition by Fabian Pascal, corrected by Dejan Sarka)

7

Natural Keys Myth

- If you do not have a natural key, you do not have a set
- WRONG!
- Just give me a set, any set!
- A set is purely logical and abstract term
  - The one who defines the set, defines what makes elements unique as well
  - Also, do not forget on abstraction

8

Redundancy Myth

- Normalization is about eliminating redundancy
- Partially wrong
- Normalization is about eliminating redundancy and incompleteness

SQL Language Myth

- SQL language is essential for relational databases
- WRONG!
- The question of using relational model in a database is orthogonal to the question which language is supported by the RDBMS
  - The "D" language proposal (Date, Darwen)
  - Query By Example (QBE)

Truncate Table – DDL or DML?

- Wikipedia: "In SQL, the TRUNCATE TABLE statement is a Data Definition Language (DDL) operation that marks the extents of a table for deallocation (empty for reuse)"
  - So it is DDL?
- WRONG!
- However: "The TRUNCATE TABLE mytable statement is logically (though not physically) equivalent to the DELETE FROM mytable statement (without a WHERE clause)"
  - In a relational model, we are manipulating with data on the logical level; the physical implementation is left to the database management system
- So TRUNCATE TABLE is DML!

11

ALTER TABLE SWITCH – DDL or DML?

- Syntax is DDL
  - SQL Server changes system pages
  - So it is DDL?
- WRONG!
- However, logically, you just move the data from one table or partition to another table or partition
  - You do not change the schema at all
- Therefore, semantically, this is a DML statement

12

Set Operators (1)

- UNION, INTERSECT and EXCEPT operators are commonly called Set Operators
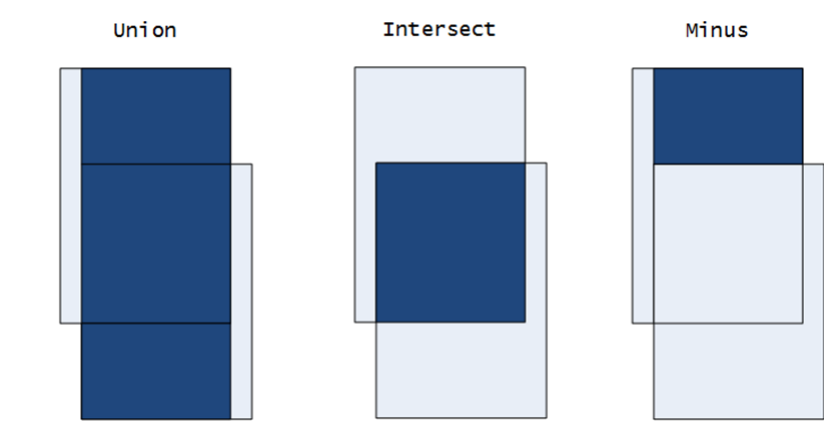 - Represented by Venn diagrams
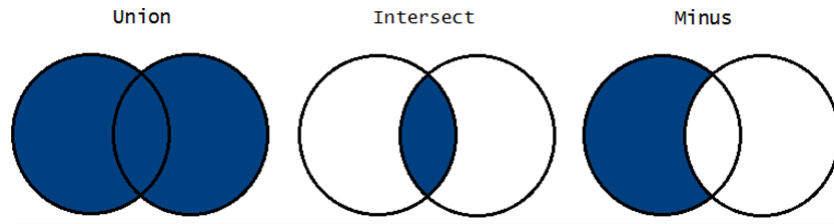
- WRONG!

13

Set Operators (2)

- Set operators work on sets and produce a set
- Relational operators work on relations and produce a relation; SQL operators UNION, INTERSECT and EXCEPT produce relations
- So they are relational operators!

14

Set Operators (3)

- Need to change the presentation!

Denormalization Myth

- Star schema is denormalized
- Not necessary
- Dimensions and fact tables
  - Dimension are denormalized only if you have rigid attribute relationships in natural hierarchies
  - Transaction fact tables are denormalized only if you include degenerate dimensions

Fact Table Myth

- Fact table is called fact table because it holds facts
- Yes, but... but... but...
- Don't all tables hold facts?
  - Fact tables should have different name
  - Central tables, for example

SCD Myth

- There are three solutions for the SCD problem (type 1, type 2, type 3)
- WRONG!
- There are more and better solutions
  - Create another dimension from the type 2 attributes, and leave them in the original dimension as well
  - Scalable solution that encompasses both type 1 and type 2

Class-Table Myth

- A class is the same thing as a table
- WRONG!
- A class is the same thing as a domain (data type)
  - The definitions are the same
  - Relational model is not limited to "simple" types
  - Relational model allows "typed" tables

Table-Class Myth

- A table is the same thing as a class
- WRONG!
- A table is a variable
  - It behaves like a variable (logically, of course)
  - Update means replacing on set with another
  - 1st Zermelo - Fraenkel axiom

Subtype Myth

- A subtype is the same thing as a subclass
- WRONG!
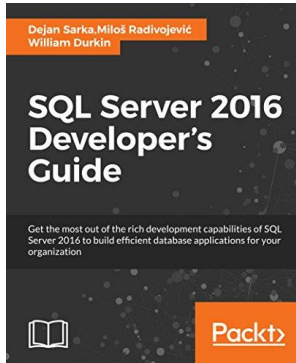- What is inherited by a subclass?
- What is inherited by a subtype?

21

Myth above All Myths

- In this room, there are people who can drink more beer and schnapps than your speaker
- Maybeeee......
- But it has to be proven!

22

Q & A

SQL Server 2016 Developer's Guide

Dejan Sarka, Miloš Radivojević, William Durkin

Get the most out of the rich development capabilities of SQL Server 2016 to build efficient database applications for your organization

Packt>

- Thank you!

- Only for SQL Grillen attendees: 40% discount on e-book!
  - Discount Code: **SQL2016DG40**
  - Start date: June 1st 2017
    Expiry date: June 30th 2017
  - Create a login on the Packt site www.packtpub.com and add the book to the cart
  - Click "View Cart"
  - "Do you have a promo code?" field enter (code provided above)
  - Click the "Apply" button to apply the discount