

# Data Types in SQL-92

by

Joe Celko

copyright 2006

# Joe Celko - Articles

- Member of ANSI X3H2 since 1987
- SQL for Smarties - DBMS Magazine
- Celko on SQL - DBP&D
- SQL Puzzle - Boxes & Arrows
- DBMS/Report - Systems Integration
- WATCOM SQL Column - PBDJ
- Celko on Software - COMPUTING(UK)
- Celko - Intelligent Enterprise
- SELECT FROM Austin - DB/M (Netherlands)

# Joe Celko - Books

- SQL for Smarties - 1995, 1999, 2005
- Instant SQL (Wrox Press)
- SQL Puzzles & Answers – 1997
- Data & Databases - 1999
- Trees & Hierarchies in SQL- 2003
- SQL Programming Style – 2005
- Analytics & OLAP in SQL - 2006

# Definitions

- A distributed database is usually located in several different physical locations, but behaves as one logical unit.
- There are no ISO or ANSI standards for this, so each vendor is on his own.

# History -1

- In the old days (pre-1960's) data was highly centralized on mainframes
- Mini-computers in (1970's) made distribution possible because each branch or group within an organization could afford its own machine.

## History -2

- First distributed mini-systems were file based.
  - Headquarters downloads current price lists overnight
  - Stores uploads daily sales and orders over night
- This used phone lines, modems, pretty simple stuff

## History -3

- Early Operating Systems were not good at tricky file handling.
- These systems tended to move entire files and not just the information needed at each store, so you had high network traffic on your telephone lines
- Not good at verifying the work had been done

## History -4

- “Buddy system” began to appear.
- Headquarters sends out a file
  - Store X fails to get the file for whatever reason
  - Store X asks his buddy Store Y for a copy of the file
- Any individual tuning was done by programs -  
- no databases yet!

# Basic Distribution Models -1

- Subscription = X is on the mailing list for updates and gets them whenever they come out. Push technology.
- By Request = X must request the update from a trusted source. Notice that the information might be assembled on the fly! Pull Technology.

## Basic Distribution Models -2

- A single set of trusted sources - usually one database at headquarters - provides the new data.
- Example: Wal-Mart sells products at one price in every store, so a central price list makes sense.

## Basic Distribution Models -3

- The network has timestamps and validations so that any node can be a data source or destination.
- Example: Auction systems which have to include shipping and local taxes from each node to the destination of the winning bids.

# Problems -1

- Backup of a distributed database is a nightmare.
  - It must come back to its own internally consistent state (like a single database)
  - It must be consistent with every other database with which it is linked.

## Problems -2

- Co-ordination of databases is tricky because it comes in many flavors
- Problems occur in
  - Temporal
  - Logical
  - Legal

## Problems -3

- Temporal problems
- SQL-92 uses only UTC timestamps, but real products do not.
- Lawful time is variable -- Australia in 2000 is postponing DST for the Olympics, the UE is changing rules, etc.
- There are commercial timing services to synchronize your networks

## Problems -4

- Logical Problems accross databases
- One salesman shows up in two cities at once.
- There are special normal forms for horizontal partitions of a table. See Smith and [2-3] NF, based on unions operations

# Questions & Answers

?