Database Technology I
Happy Fifth Birthday to XML

10 February 2003: Celebrate the fifth birthday of the Extensible Markup Language (XML), first published as a W3C Recommendation on 10 February 1998. Visit the XML home page. Read about XML's growth in this article by Dave Hollander and C. M. Sperberg-McQueen, participants in the W3C XML Working Group who wrote the original twenty-five page XML specification. The authors believe, "Just as interchangeable parts drove the Industrial Age, reusable information powers the Information Age." (News archive)

http://www.org/2003/02/xml-at-5.html
DATABASE TECHNOLOGY I

• Object and relational technology

• The database schema

• Accessing databases: SQL, ODBC, and JDBC

• Size, performance, and other issues
Object and relational technology

- **Database design strategies**
  - Database paradigm:
    - RDB - ORDB - ODB
  - Interconnect paradigm:
    - JAVA - CORBA
  - Query design in Client-Server environment

- **Hardware functionality**
  - Hard disks, RAID, & striping
  - JBOD
Database Functionality

- Relational vs Object-Oriented Models
  - Object-oriented development environment
  - Unified model vs. multiple parts

- Features of Current Relational Models
  - SQL query
  - Replication servers
  - Higher-level tools

- Primary Issues in Implementation
  - Heterogeneous delivery systems
  - Heterogeneous storage systems
  - Legacy systems
  - User interface commonality
The Object-Relational DBMS . . .

- Supports Queries
- Supports Complex Data
- Supports Standards
  - SQL
  - Legacy data
  - Client-server
  - Development tools
- Supports Open Tools
  - ODBC
  - Java
  - Internet

Adapted from *Object-Relational DBMSs: The Next Wave* by Michael Stonebreaker, Morgan Kauffman, Publ., San Francisco, 1996.
Simplicity of the object-relational paradigm

Example: archiving DICOM information

Data models are based on tables where only simple data types (integer, string, data/time, etc.) are allowed.

<table>
<thead>
<tr>
<th>Patient Name</th>
<th>DOB</th>
<th>Image</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mary</td>
<td>3/2/56</td>
<td>.....</td>
</tr>
<tr>
<td>Bob</td>
<td>5/6/79</td>
<td>pointer</td>
</tr>
<tr>
<td>Joe</td>
<td>2/2/74</td>
<td>.....</td>
</tr>
</tbody>
</table>

Images must reside outside of the database and are referenced by pointers.
The Object-Relational Paradigm

No Loss of connectivity or relationships.

No change in programming methods between text and images.

**ORBDB**

<table>
<thead>
<tr>
<th>Patient Name</th>
<th>Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mary</td>
<td>.....</td>
</tr>
<tr>
<td>Bob</td>
<td></td>
</tr>
<tr>
<td>Joe</td>
<td>.....</td>
</tr>
</tbody>
</table>

![Smart compound image object]

© cfdewey 2005
Using an ORDB

- Must build a definition within database
- Provides database connection and session management
- Direct access using SQL-2
- Functions to disassemble and reconstruct complete objects
- Functions to analyze, manipulate images and other objects
- Basis functions for developing new opaque data types and methods
- Handles most communications with database applications

DICOM ORDB example
**SQL: Structured Query Language**

- Relatively Simple Set of Commands
- Example:
  
  \[
  \text{Select NAME from EMPLOYEE} \\
  \text{where SALARY > 60000}
  \]

- Many Third-Party Tools For Writing Applications
  - ODBC, JDBC
- Oracle and IBM DB2 have JDBC and “blade” technology
Extended database with replication

- Server
  - Primary data
  - Networked replication engines
- Client
- Local server
  - Replicated data
- Client
Replication Servers

- Maintain copies of data from central database

- Synchronous
  - Updates to all copies on each transaction
  - Two-phase commit
  - All or nothing

- Asynchronous
  - Source first, target later
  - Log transfer manager
  - Seconds to minutes
  - Avoids hangups with single sites
Caching

- Maintain local copies of central database
  - Requires synchronous replication on server updates
  - Much of the updated information never used locally

- Synchronous
  - Updates cached copy on each transaction
  - All or nothing
  - Guaranteed fresh
  - Can set local rules for retention
  - Well known browser technology

- Akami
  - Smart caching
  - Dynamic allocation through DNS