



Engenharia de Dados e Conhecimento

XML DATABASES



- Why do I need to store XML in a database?
- Because a lot of information exchanged by systems is in XML format.
So, it's very reasonable to save that information as it is.
- Advantages
 - It spares time by no need of processing and transforming data
 - Support of backups, transactions, etc.
- Disadvantages
 - Format of store (more size)
 - (In)compatibilities with traditional RDBMSs



- How to store XML structured information in databases?
- Four distinct ways:
 - Relational DB, storing the “text/xml” data in a table field as a string
 - Relational DB with internal schema parsing
 - Native XML Database (NXD)
 - Files System



- Allows to use functionalities from traditional RDBMs
- It hasn't efficient mechanisms to search and update the XML data
- It's a "forcing" solution for DBs
- Examples
 - MySQL



- Each DBMS process the information from the XML Schema and creates an internal representation of the XML structure
- If there isn't any schema, the representation is created according to the XML documents



- Advantages
 - Can consume XML data
 - Allows XQuery queries, resulting in sets of nodes and not in sets of registers
 - Validates the operations, in concordance with the XML Schema (if it exists)
 - Disadvantages
 - Architecture dependent
 - Requires conversion from XQuery to SQL or internal mechanisms
- htz@ua.pt • It's an adaptation from the RDBM model



- Examples of RDBs
 - SQL Server 2005 +
 - Oracle 9i +
 - PostgreSQL 8+
 - ...



- Advantages
 - Input, output and stored formats are all the same
 - Supports transactions
 - Validation according with XML Schema
 - Very good performance
- Disadvantages
 - Systems created from ground
 - They haven't the trust of the well established RDBMs



- Examples of NXDs
 - BaseX
 - eXist
 - Sedna
 - Berkeley DB (Oracle)
 - EMC xDB
 - MarkLogic Server
 - ...



- Advantages
 - Input, output and stored formats are all the same
 - Very fast access to information
- Disadvantages
 - Doesn't support transactions, concurrency, ...
 - Need of external tools to search and update
 - Doesn't support validation



	NXD	XEnDB	Files System
Transactions	Yes	Yes	No
XML Insert	Direct insertion or light conversion to internal structure	Conversion to relational structure	Direct insertion
XML Retrieve	Direct retrieve	Conversion from relational to XML	Direct retrieve
XML Update	XQuery/XUpdate	SQL or Converted XQuery/XUpdate	File replace
XML Queries / Indexes	Yes	Yes	External tool
XML Document in-out	Generally YES	No / Can be	Always