

XML Databases And Effective search Engine

Review Report

Submitted by:

R.Srinivasan

200272129

*MSC Integrated
Computer Science*

Modules

1. Converting Database table to XML

Three approaches:

- 1. Using Map file*
- 2. Using the String class*
- 3. Using DOM Tree*

2. Parsing XML and Inserting it into Database table

3. Text file to XML

4. HTML to XML using Tidy Parser

5. Constructing DOM tree (Generalized for any XML file) Using JTree of Java swing

6. Search Engine

- 1. Using KMP Algorithm (Converting XML
To String and Searching for the Pattern)*
- 2. Specific Tag Based Search (Using DOMTree)*

Introduction

- *XML was originally designed to exchange data over the World Wide Web*
- *Currently, it has become a public and widely accepted standard*
- *It is Used for exchanging data between any number of computer systems*

Storing XML

- *File system*
 - Text files - differently represented during processing*
 - Text files - with additional indexes*
- *Database systems*
- *Relational*
- *Object-oriented, etc.*

File System

Advantages

Easy to store

Document granularity is maintained

Disadvantages

Query processing

Database System

Advantages

Well understood technology

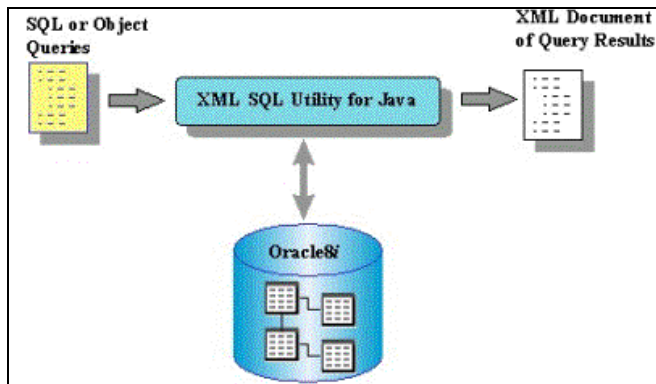
Optimized for query processing

Disadvantages

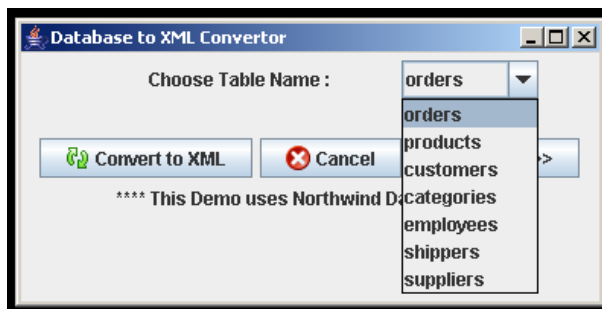
XML does not directly map to any data model

Not amenable to a rigid schema

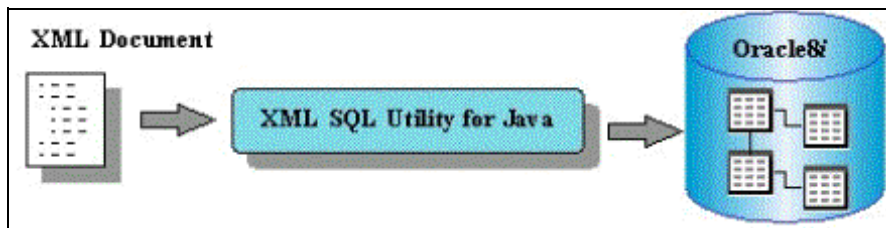
Database table to XML (Module 1)



Database to XML



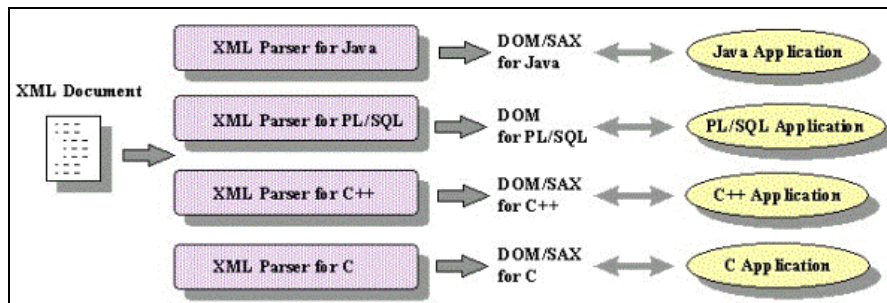
XML to Database (Module 2)



Java XML Parsers

1. SAX
2. DOM

XML Parsers



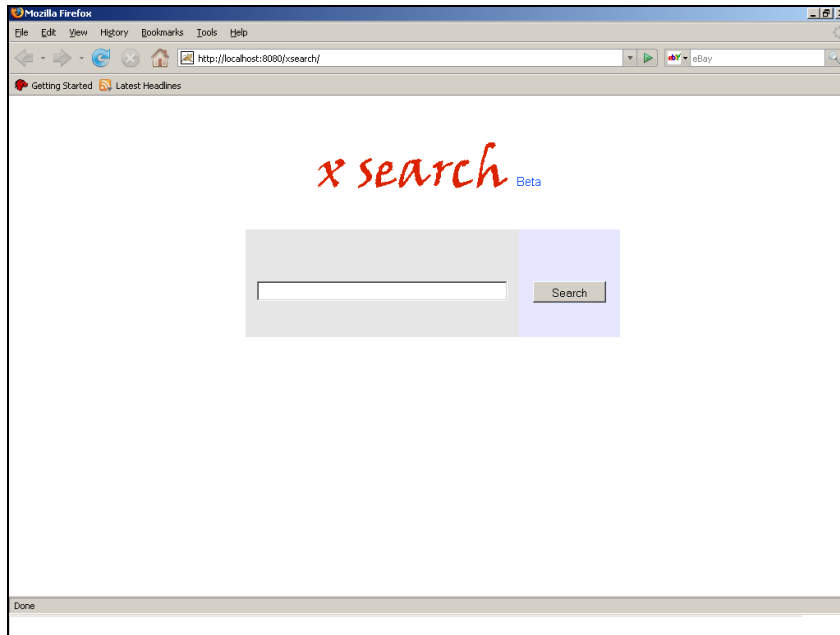
X Search

- *Java Servlets*
- *KMP String matching algorithm*
- *DOM Tree Searching*
- *TOMCAT*

Future Search Engines

- *More and More Search Engines*
- *Location sensitive Searches*
- *Intelligent Search Engines*
- *More Accurate Results*
- *New Areas of Search*
 1. *Draw a diagram and search for similar pictures in the Web*
 2. *Hum a song find Matches in the web*

X Search



Conclusions

- *A more Sophisticated XML Database API is written using JAVA*
- *An XML Search Engine is built*
- *Native XML Databases can perform better than relational DB (only for XML Data)*
- *Searching XML can give more accurate results*