

CHƯƠNG 9

Giới thiệu các vấn đề liên quan đến CSDL

Slide bài giảng môn Cơ sở dữ liệu

© Bộ môn Hệ Thống Thông Tin - Khoa Công Nghệ Thông Tin - Trường Đại học Khoa học Tự nhiên

Nội dung

- Cơ sở dữ liệu hướng đối tượng
- Kho dữ liệu
- Cơ sở dữ liệu và Internet
- Cơ sở dữ liệu XML

© Bộ môn HTTT - Khoa CNTT - Trường ĐH KHTN

2

Cơ sở dữ liệu hướng đối tượng

Slide bài giảng môn Cơ sở dữ liệu

© Bộ môn Hệ Thống Thông Tin - Khoa Công Nghệ Thông Tin - Trường Đại học Khoa học Tự nhiên 3

Các ứng dụng cơ sở dữ liệu nâng cao

- Computer-Aided Design (CAD)
- Computer-Aided Manufacturing (CAM)
- Computer-Aided Software Engineering (CASE)
- Network Management Systems
- Office Information Systems (OIS) and Multimedia Systems
- Digital Publishing
- Geographic Information Systems (GIS)
- Interactive and Dynamic Web sites
- Other applications with complex and interrelated objects and procedural data.

© Bộ môn HTTT - Khoa CNTT - Trường ĐH KHTN

4

Sự bất cập của cơ sở dữ liệu quan hệ



- Tính biểu diễn thực thể của thế giới thực kém
- Các thao tác trên quan hệ ít và khó mở rộng
- RDBMS không hỗ trợ việc truy xuất theo kiểu duyệt (navigation)

Các khái niệm của hướng đối tượng



- Trừu tượng (Abstraction), Bao bọc (encapsulation), Che dấu thông tin (information hiding).
- Đối tượng (Object) và thuộc tính (attributes)
- Định danh đối tượng (Object identity).
- Phương thức (Methods) và thông điệp (messages).
- Lớp (Classes), lớp con (subclasses), lớp cha (superclasses), thừa kế (inheritance).
- Overloading.
- Đa hình (Polymorphism) và liên kết động (dynamic binding).

Đối tượng (Object) - Lớp (Class)

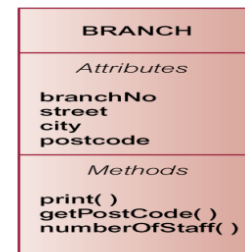


- **Đối tượng:**
 - Là một đối tượng thực tế hay trừu tượng được xác định bằng: một tập các tính chất (attribute) và một tập các ứng xử (behavior)
 - Ví dụ : Thầy giáo A là một đối tượng có các tính chất như : tên, tuổi, học vị... và các các ứng xử như: giảng bài, chấm bài, soạn bài...
- **Lớp**
 - Là tập các đối tượng tương tự nhau (cùng tập tính chất và tập ứng xử)
 - Một đối tượng là một xuất hiện (instance) của một lớp

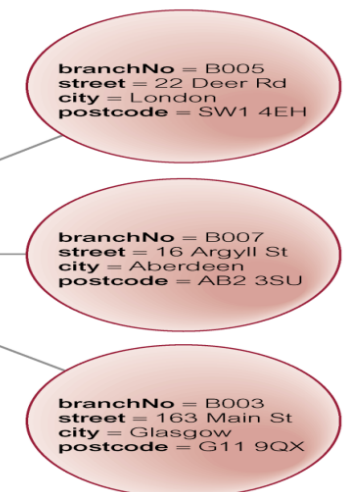
Ví dụ



CLASS DEFINITION

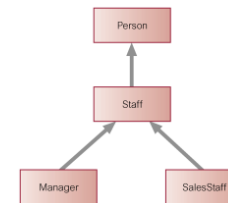


CLASS INSTANCES

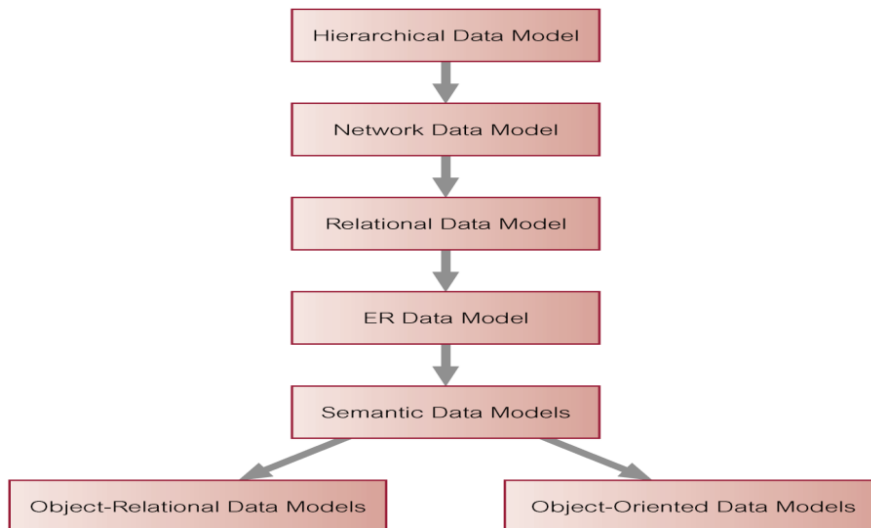
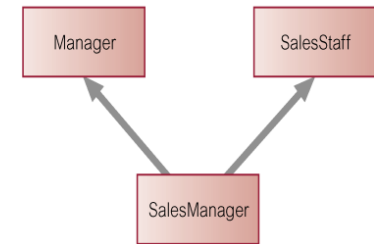


- Phương thức (Method) xác định một ứng xử của đối tượng – thường là một tập nhóm các chức năng được bao bọc/đóng gói
- Thông điệp (message) là một yêu cầu từ một đối tượng đến một đối tượng khác yêu cầu thực hiện một phương thức của đối tượng đó

Đơn thừa kế



Đa thừa kế



- **Object-Oriented Data Model (OODM)**
 - Mô hình dữ liệu hỗ trợ ngữ ngữ của đối tượng hỗ trợ cho lập trình hướng đối tượng.
- **Object-Oriented Database (OODB)**
 - Lưu trữ và chia sẻ tập các đối tượng được định nghĩa bằng ODM.
- **Object-Oriented DBMS (OODBMS)**
 - Hệ quản trị CSDL hướng đối tượng.

- Mở rộng hệ cơ sở dữ liệu quan hệ để hỗ trợ các đặc trưng của đối tượng đối với các thuộc tính/yếu tố dữ liệu đặc biệt
- Mở rộng ngôn ngữ SQL3 để hỗ trợ truy vấn đối tượng

- RDBMSs hiện chiếm lĩnh thị trường công nghệ cơ sở dữ liệu ước lượng khoảng 15 – 20 tỉ dollars một năm (50 tỉ thêm các công cụ bán kèm) và tăng trưởng 25% năm.
- Thị trường OODBMS còn nhỏ với khoảng 150 triệu dollars năm 1996 và chỉ chiếm 3% thị trường vào năm 1997
- Một vài chuyên gia về thị trường OODBMS cho rằng nó sẽ tăng trưởng trên 50% một năm nhưng cũng khó vượt được RDBMS.

Kho dữ liệu (Data Warehouse)

- Kho dữ liệu là một cơ sở dữ liệu hỗ trợ ra quyết định độc lập với cơ sở dữ liệu nghiệp vụ của doanh nghiệp
- Tính chất của kho dữ liệu
 - Hướng chủ thể (subject-oriented)
 - Tích hợp (Integrated)
 - Theo thời gian (time-variants)
- Đây là tập dữ liệu hỗ trợ cho doanh nghiệp ra quyết định

Dữ liệu hướng chủ thể



- Dữ liệu trong kho dữ liệu được tổ chức theo các chủ thể chính của doanh nghiệp (ví dụ: khách hàng, sản phẩm, bán hàng...) hơn là theo các lĩnh vực ứng dụng chính (ví dụ: quản lý tồn kho, lập hóa đơn...)
- Dữ liệu cần thiết cho việc ra quyết định chứ không phải dữ liệu cho việc cho các nghiệp vụ

Dữ liệu tích hợp



- Dữ liệu trong Kho dữ liệu được tích hợp từ nhiều nguồn dữ liệu tác nghiệp của doanh nghiệp
- Sự tích hợp phải bảo đảm tính nhất quán và thống nhất

Dữ liệu theo thời gian



- Dữ liệu trong kho dữ liệu chỉ chính xác trong một khoảng thời gian nào đó
- Dữ liệu liên kết với thời gian. Các sự kiện rút trích được từ dữ liệu theo thời gian



Cơ sở dữ liệu và công nghệ Internet

Web và Internet

- Là một cơ sở dữ liệu cực lớn/ hệ thống tập tin cực lớn
- Độc lập với nền tảng
- Trao đổi thông qua các nghi thức xác định trước

URL

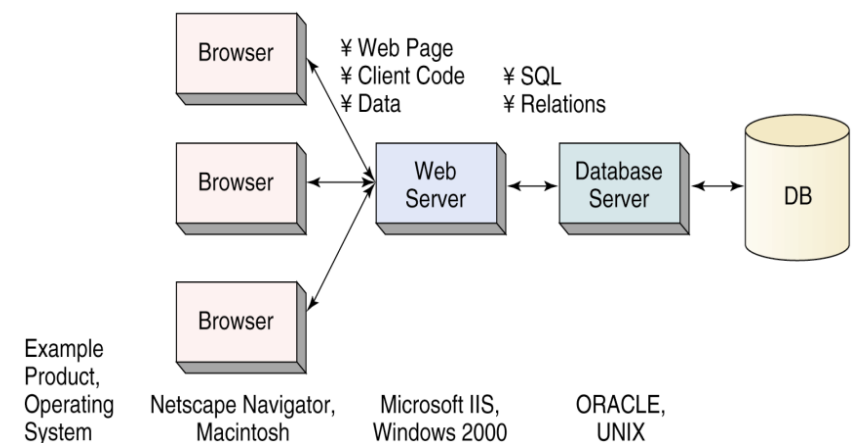
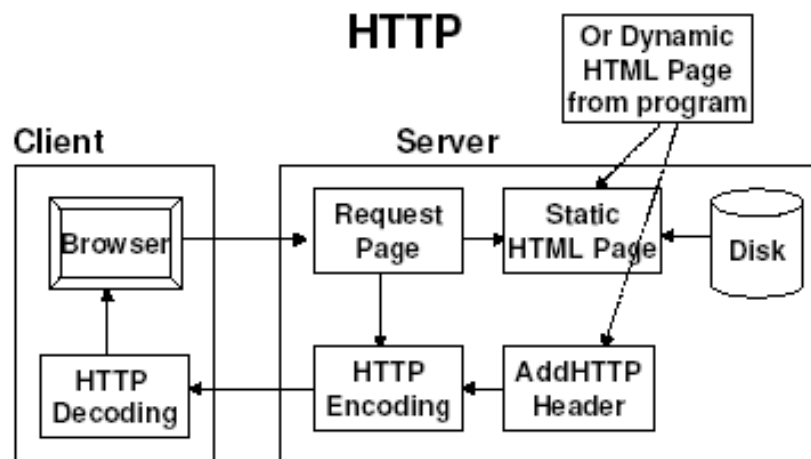
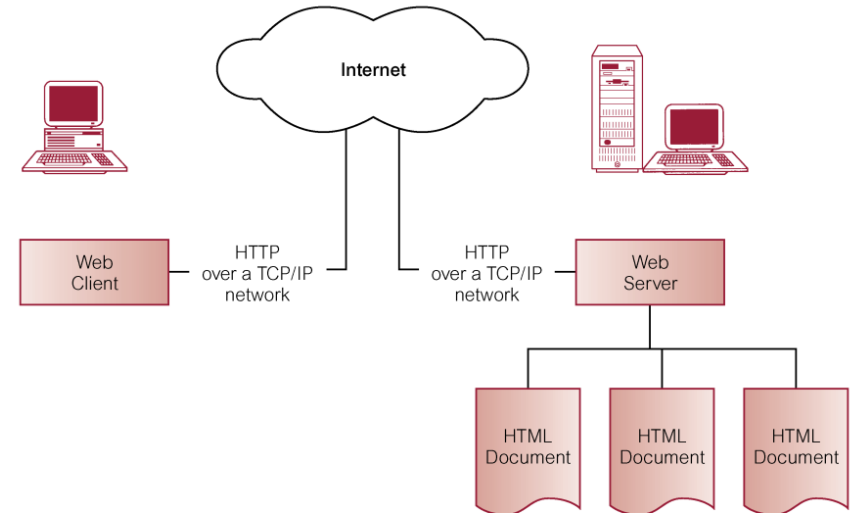
- Là địa chỉ duy nhất để xác định một tài nguyên (resource) trên Internet

Ngôn ngữ đánh dấu (Markup Language)

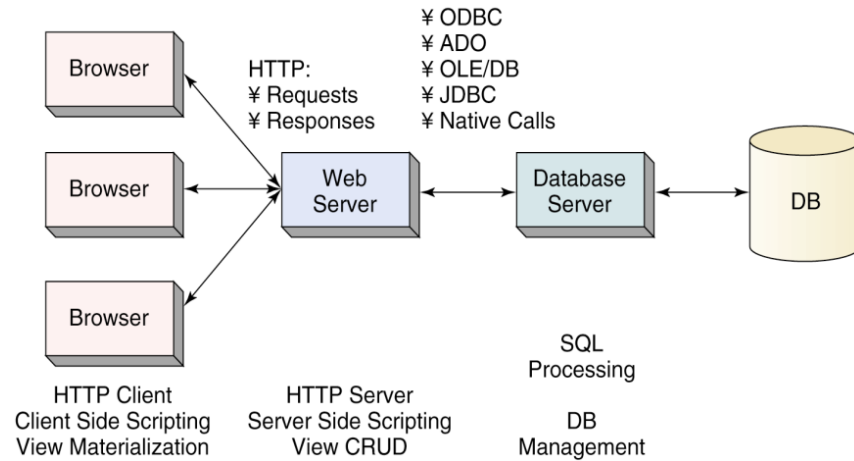
- HTML
- XML

Nghi thức

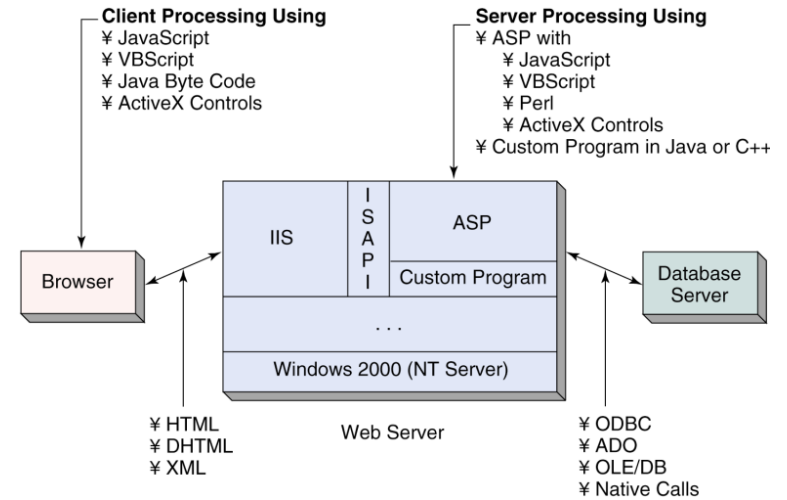
- HTTP



Chức năng của các lớp



Các chuẩn và ngôn ngữ trên Web Server

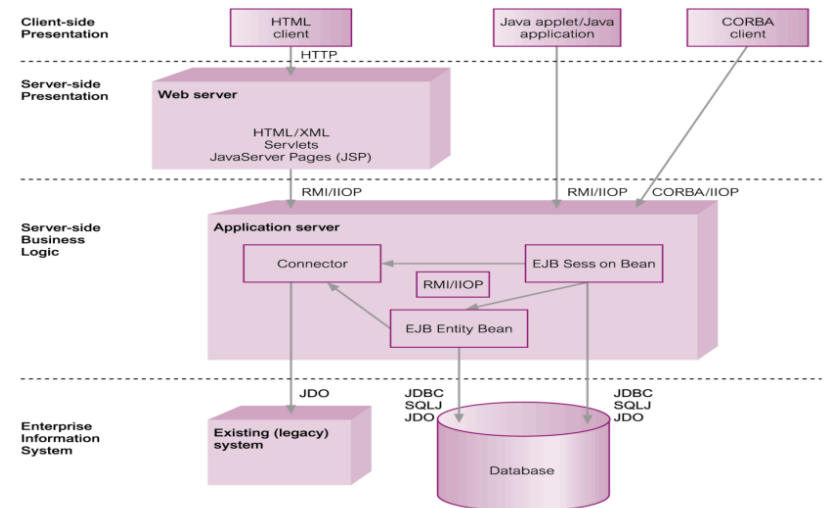


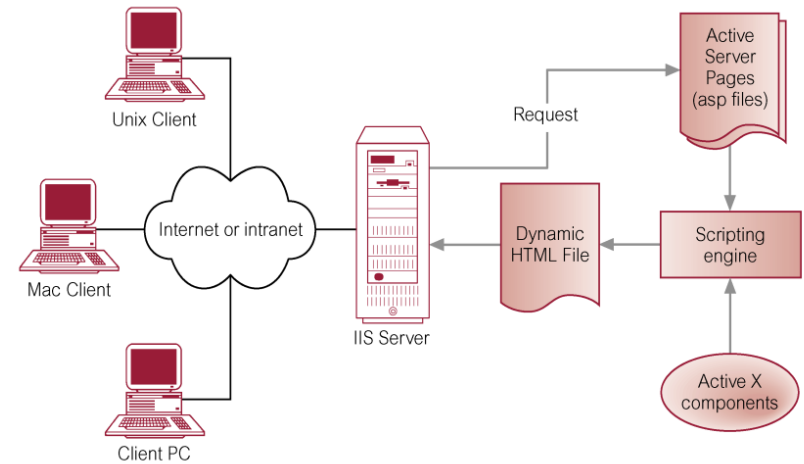
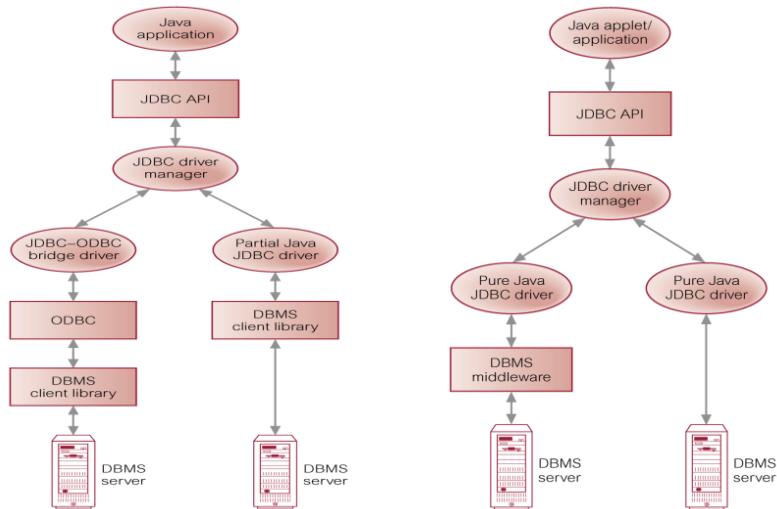
Một số công nghệ kết nối Internet và DB



- Java 2 Platform
- JDBC
- ASP
- .NET Framework

Java 2 Platform





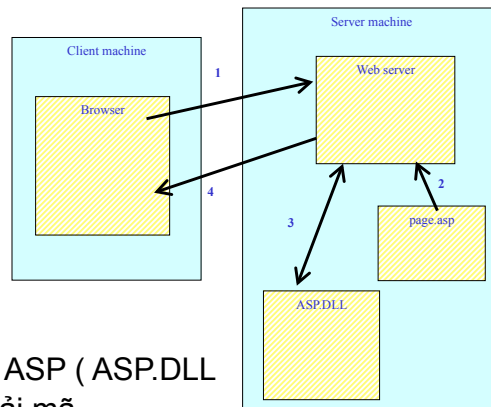
ASPs hoạt động thế nào

1. Trình duyệt yêu cầu một trang ASP từ server.
click vào liên kết (link).
Một form được gửi đến server.

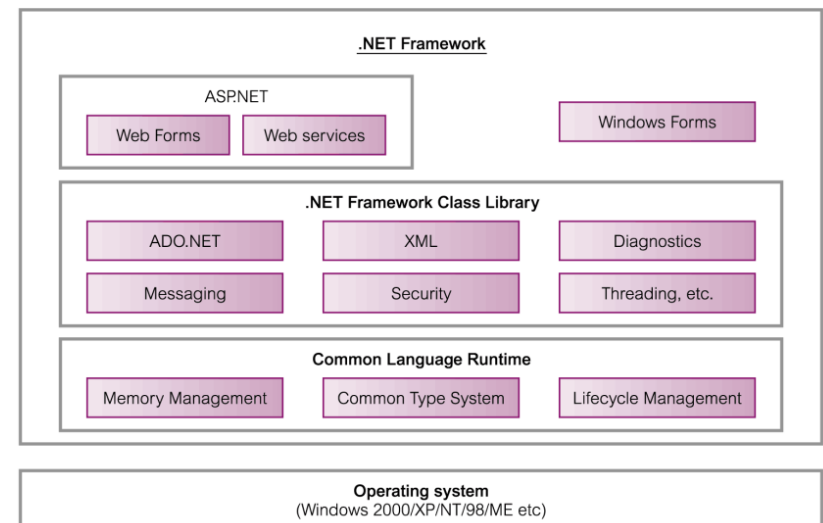
2. server lấy trang yêu cầu

3. Server sử dụng bộ xử lý ASP (ASP.DLL trên Windows) để phân giải mã.

4. Trang HTML kết quả sẽ được gửi trả về client.



.NET Framework



XML

Slide bài giảng môn Cơ sở dữ liệu

© Bộ môn Hệ Thống Thông Tin - Khoa Công Nghệ Thông Tin - Trường Đại học Khoa học Tự nhiên 33

Lịch sử phát triển XML

- XML được tạo vào năm để vượt qua các giới hạn của SGML và HTML
- SGML thì lớn và phức tạp
- HTML tập trung vào cấu trúc, ngữ nghĩa và siêu dữ liệu định dạng nó không luôn được “well-formed”
- XML không thay thế HTML trong các trình duyệt web, nhưng được dùng cho các lĩnh vực khác đặc biệt là trong trao đổi dữ liệu (data interchange)

Mô hình tài liệu XML

- XML cung cấp các mở rộng (extensions) để tuuwong tác với siêu văn bản và nhiều ngôn ngữ khác
- Hai khái niệm cấu trúc chính của XML:
 - **Thành phần (Elements)** :Complex & Simple)
 - **Thuộc tính (Attributes)**
- Mô hình dữ liệu phân cấp XML
 - Các node bên trong biểu diễn cho các thành phần phức, các node là biểu diễn cho các thành phần đơn.
- Tài liệu XML được gọi là bán cấu trúc (semi-structured)
- Các API cho mô hình tài liệu XML hỗ trợ tìm kiếm, thêm, xóa và sửa

Các loại tài liệu XML

- Có 3 loại tài liệu XML chính:
 1. Tài liệu XML về dữ liệu (Data-centric) :
Có nhiều phần tử dữ liệu nhỏ có cùng cấu trúc
 2. Tài liệu XML về văn bản (Document-centric):
Các tài liệu với số lượng lớn các văn bản có tựa và không có thành phần dữ liệu có cấu trúc
 3. Tài liệu XML hỗn hợp (Hybrid):
Có phần chứa dữ liệu có cấu trúc có phần chứa văn bản phi cấu trúc

Có hai loại XML

- Well-Formed XML

- Bắt đầu bằng phần khai báo XML
- Tiếp theo là chú thích cú pháp về mô hình cây
- Tiêu chí cho tài liệu XML well-formed

- Valid XML

- Một tiêu chí để kiểm chứng (Valid) một tài liệu XML là
 - well-formed
 - Theo sau một đặc tả cấu trúc trong một tập tin DTD hoặc tập tin lược đồ XML (XML Schema).

37

- Được xây dựng bằng ngôn ngữ định nghĩa DTD (DTD definition language)
- Là một phần của chuẩn XML
- Can mirror the structure of a relation and format Queries output to generate a relation
- Các đối tượng văn bản (text object) phải well-formed và valid

38

```
<?xml version= "1.0" encoding= "UTF-8" standalone= "yes"?>
<?xml:stylesheet type = "text/xsl" href = "staff_list.xsl"?>
<!DOCTYPE STAFFLIST SYSTEM "staff_list.dtd">
<STAFFLIST>
  <STAFF branchNo = "B005">
    <STAFFNO>SL21</STAFFNO>
    <NAME>
      <FNAME>John</FNAME><LNAME>White</LNAME>
    </NAME>
    <POSITION>Manager</POSITION>
    <DOB>1-Oct-45</DOB>
    <SALARY>30000</SALARY>
  </STAFF>
  <STAFF branchNo = "B003">
    <STAFFNO>SG37</STAFFNO>
    <NAME>
      <FNAME>Ann</FNAME><LNAME>Beech</LNAME>
    </NAME>
    <POSITION>Assistant</POSITION>
    <SALARY>12000</SALARY>
  </STAFF>
</STAFFLIST>
```

Một tập tin XML DTD mô tả cấu trúc của một dự án

```
<!DOCTYPE projects [
  <!ELEMENT projects (project+)>
  <!ELEMENT project (Name, Number, Location, DeptNo?, Workers)>
  <!ELEMENT Name (#PCDATA)>
  <!ELEMENT Number (#PCDATA)>
  <!ELEMENT Location (#PCDATA)>
  <!ELEMENT DeptNo (#PCDATA)>
  <!ELEMENT Workers (Worker*)>
  <!ELEMENT Worker (SSN, LastName?, FirstName?, hours)>
  <!ELEMENT SSN (#PCDATA)>
  <!ELEMENT LastName (#PCDATA)>
  <!ELEMENT FirstName (#PCDATA)>
  <!ELEMENT hours (#PCDATA)>
] >
```

- Lọc đồ XML có thể được phân tích bởi bộ phân giải XML (XML parser)
- Thao tác dữ liệu XML
 - XQuery dựa trên Xpath
 - XQuery hỗ trợ “select, project, and join”
 - XUpdate hỗ trợ cập nhật (updates)

