# Chương 1: Tổng quan về công nghệ

## Giới thiệu về công nghệ Virtuoso.

Virtuoso là 1 công nghệ lưu trữ dữ liệu tiên tiến, với ưu điểm là có tốc độ vượt trội. Virtuoso là một giải pháp hiện đại cho việc truy cập dữ liệu, ảo hóa, hội nhập và đa mô hình quản lý cơ sở dữ liệu quan hệ (bảng SQL hoặc đồ thị RDF).

* **Nhanh nhẹn thông qua việc khai thác trí tuệ nhân tạo**.
* **Web Application Deployment**: có hỗ trợ ảo hóa của các hành động khác nhau trên một loạt các API liên quan đến Web Services.
* **Bảo mật dữ liệu & An ninh**: qua sử dụng kiểm soát truy cập thuộc tính dựa trên hạt mịn để cung cấp truy cập dữ liệu vô song và an ninh thao tác thông qua sử dụng các tiêu chuẩn mở hiện có như TLS và HTTPS
* **Tối đa hóa hệ thống:** Đảm bảo kết nối tiêu chuẩn hiện có như: (URI, URL, HTTP, SPARQL, SQL, ODBC, JDBC, ADO.NET, OLEDB, XMLA, XQuery, XPath, XSLT).

## Một số thuật ngữ

Một số thuật ngữ cần được làm rõ trước khi cài đặt:

### RDF (The Resource Description Framework)

RDF là được hiểu là 1 tiêu chuẩn được cung cấp bởi tổ chức W3C (www.w3c.org).

Nó cung cấp một phương tiện đơn giản để mô tả các đối tượng tài nguyên đa diện và để liên kết chúng thành các biểu đồ quan hệ phức tạp.

### IRI

* IRI là 1 phần cốt lõi của RDF, nhờ có IRI mà RDF trở nên vô cũng mạnh mẽ.
* IRI có 2 phần cơ bản, prefix và local. Một IRI có dạng như sau:

<http://abc.com/DEMO>

* Trong đó <http://abc.com/> là prefix, đây là 1 đường dẫn bắt kì, nhưng phải bắt đầu bằng http://
* DEMO là phần local.

### Literal

Literal cũng là 1 dạng RDF, nó thường được dùng để lưu trữ các dữ liệu có dạng String là chủ yếu, Literal chỉ có 1 thành phần, chính là thành phần mang ý nghĩa của literal.

### Tripple

Tripple là 1 dạng lưu trữ phổ biến trên các server ở dạng RDF. Tripple bao gồm 3 trường:

* Subject: Trường đầu tiên chứa URI xác định duy nhất tài nguyên mà bộ ba này mô tả.
* Predicate: Trường thứ hai chứa URI xác định một thuộc tính của tài nguyên này, chẳng hạn như màu sắc hoặc kích thước của nó hoặc mối quan hệ giữa tài nguyên này với tài nguyên khác, chẳng hạn như nguồn gốc hoặc quyền sở hữu.
* Object: Trường thứ ba là giá trị. Nó có thể là một giá trị theo nghĩa đen, chẳng hạn như "màu đỏ" hoặc URI của tài nguyên được liên kết.

### Resource description

Được định nghĩa là một tập hợp các triples mà tất cả đều có cùng một URI trong trường chủ đề. Nói cách khác, triples đều mô tả các thuộc tính của cùng một thứ.

### Statement

Là một đối tượng Java phía máy khách mô tả triple.

## Hướng dẫn cài đặt Virtuoso (phiên bản mới nhất – 8.2)

### Hướng dẫn cài đặt trên linux – linux mint 19.1

• Tạo thư mục virtuoso chứa các gói cài đặt với lệnh sau:

$ sudo mkdir virtuoso

• Lên trang chủ tải Virtuoso tại đường dẫn: *https://shop.openlinksw.com/license\_generator/virtuoso-download/*

(Chọn bản 8.2 giành cho Linux)

• Khi đó ta được file lovpz2zz.tar

• Copy file mới tải về vào thư mục “virtuoso” vừa tạo bên trên với câu lệnh:

$ sudo cp lovpz2zz.tar /virtuoso

• Sau đó giải nén file mới copy vào thư mục “virtuoso” với câu lệnh sau:

$ sudo tar -xvf lovpz2zz.tar

• Sau khi giải nén xong ta sẽ được 2 file đó là install.sh và universal-server.taz

• File install.sh là file để ta cài đăt Virtuoso, ta thay quyền cho nó bằng lệnh:

$ sudo chmod +x install.sh

• Sau đó ta chạy file install.sh (nó sẽ tạo ra rất nhiều file) bằng lệnh sau:

$ ./install.sh

• Sau câu lệnh đó thì ta đã cài đặt hoàn thành, nó sẽ tự truy nhập trên Ứng dụng truy cập Web mặc định để vào thay đổi tài khoản và mật khẩu (ta tắt nó đi, coi như mặc định tên tài khoản là dba và mật khẩu cũng là dba)

• Để chạy server ta vào thư mục /bin chạy lệnh sau:

$ sudo ./virtuoso-start.sh (tương tự lệnh sudo ./virtuoso-stop.sh để dừng server)

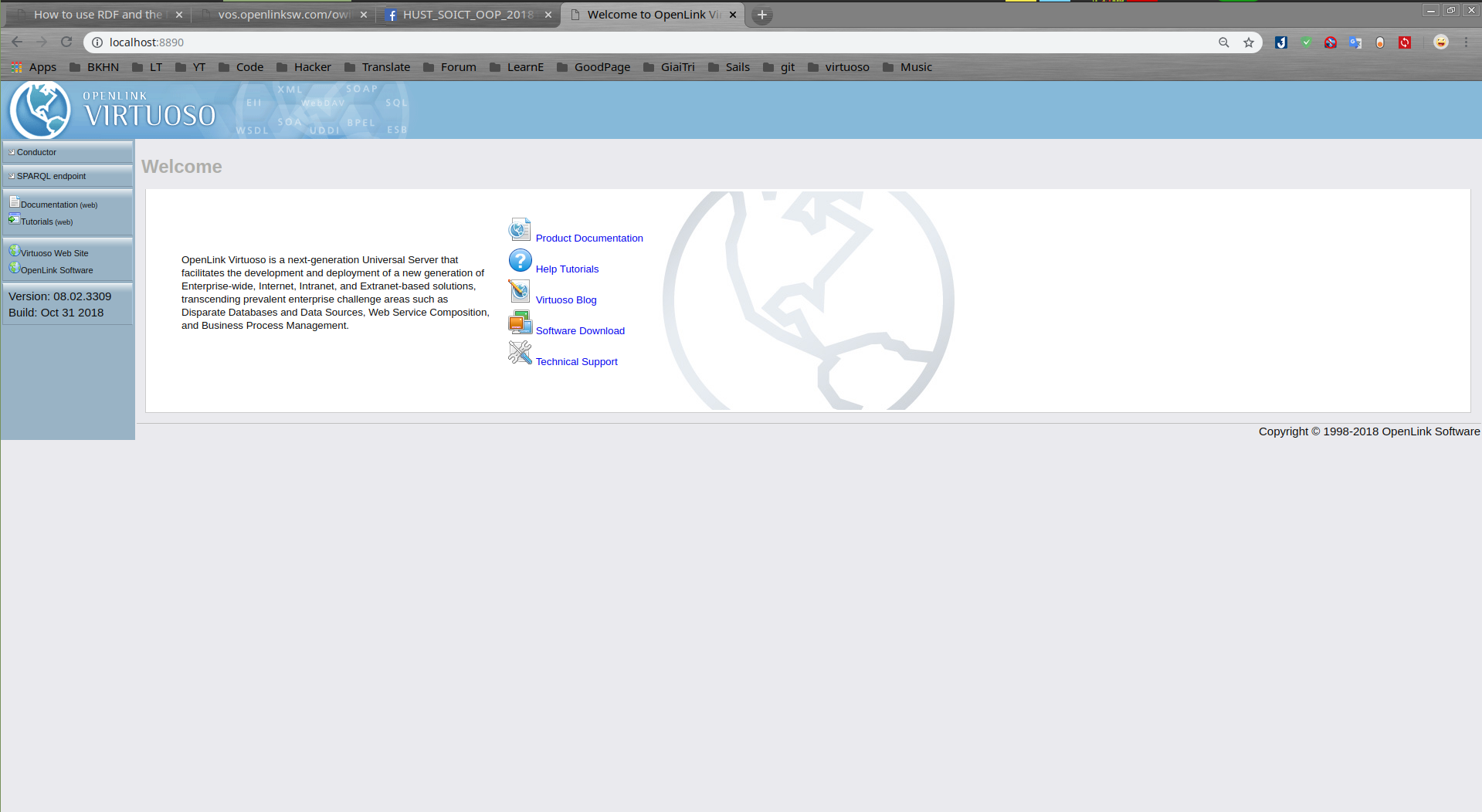
• Để kiểm tra Server có hoạt động hay không chúng ta vào trang http://localhost:8890 để kiểm tra, nhập tên với mật khẩu dba đã mặc định ở trên tạo ở mục trên.

• Để gỡ Virtuoso, ta chỉ cần xóa thư mục “virtuoso” với lệnh sau:

$ sudo rm -rf virtuoso

### Hướng dẫn cài đặt trên windown 10 version 1809

* Bước 1: Truy cập link https://shop.openlinksw.com/license\_generator/virtuoso-download/ .
* Bước 2: chọn phiên bản và hệ điều hành tương ứng, sau đó bấm download và tiến hành cài đặt.
* Bước 3: trên trình duyệt web gõ localhost:8890/ nếu hiển thị như bên dưới thì thành công.



* Do nhóm em hầu hết mọi thành viên đều sử dụng linux, nên việc cài trên windown không được thực hiện bởi thành viên nào cả. Việc cài trên chỉ là demo trên máy ảo.

## Các kĩ thuật sử dụng

### Cách import thư viện bằng Maven

Vào eclipse, chọn File menu, chọn New -> New Project… . Sau đó chọn Maven Project. Chọn tùy chọn “Create a simple project” và nhấn Next.

Trong màn hình tiếp theo bạn phải xác định thêm chi tiết về Maven project như group id-, artifact id-, name- với các thông số dưới đây bạn có thể điền tùy ý:

* group id (ví dụ: org.example)
* artifact id (ví dụ: rdf4j-getting-started)
* name (ví dụ: HelloRDF4J)

Sau đó nhấn Finish để hoàn tất.

### Cách import thư viện bằng file jar

Trước tiên, truy cập vào trang: [*http://vos.openlinksw.com/owiki/wiki/VOS/VirtSesame2Provider*](http://vos.openlinksw.com/owiki/wiki/VOS/VirtSesame2Provider)

Tìm và tải 2 file jar: virt\_rdf4j.jar và virtjdbc4\_2.jar

Đưa 2 file jar trên vào thư mục lib của project.

Chọn chuột phải vào project, chọn Build path -> Configure build path.

* Chọn Libraries và chọn Add JARs…
* Sau đó thêm thư viện bạn muốn chọn (2 file tải trên).
* Sau khi thêm bạn cần update lại project: Chuột phải vào project của bạn chọn Maven >Update project.

## Cách truy vấn

* Virtuoso lưu trữ dữ liệu dưới dạng các “triples”, lưu trữ theo định dạng RDF với 3 trường subject-predicate-object. Để truy xuất các “RDF triples”, ta sử dụng truy vấn SparQL.

Ví dụ về 1 câu lệnh SparQL: “SELECT ?s ?p ?o WHERE {?s ?p ?o. }”

* Mệnh đề SELECT trả về biến ?s, ?p, ?o. Các biến được liên kết với các giá trị subject, predicate, object của 1 “triple” thỏa mãn trong mệnh đề WHERE. Dấu chấm (.) biểu thị sự kết thúc cho ràng buộc.

Ví dụ 1 câu lệnh sử dụng SPARQL filter match:

* Obama đã đến thăm những nước nào?:

String query = "PREFIX label:<http://www.example.org/Virtuoso# /> \n";

* + - += "PREFIX re:<http://www.example.org/relationship/> \n";
    - += "SELECT ?s ?p ?o \n";
    - += "WHERE { \n";
    - += " ?person label:label ?s .\n";
    - += " ?person ?p ?country .\n";
    - += " ?country label:label ?o .\n";
    - += " FILTER regex(?s, \" Obama\") ";
    - += " FILTER (?p =re:visit)";
    - += "}";
* Phương thức: prepareTupleQuery(QueryLanguage.SPARQL, queryString) của đối tượng RepositoryConnection với tham số đầu vào là ngôn ngữ sử dụng (ở đây là SPARQL) và câu lệnh. Phương thức này trả về 1 đối tượng TupleQuery.
* Sau đó sử dụng phương thức evaluate() của TupleQuery để trả về 1 tập kết quả tương tự như ResultSet trong JDBC, nó trả về 1 TupleQueryResult mà ta có thể truy xuất kết quả với phương thức getValue().

VD:

TupleQuery tupleQuery = connection.prepareTupleQuery(QueryLanguage.SPARQL, query);

TupleQueryResult result = tupleQuery.evaluate();

while (result.hasNext()) {

BindingSet bind = result.next();  
Value s = bind.getValue("s");  
Value p = bind.getValue("p");

Value o = bind.getValue("o");

System.out.prindln(“%s %s %s”, s, p, o);

}

## Cách giao tiếp bằng API

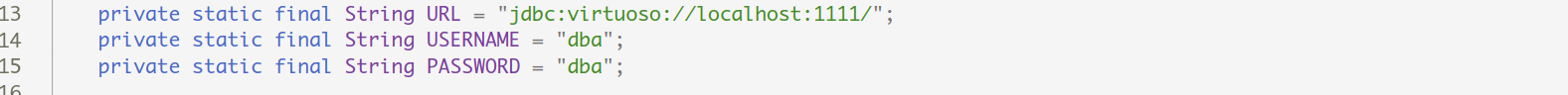
Giao tiếp API của Eclipse, link hướng dẫn cụ thể: [*http://vos.openlinksw.com/owiki/wiki/VOS/VirtSesame2Provider*](http://vos.openlinksw.com/owiki/wiki/VOS/VirtSesame2Provider)

Sau khi tạo 1 project java với tùy chọn là maven và import 2 thư viện của công nghệ virtuoso như hướng dẫn trên.

1 số thư viện cần import:



1 số cấu hình định nghĩa



Kết nối:



Ngắt kết nối:

