

ĐẠI HỌC QUỐC GIA TP.HCM
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN
KHOA: HỆ THỐNG THÔNG TIN



BÁO CÁO TIỂU LUẬN CUỐI KÌ
CƠ SỞ DỮ LIỆU PHÂN TÁN - IS211.012

Đề tài
Hệ quản trị cơ sở dữ liệu NoSQL MarkLogic Server

Sinh viên thực hiện	MSSV
Nguyễn Hữu Phụng	21522481
Trần Minh Quang	21522519
Đặng Lưu Hà	21520798

GVHD: NGUYỄN HỒ DUY TRI

MỤC LỤC

PHẦN 1. GIỚI THIỆU.....	3
1.1. Tên HQT CSDL.....	3
1.2. Lịch sử hình thành:	3
1.3. Tác giả, tổ chức quản lý:.....	4
1.4. Mô hình lưu trữ (key – value, document, graph, hybrid...)	4
1.5. Ngôn ngữ thao tác với dữ liệu	4
1.6. Cơ chế phân tán (cách phân tán, mô hình phân tán dữ liệu.....)	5
PHẦN 2. HƯỚNG DẪN CÀI ĐẶT (MÔ TẢ CHI TIẾT CẤU HÌNH MÁY, CÁC BƯỚC THỰC HIỆN, HÌNH MINH HỌA...)	7
2.1. Cài đặt trên một máy	7
2.2. Cài đặt trên cụm máy phân tán	13
PHẦN 3. THỰC NGHIỆM MÔ PHỎNG PHÂN TÁN	17
3.1. Mô tả bài toán đặt ra với dữ liệu	17
3.2. Mô tả cấu trúc dữ liệu sử dụng	17
3.3. Các bước thực nghiệm chi tiết từ lúc đưa dữ liệu vào đến lúc truy vấn phân tán, kèm theo minh họa	20
3.3.1. Tạo cơ sở dữ liệu	20
3.3.2. Truy vấn dữ liệu của Host B từ Host A và ngược lại	23
KẾT LUẬN	30
TÀI LIỆU THAM KHẢO.....	31

PHẦN 1. GIỚI THIỆU

1.1. Tên HQT CSDL

MarkLogic Server là cơ sở dữ liệu hướng tài liệu được phát triển bởi MarkLogic. Đó là cơ sở dữ liệu đa mô hình NoSQL được phát triển từ cơ sở dữ liệu XML để lưu trữ nguyên bản các tài liệu XML, JSON và bộ ba RDF, mô hình dữ liệu cho ngữ nghĩa.

MarkLogic được thiết kế để trở thành trung tâm dữ liệu cho dữ liệu vận hành và phân tích.

1.2. Lịch sử hình thành:

MarkLogic Corporation được thành lập vào năm 2001 bởi Christopher Lindblad và Paul Pedersen. Hiện tại, MarkLogic Corporation có trụ sở tại San Carlos, California, Hoa Kỳ và có hơn 500 nhân viên. Nó được mua lại bởi Progress Software vào tháng 2 năm 2023 với giá 355 triệu đô la.

MarkLogic Server là một cơ sở dữ liệu NoSQL do MarkLogic Corporation phát triển và cung cấp.

MarkLogic Server đã phát triển qua các phiên bản sau:

2005	MarkLogic Server 3.0
2006	MarkLogic Server 3.1
2007	MarkLogic Server 3.2
2008	MarkLogic Server 4.0
2009	MarkLogic Server 4.1
2010	MarkLogic Server 4.2
2011	MarkLogic Server 5.0
2012	MarkLogic Server 6.0
2013	MarkLogic Server 7.0
2015	MarkLogic Server 8.0
2017	MarkLogic Server 9.0
2017	MarkLogic Server 10.0
2022	MarkLogic Server 11.0

1.3. Tác giả, tổ chức quản lý:

- MarkLogic Corporation (2001).
- Progress Software (02/2023).

1.4. Mô hình lưu trữ (key – value, document, graph, hybrid...)

MarkLogic Server sử dụng mô hình lưu trữ multi-model gồm các mô hình lưu trữ:

Mô hình tài liệu (Documents)

Tài liệu là đơn vị dữ liệu cơ bản của MarkLogic Server. Tài liệu có thể được định dạng bằng XML, JSON, hoặc bất kỳ định dạng văn bản nào khác.

Mô hình đồ thị ngữ nghĩa (RDF triples)

MarkLogic Server sử dụng đồ thị ngữ nghĩa để lưu trữ mối quan hệ giữa các tài liệu. Trong MarkLogic Server, các tài liệu được đại diện bởi các nút trong đồ thị ngữ nghĩa. Các mối quan hệ giữa các tài liệu được đại diện bởi các cạnh trong đồ thị ngữ nghĩa.

Đồ thị ngữ nghĩa là dạng cơ sở dữ liệu đồ thị lưu trữ các sự việc theo ngữ nghĩa.

Ví dụ: Một người cụ thể đã tham dự một trường đại học cụ thể, nằm ở một thành phố cụ thể, thuộc về một quốc gia cụ thể và có một phân loại cụ thể.

1.5. Ngôn ngữ thao tác với dữ liệu

MarkLogic Server cho phép người dùng sử dụng các ngôn ngữ để truy vấn dữ liệu như.

XQuery

Là một ngôn ngữ truy vấn mạnh mẽ và linh hoạt có thể được sử dụng để thao tác với dữ liệu XML và tài liệu nói chung.

XSLT

Là một ngôn ngữ bảng mẫu giúp biến đổi nội dung trong quá trình nhập và xuất.

SPARQL

Là một ngôn ngữ truy vấn giống SQL để truy xuất dữ liệu ngữ nghĩa.

JavaScripts

Là một ngôn ngữ động, dựa trên đối tượng, vượt trội trong việc làm việc với JSON.

1.6. Cơ chế phân tán (cách phân tán, mô hình phân tán dữ liệu...)

MarkLogic Server sử dụng cơ chế phân tán dựa trên cụm. Mỗi cụm (cluster) có nhiều máy (máy chủ, host), mỗi host trong cụm có khi được gọi là nút (node) và mỗi nút trong cụm có bản sao riêng chứa tất cả thông tin cấu hình cho toàn bộ cụm.

Khi được triển khai dưới dạng cụm (cluster), Host sẽ triển khai kiến trúc shared-nothing. Không có máy chủ duy nhất phụ trách. mỗi máy chủ giao tiếp với mọi máy chủ khác và mỗi nút trong cụm duy trì bản sao cấu hình của riêng nó. Cơ sở dữ liệu bảo mật cũng như tất cả các cơ sở dữ liệu khác trong cụm đều có sẵn cho mỗi nút trong cụm. Kiến trúc shared-nothing này có lợi thế lớn khi nói đến khả năng mở rộng và tính sẵn sàng. Khi nhu cầu mở rộng tăng lên, chỉ cần thêm nhiều nút hơn vào cụm.

Cụm (cluster) là một tập hợp gồm một hoặc nhiều máy chủ (Host) sẽ hoạt động cùng nhau như một tổng thể thống nhất.

Máy chủ (Host) là một phiên bản MarkLogic duy nhất chạy trên một máy.

MarkLogic cho phép cấu hình nhiều máy chủ.

Nhóm (Group) là một tập hợp các máy chủ. Nhóm được sử dụng để đơn giản hóa việc quản lý cụm.

Forest là một kho lưu trữ tài liệu. Mỗi forest được quản lý bởi một máy chủ duy nhất. Máy chủ có thể quản lý nhiều forest cùng một lúc.

Database là một tập hợp gồm một hoặc nhiều forest. Mỗi forest trong cơ sở dữ liệu phải được cấu hình nhất quán.

Collection (Stand) giúp tổ chức và quản lý dữ liệu. Mỗi document có thể thuộc vào một hoặc nhiều collection. Collection giúp tổ chức các document thành các nhóm có chủ đề hoặc mục đích cụ thể.

PHẦN 2. HƯỚNG DẪN CÀI ĐẶT (MÔ TẢ CHI TIẾT CẤU HÌNH MÁY, CÁC BƯỚC THỰC HIỆN, HÌNH MINH HỌA...)

2.1. Cài đặt trên một máy

Yêu cầu cấu hình:

- Hệ điều hành:

Hỗ trợ trên nhiều hệ điều hành như Windows, Linux và macOS.

- Bộ xử lý:

Tối thiểu CPU có 4 lõi.

- Bộ nhớ RAM:

Ít nhất 8 GB RAM.

- Dung lượng đĩa cứng:

Ít nhất 20 GB dung lượng đĩa cứng.

Bước 1: Truy cập vào liên kết <https://developer.marklogic.com/products/marklogic-server> và chọn bản cài đặt phù hợp.

MarkLogic 11

11.1.0, Released October 2023

[What's New? »](#)


[Read about the Free Developer License »](#)

[Review System Requirements »](#)

Cloud Platform

 [Deploying on the Cloud](#)

Microsoft Windows (7, 10, Server 2012, 2016 and 2019)

 MarkLogic Server x64 (AMD64, Intel EM64T) 64-bit Windows Installer	694.6 MB  (SHA1)
 MarkLogic Server (GPU-enabled) x64 (AMD64, Intel EM64T) 64-bit Windows Installer	695.5 MB  (SHA1)
 Converters & Filters for Microsoft Windows	56.9 MB  (SHA1)

Mac OS X

 MarkLogic Server OS X disk image	184.8 MB  (SHA1)
 Filters for OS X	39.6 MB  (SHA1)

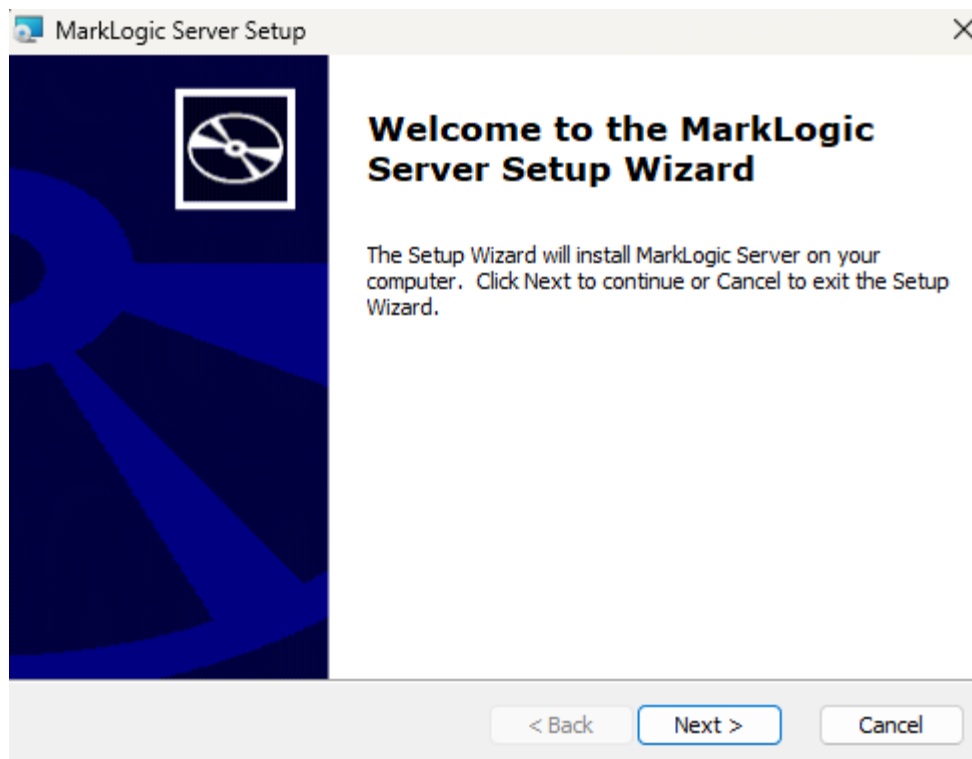
Red Hat Enterprise Linux / CentOS, Version 7 and 8

 MarkLogic Server x64 (AMD64, Intel EM64T) 64-bit Linux RPM	659 MB  (SHA1)
 Converters & Filters for Linux	74.3 MB  (SHA1)

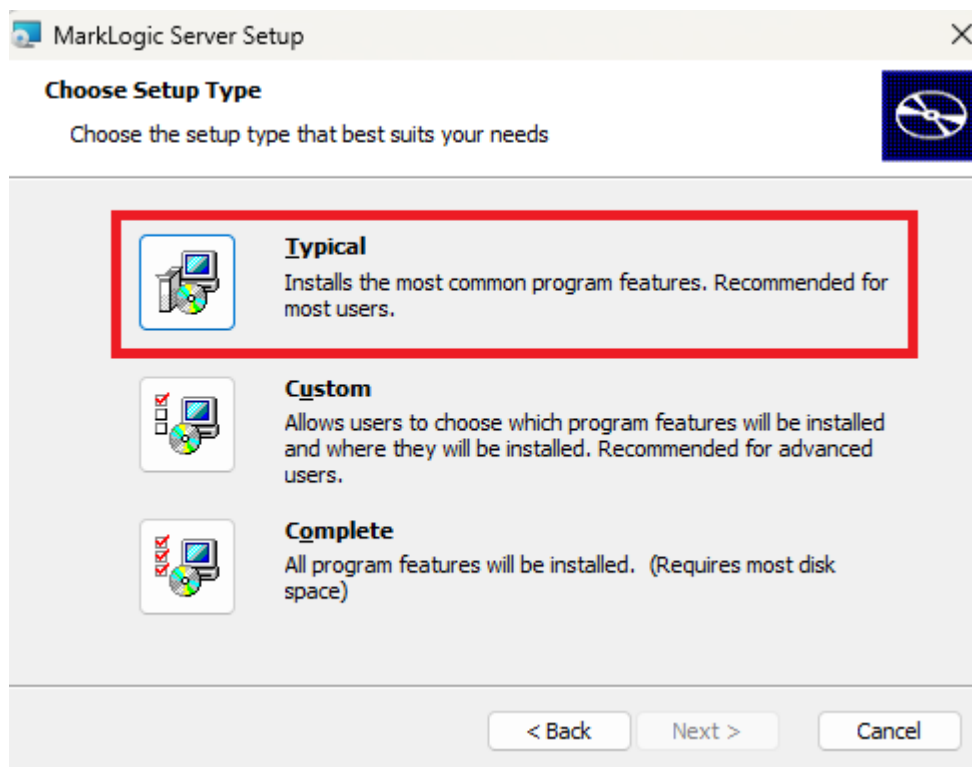
Machine Learning Models

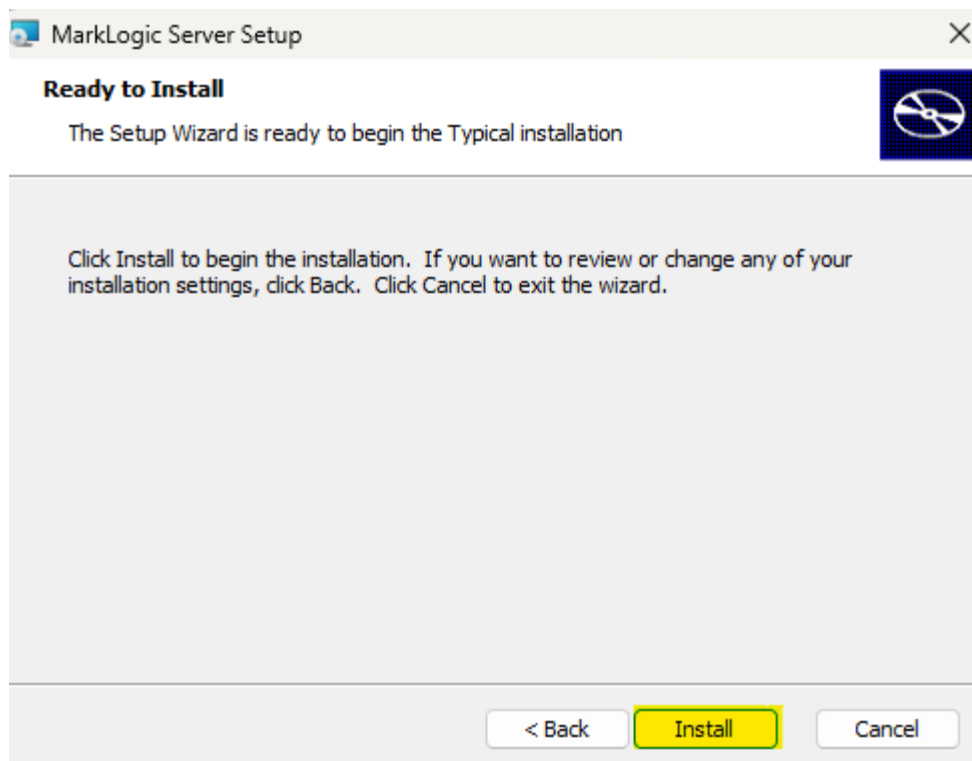
 Example of converting PyTorch model to ONNX	3.0 GB  (SHA1)
---	---

Bước 2: Chạy trình cài đặt cho MarkLogic Server.

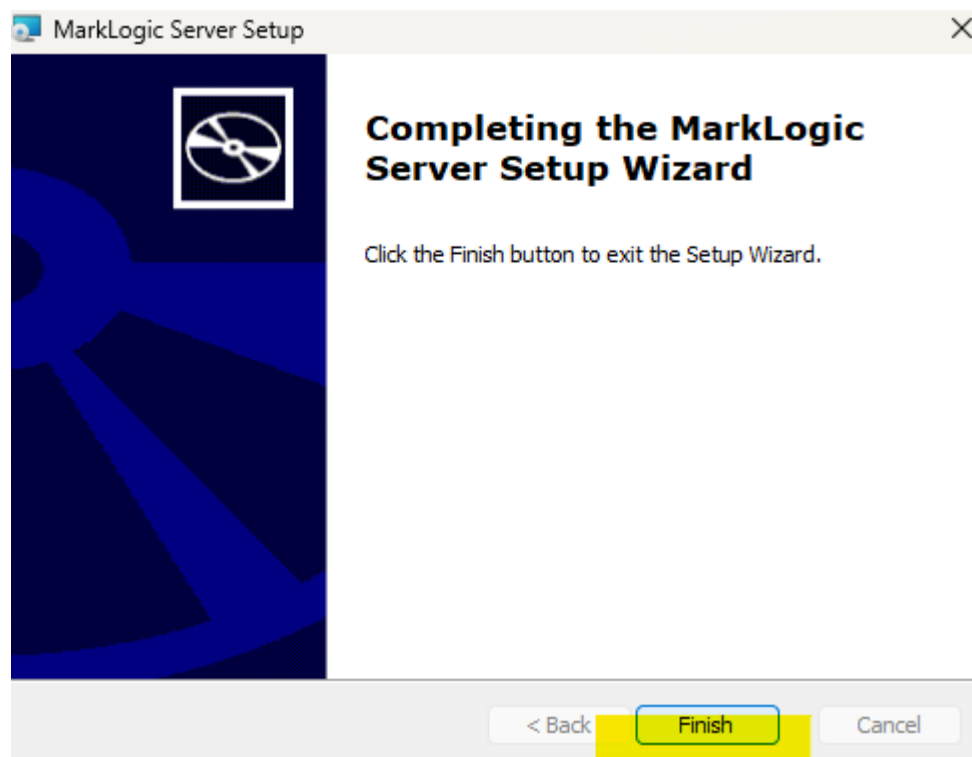


Bước 3: Click next và chọn Typical. Rồi chọn Install.

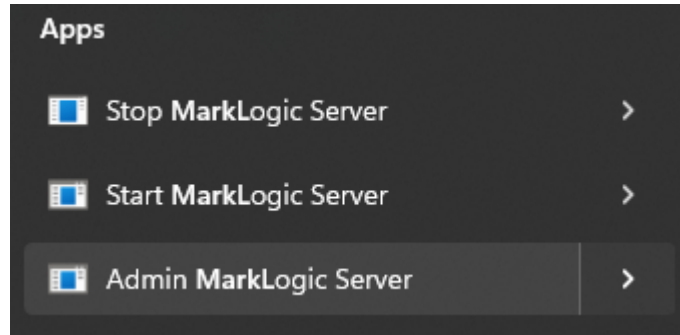




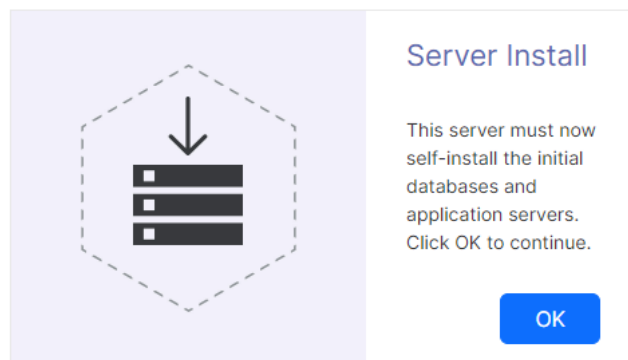
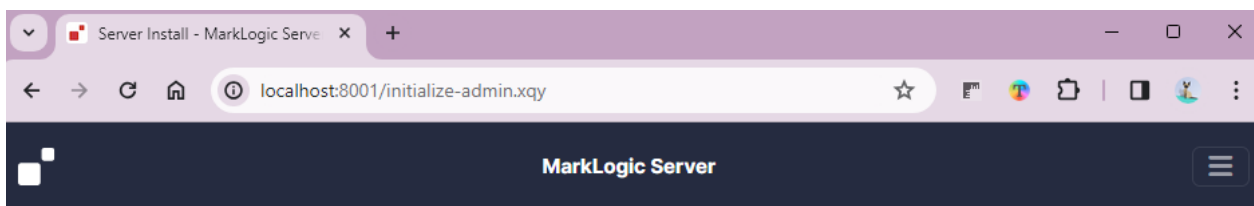
Bước 4: Click Finish để hoàn tất quá trình cài đặt.



Bước 5: Khởi động MarkLogic Server bằng cách tìm kiếm MarkLogic Server ở thanh tìm kiếm trên taskbar, Right-click vào Start MarkLogic Server rồi chọn Run as administrator.



Bước 6: Vào trình duyệt web, truy cập <http://localhost:8001>, nhấn Ok.



Bước 7: Nhấn Skip và điền thông tin cho admin ở trang tiếp theo rồi nhấn Ok.

Join a Cluster

Now that MarkLogic Server is installed on this host, you can join an existing cluster. In order to do so, enter the host name of one of the cluster's hosts and provide the port number of that host's administration interface.

Press 'Skip' if you do not wish to join a cluster.

Host Name	<input type="text"/>
	One of the target cluster's hosts Required.
Admin Port	<input type="text" value="8001"/>
	Port for admin interface on server Required.
Protocol	<input type="text" value="http"/>
	Whether the host's admin server has SSL enabled.

Skip

OK

Security Setup

MarkLogic Server has detected that Administration has not been secured. Please supply a user name and password for the Administrative user to set up security.

You also need to specify a realm for this security database. This is the realm that will be displayed to clients authenticating against this database. Since this value is used in password hashes it is recommended that you not change this value once it is set. Please read the further documentation about realms.

Admin	<input type="text" value="admin"/>
	User/login name (unique) Required. You must supply a value for user-name.
Admin Password	<input type="password" value="*****"/>
	Encrypted Password. Required.
Confirm Admin Password	<input type="password" value="*****"/>
	Encrypted Password. Required.
Realm	<input type="text" value="public"/>
	The authentication realm.

MarkLogic Server comes with a built-in PKCS#11 wallet, please provide a password to secure it.

Wallet Password	<input type="password" value="*****"/>
	Encrypted Password. Required.
Confirm Wallet Password	<input type="password" value="*****"/>
	Encrypted Password. Required.

Encrypt Security Database	<input type="checkbox"/>
Encrypt Audit Log	<input type="checkbox"/>

OK

Bước 8: Đăng nhập vào admin interface.

Đăng nhập

http://localhost:8001

Tên người dùng

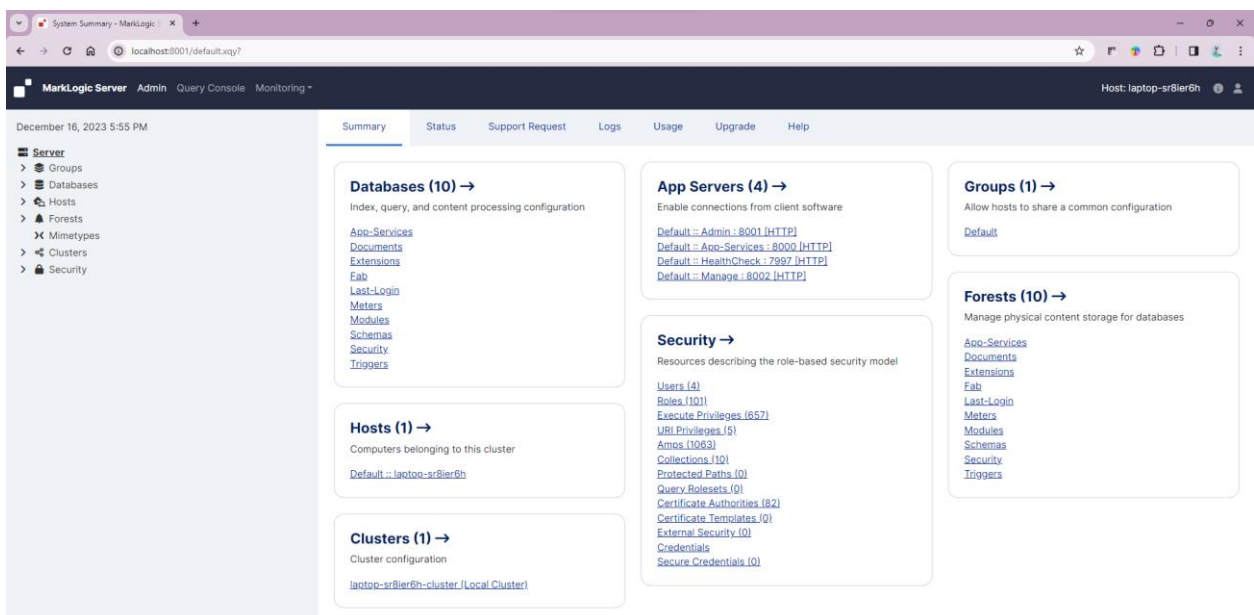
admin

Mật khẩu

.....

Đăng nhập

Hủy



2.2. Cài đặt trên cụm máy phân tán

Yêu cầu:

Số lượng nút (Node):

Tùy thuộc vào quy mô dự án. Càng nhiều nút, càng tăng khả năng chịu lỗi và tăng hiệu suất.

Mạng:

Đảm bảo có kết nối mạng ổn định giữa các nút.

Bộ xử lý và RAM cho mỗi nút:

Tương tự như localhost, mỗi nút cần ít nhất 4 lõi CPU và 8 GB RAM.

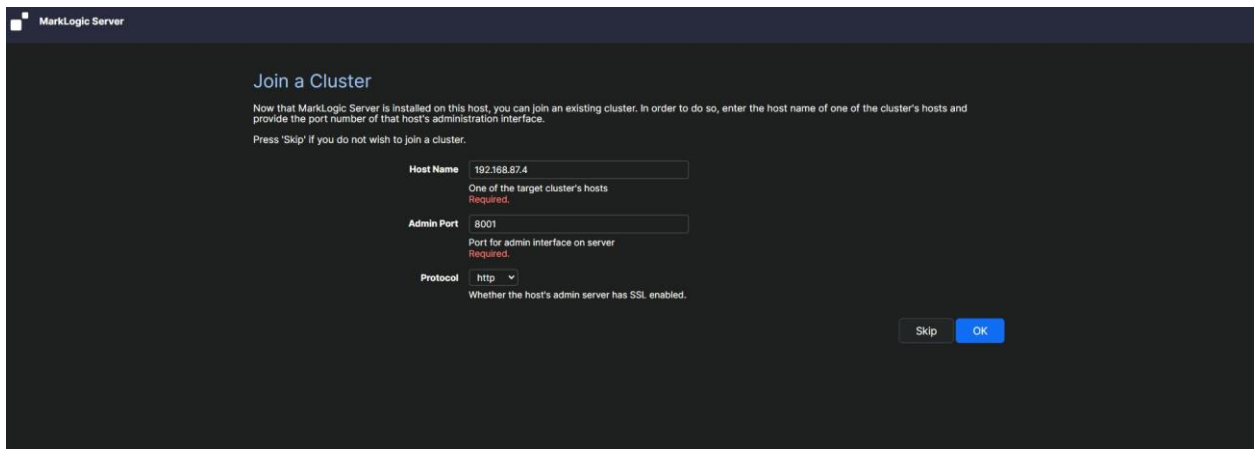
Firewall và Bảo mật:

Cấu hình tường lửa để cho phép giao tiếp giữa các nút.

Xác thực và quản lý quyền truy cập một cách an toàn.

Bước 1: Cài đặt giống như trên 1 máy từ bước 1 cho tới bước 6.

Bước 2: Tại màn hình Join a cluster, nhập địa chỉ IP của máy đầu tiên vào ô Host name và nhấn Ok.



MarkLogic Server

Join a Cluster

Now that MarkLogic Server is installed on this host, you can join an existing cluster. In order to do so, enter the host name of one of the cluster's hosts and provide the port number of that host's administration interface.
Press 'Skip' if you do not wish to join a cluster.

Host Name 192.168.87.4
One of the target cluster's hosts
Required.

Admin Port 8001
Port for admin interface on server
Required.

Protocol http
Whether the host's admin server has SSL enabled.

Skip OK

Bước 3: Đăng nhập vào admin interface bằng tên người dùng và mật khẩu mà máy 1 đã tạo.

Sign in

http://192.168.87.4:8001

Your connection to this site is not private

Username

Password

Sign in

Cancel

Bước 4: Nhấn Ok cho đến khi vào được giao diện Admin Interface.

MarkLogic ServerAdminQuery ConsoleMonitoring

Host: laptop-sr8ier6h

Join a Cluster

In order to join this MarkLogic cluster, the new host must be assigned to one of the cluster's groups and its name confirmed. Please select a group to join and confirm the new host's name.

Group

Default

The groups in the cluster.

Host Name

kuaq

The name of the new host in the cluster.

Zone

The zone of the new host in the cluster (optional).

Cancel

OK

MarkLogic ServerAdminQuery ConsoleMonitoring

Host: laptop-sr8ier6h

Join a Cluster

You are about to join this MarkLogic cluster. To proceed click OK.

Cancel

OK

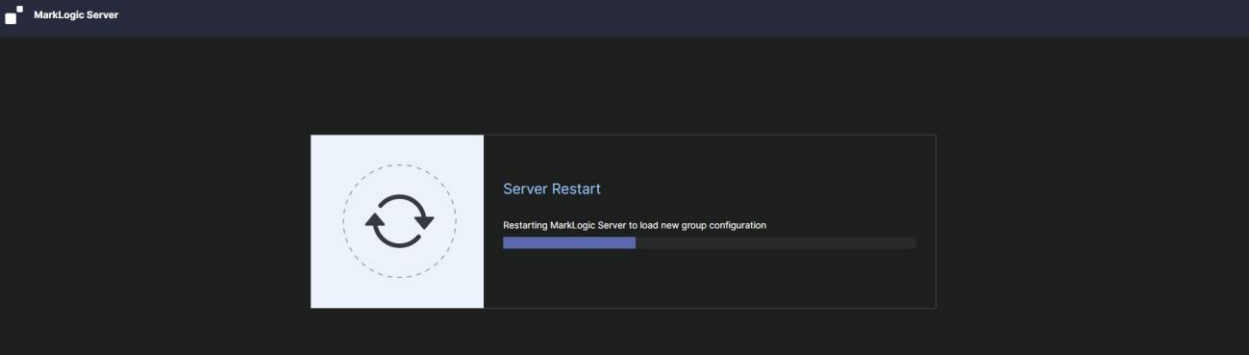
MarkLogic ServerAdminQuery ConsoleMonitoring

Host: laptop-sr8ier6h

Join the Cluster

A new host has joined the MarkLogic cluster. Press 'OK' to transfer cluster configuration information to the new host.

OK



MarkLogic Server Admin Query Console Monitoring

Host: kuaq

December 23, 2023 4:22 PM

Server

- Groups
- Databases
- Hosts
 - desktop-54e2938
 - kuaq
 - laptop-sr8ier6h
- Forests
- Mimetypes
- Clusters
- Security

Summary Help

Hostname	Software Version	Status	Group	Zone	Forests
kuaq	11010000	local	Default		chinhanh3
laptop-sr8ier6h	11010000	connected	Default		App-Services chinhanh1 CN1 Documents Extensions Fab Last-Login Meters Modules Schemas Security IDggets
desktop-54e2938		disconnected			Chihhanh2

Essential Enterprise 111.0 - Copyright © 2023 MarkLogic Corporation

PHẦN 3. THỰC NGHIỆM MÔ PHỎNG PHÂN TÁN

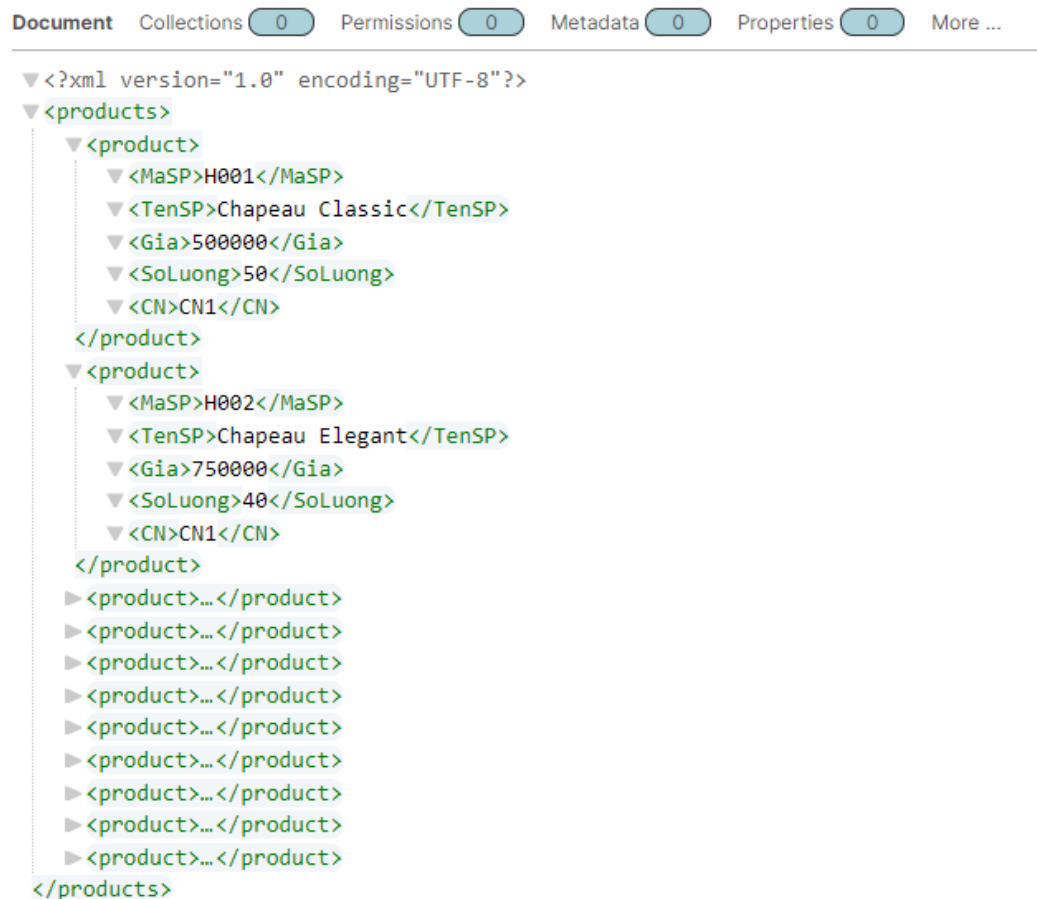
3.1. Mô tả bài toán đặt ra với dữ liệu

Xây dựng một hệ thống chuỗi cửa hàng bán cà phê gồm 2 cửa hàng ở vị trí khác nhau, mỗi cửa hàng có danh sách các mặt hàng cà phê có sẵn tại cửa hàng đó. Thiết kế cơ sở dữ liệu phân tán bao gồm nhiều cửa hàng, cho phép các cửa hàng truy xuất, tra cứu dữ liệu của các cửa hàng trong chuỗi cửa hàng với nhau thông qua bất kì một cửa hàng trong chuỗi để biết được mặt hàng cà phê đó có sẵn ở cửa hàng nào.

3.2. Mô tả cấu trúc dữ liệu sử dụng

Cấu trúc dữ liệu được chia thành 4 tài liệu gồm: 3 tài liệu lưu trữ thông tin của 3 loại sản phẩm (cà phê) ở cửa hàng tương ứng; 1 tài liệu lưu trữ thông tin, địa chỉ của cửa hàng trong chuỗi cửa hàng.

Document SP-Hat.xml



Document SP-Hat.xml gồm tập hợp các sản phẩm (Cà phê dạng hạt) chứa thông tin của cà phê hạt như: MaSP (mã sản phẩm), TenSP (Tên sản phẩm), Gia (Giá sản phẩm), SoLuong (Số lượng), CN (Mã chi nhánh).

Document SP-HoaTan.xml

```
▼<?xml version="1.0" encoding="UTF-8"?>
▼<products>
  ▼<product>
    ▼<MaSP>HT001</MaSP>
    ▼<TenSP>Hoa Tan Premium</TenSP>
    ▼<Gia>1200000</Gia>
    ▼<SoLuong>30</SoLuong>
    ▼<CN>CN1</CN>
  </product>
  ▼<product>
    ▼<MaSP>HT002</MaSP>
    ▼<TenSP>Hoa Tan Deluxe</TenSP>
    ▼<Gia>900000</Gia>
    ▼<SoLuong>25</SoLuong>
    ▼<CN>CN1</CN>
  </product>
  ▶<product>...</product>
  ▶<product>...</product>
  ▶<product>...</product>
  ▶<product>...</product>
  ▶<product>...</product>
  ▶<product>...</product>
  ▶<product>...</product>
  ▶<product>...</product>
  ▼<product>
    ▼<MaSP>HT011</MaSP>
    ▼<TenSP>Hoa Tan Supreme</TenSP>
    ▼<Gia>2000000</Gia>
    ▼<SoLuong>15</SoLuong>
    ▼<CN>CN1</CN>
  </product>
</products>
```

Document SP-Hat.xml gồm tập hợp các sản phẩm (Cà phê dạng hòa tan) chứa thông tin của cà phê hòa tan như: MaSP (mã sản phẩm), TenSP (Tên sản phẩm), Gia (Giá sản phẩm), SoLuong (Số lượng), CN (Mã chi nhánh).

Document SP-RangXay.xml

```
<?xml version="1.0" encoding="UTF-8"?>
<products>
  <product>
    <MaSP>RX001</MaSP>
    <TenSP>Rang Xay Pro</TenSP>
    <Gia>3000000</Gia>
    <SoLuong>20</SoLuong>
    <CN>CN1</CN>
  </product>
  <product>
    <MaSP>RX002</MaSP>
    <TenSP>Rang Xay Ultimate</TenSP>
    <Gia>2500000</Gia>
    <SoLuong>15</SoLuong>
    <CN>CN1</CN>
  </product>
  <product>...</product>
  <product>...</product>
  <product>...</product>
  <product>...</product>
  <product>...</product>
  <product>...</product>
  <product>...</product>
  <product>...</product>
  <product>...</product>
</products>
```

Document SP-Hat.xml gồm tập hợp các sản phẩm (Cà phê rang xay) chứa thông tin của cà phê rang xay như: MaSP (mã sản phẩm), TenSP (Tên sản phẩm), Gia (Giá sản phẩm), SoLuong (Số lượng), CN (Mã chi nhánh).

Document ChiNhanh.xml

```
▼<?xml version="1.0" encoding="UTF-8"?>
▼<stores>
  ▼<store>
    ▼<MaCH>CH1</MaCH>
    ▼<TenCH>Cửa hàng A</TenCH>
    ▼<DiaChi>123 Đường ABC, Quận 1, Thành phố Hồ Chí Minh</DiaChi>
    ▼<SDT>0123456789</SDT>
  </store>
  ▼<store>
    ▼<MaCH>CH2</MaCH>
    ▼<TenCH>Cửa hàng B</TenCH>
    ▼<DiaChi>456 Đường XYZ, Quận 2, Thành phố Hồ Chí Minh</DiaChi>
    ▼<SDT>0987654321</SDT>
  </store>
</stores>
```

Document ChiNhanh.xml gồm tập hợp các chi nhánh chứa thông tin của chi nhánh như: MaCH (Mã cửa hàng), TenCH (Tên cửa hàng), DiaChi (Địa chỉ cửa hàng), SDT (Số điện thoại của cửa hàng).

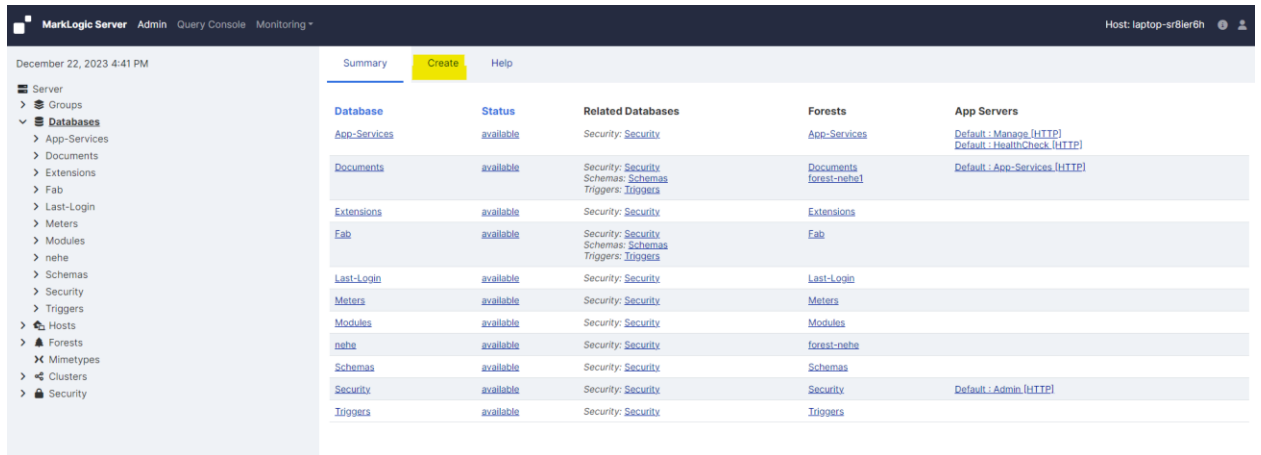
3.3. Các bước thực nghiệm chi tiết từ lúc đưa dữ liệu vào đến lúc truy vấn phân tán, kèm theo minh họa

3.3.1. Tạo cơ sở dữ liệu

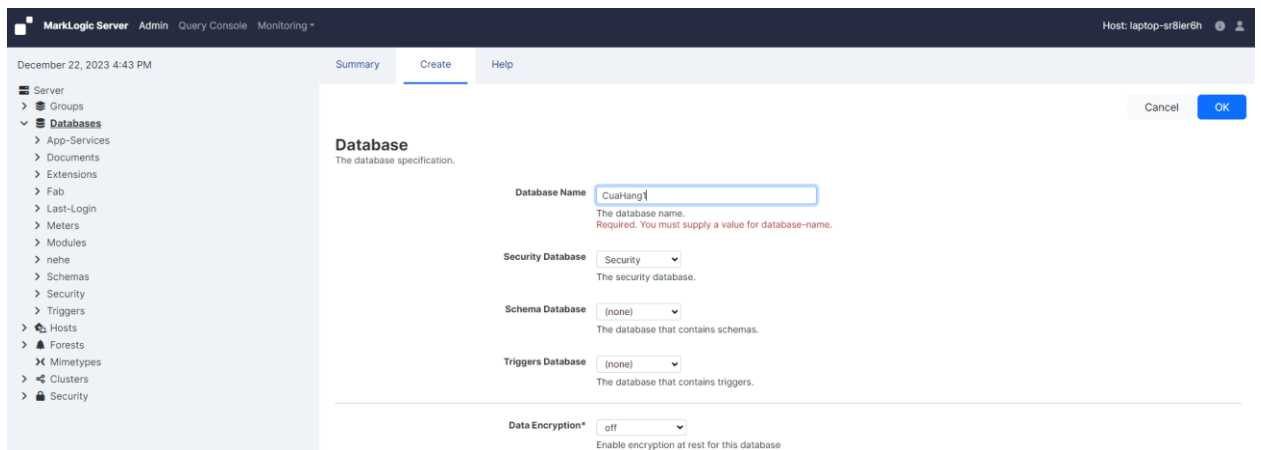
Trên Host A (CH1-cửa hàng 1).

Bước 1: Tạo cơ sở dữ liệu để lưu dữ liệu của CH1.

- Nhấn vào mục database ở thanh chọn bên trái.
- Nhấn Create để thêm mới một Database.



- Điền tên của Database rồi nhấn Ok



Bước 3: Tạo Forest để lưu trữ các tài liệu ở Cửa hàng 1

- Trên thanh chọn bên trái màn hình, nhấn vào Forest.
- Chọn tab Create và điền tên của Forest rồi nhấn Ok để thêm mới 1 Forest.

December 22, 2023 4:46 PM

Summary Create Help

Server

- Groups
- Databases
- Hosts
- Forests
 - App-Services
 - Documents
 - Extensions
 - Fab
 - forest-nehe
 - forest-nehe1
 - Last-Login
 - Meters
 - Modules
 - Schemas
 - Security
 - Triggers
 - Minetypes
 - Clusters
 - Security

Cancel OK

Create New Forests

Forest
The forest assignment specification.

Forest Name: forest-CuaHang1
The forest name.
Required. You must supply a value for forest-name.

Host: laptop-sr8ier6h
The primary host to which the forest is assigned.

Data Directory:
The optional public directory for forests.

Large Data Directory:
The optional directory for large objects in a forest.

Fast Data Directory:
The optional smaller but faster directory for forests.

Fast Data Max Size: 0
Maximum allowable size (in megabytes) of the data to be kept by this forest in the fast-data-directory, or 0 for no limit.

Bước 4: Gán cơ sở dữ liệu cho Forest vừa tạo

- Vào tab summary và tìm đến forest vừa tạo.
- Chọn database phù hợp cho forest vừa tạo ở cột database rồi nhấn Ok.

December 22, 2023 4:48 PM

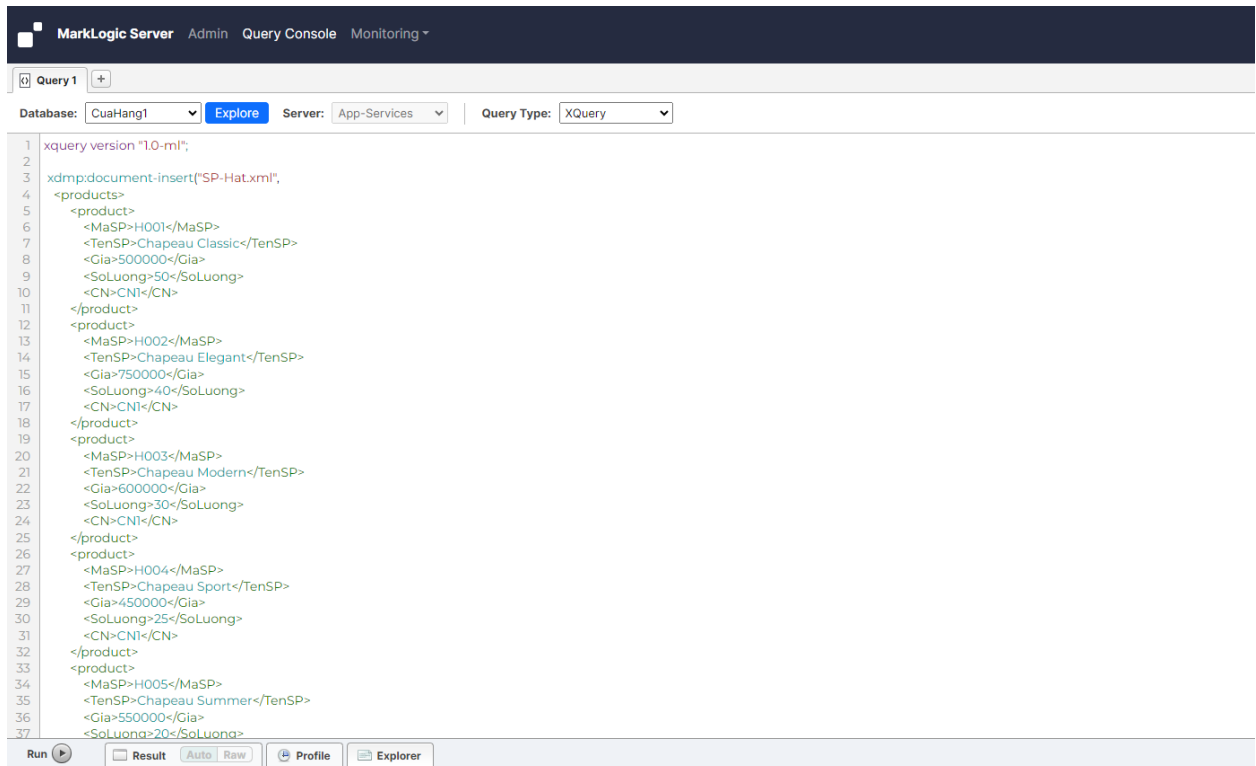
Summary Create Help

Cancel OK

Forest	Status	Database	Primary Host	Master Forest	Free Space	Data Dir
App-Services	open	App-Services	laptop-sr8ier6h		171,868 MB	D:\App\MarkLogic\Data
Documents	open	Documents	laptop-sr8ier6h		171,868 MB	D:\App\MarkLogic\Data
Extensions	open	Extensions	laptop-sr8ier6h		171,868 MB	D:\App\MarkLogic\Data
Fab	open	Fab	laptop-sr8ier6h		171,868 MB	D:\App\MarkLogic\Data
forest-CuaHang1	open		laptop-sr8ier6h		171,868 MB	D:\App\MarkLogic\Data
forest-nehe	open	nehe	laptop-sr8ier6h		171,868 MB	D:\App\MarkLogic\Data
forest-nehe1	open	Documents	laptop-sr8ier6h		171,868 MB	D:\App\MarkLogic\Data
Last-Login	open	Last-Login	laptop-sr8ier6h		171,868 MB	D:\App\MarkLogic\Data
Meters	open	Meters	laptop-sr8ier6h		171,868 MB	D:\App\MarkLogic>Data
Modules	open	Modules	laptop-sr8ier6h		171,868 MB	D:\App\MarkLogic>Data
Schemas	open	Schemas	laptop-sr8ier6h		171,868 MB	D:\App\MarkLogic>Data
Security	open	Security	laptop-sr8ier6h		171,868 MB	D:\App\MarkLogic>Data
Triggers	open	Triggers	laptop-sr8ier6h		171,868 MB	D:\App\MarkLogic>Data

Bước 5: Thêm tài liệu vào database của cửa hàng 1.

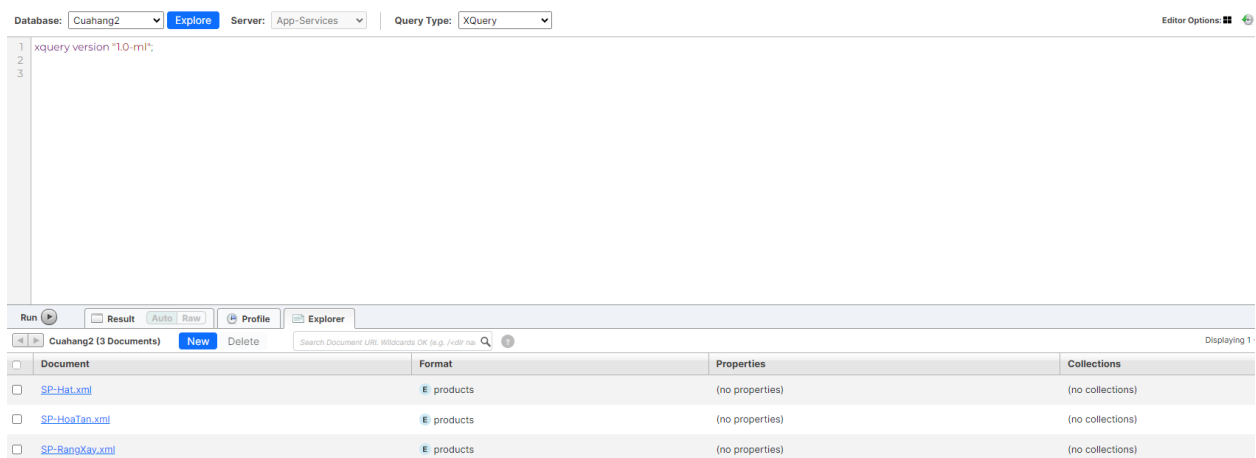
- Trên thanh navbar, nhấn vào Query console và đăng nhập bằng tài khoản mà host 1 đã tạo.
- Chọn cơ sở dữ liệu cần thêm trên thanh công cụ và sử dụng đoạn mã XQuery để thêm mới tài liệu vào cơ sở dữ liệu.



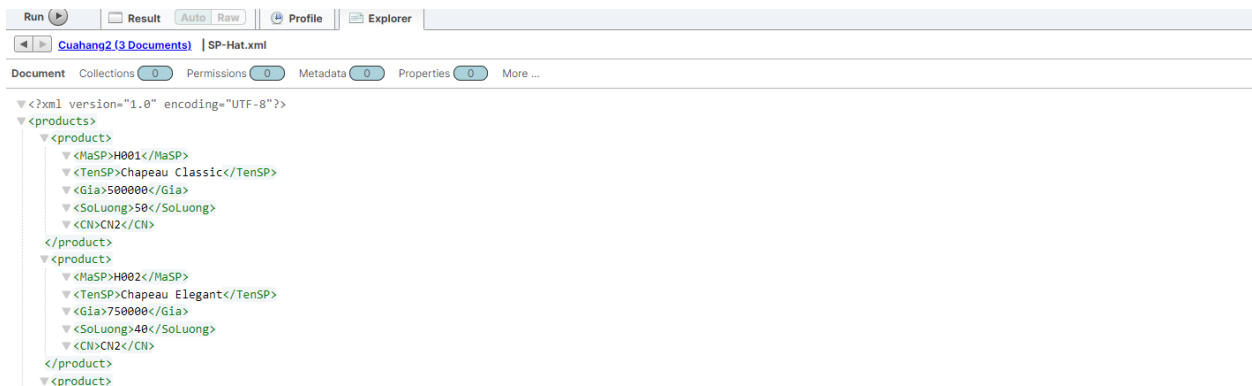
- Trên Host B (CH2-cửa hàng 2) Thực hiện tương tự như Host A nhưng dữ liệu thêm vào là dữ liệu của host B của Host B.

3.3.2. Truy vấn dữ liệu của Host B từ Host A và ngược lại

- *Xem dữ liệu của Host B từ Host A bằng cửa sổ Explore.*
- Trên Host A, chọn Database của Host B và nhấn Explore.

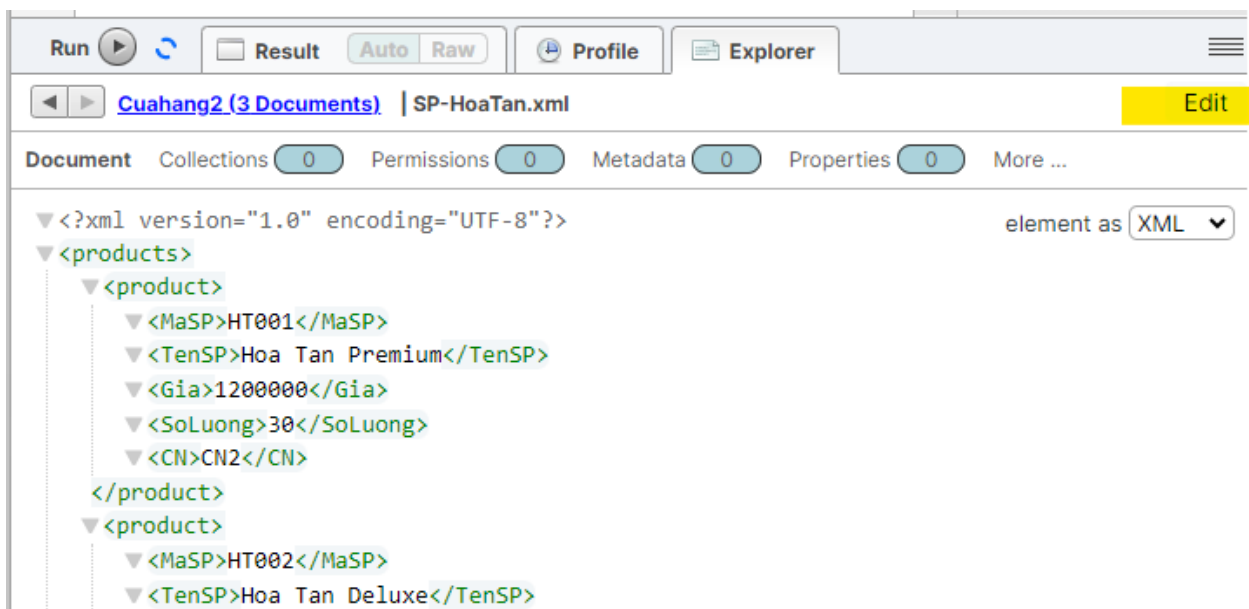


- Ở cửa sổ Results hiện toàn bộ tài liệu có trong Database của Host B. Để xem dữ liệu trong host B, nhấn vào tài liệu cần xem.



Thêm, xóa, sửa dữ liệu vào tài liệu trên cơ sở dữ liệu trên Host B

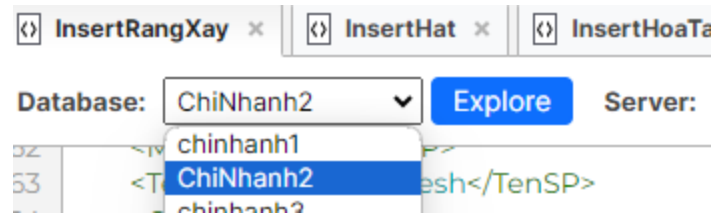
Cách 1: Chọn nút edit sau khi Explore tài liệu trên host B và thực hiện thao tác thêm, xóa, sửa.



Cách 2: Dùng câu lệnh XQuery để thực hiện thao tác thêm, xóa, sửa

Thêm mới 1 sản phẩm

Bước 1: Chọn Database là database của Cửa hàng 2



Database mà Host A muốn truy vấn chính là Database của Host B mà Host B đã tạo. Mọi câu lệnh thực hiện trên Host A sẽ tác động vào tài liệu nằm trong cơ sở dữ liệu của Host B mà Host A đã chọn.

Bước 2: Dùng câu lệnh XQuery để thêm mới 1 sản phẩm

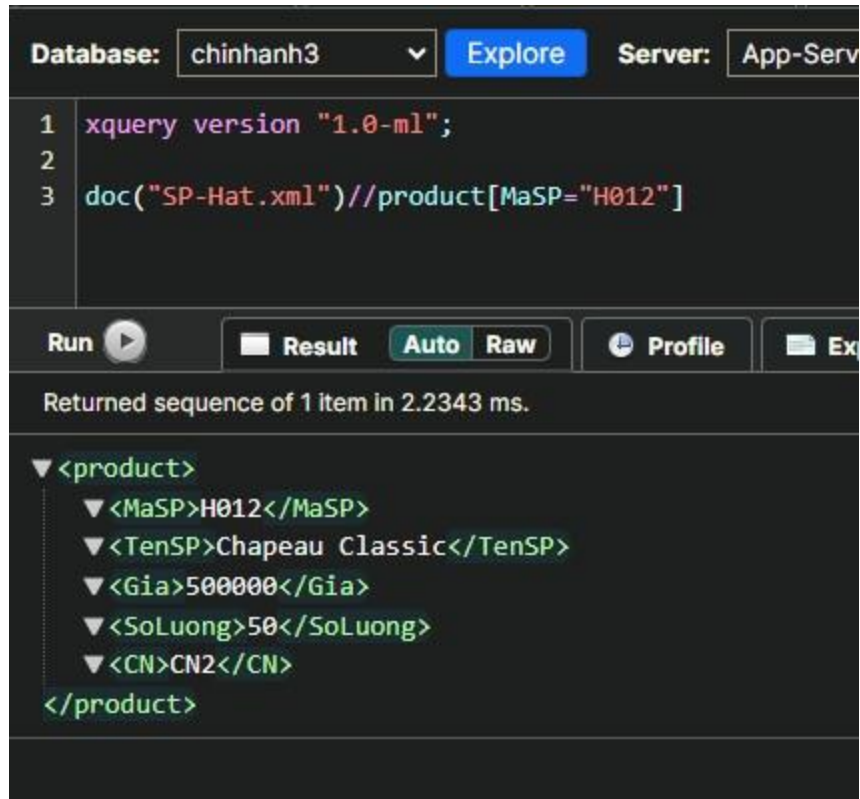
```
xquery version "1.0-ml";

xdmp:node-insert-child(doc("SP-Hat.xml/products"),
  <product>
    <MaSP>H0012</MaSP>
    <TenSP>Chapeau Classic</TenSP>
    <Gia>500000</Gia>
    <SoLuong>50</SoLuong>
    <CN>CN2</CN>
  </product>)
```

Câu lệnh trên được dùng để thêm mới một sản phẩm với các thông tin vào tài liệu “SP-Hat.xml”.

Hàm *node-insert-child()* dùng để thêm một node mới vào tài liệu với các tham số truyền vào là URI đến nơi cần thêm node vào và node cần thêm vào.

Bước 3: Kiểm tra trên Host B bằng cách dùng câu lệnh truy vấn sản phẩm vừa thêm trên tài liệu nằm ở cơ sở dữ liệu trên Host B (cũng có thể thực hiện trên Host A).

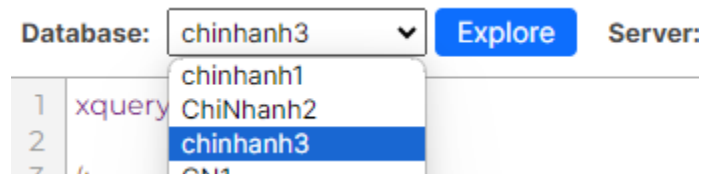


Câu lệnh trên dùng để xem sản phẩm có mã sản phẩm [MaSP] là “H012”.

Cửa sổ Result hiện kết quả của câu lệnh XQuery.

Xóa một sản phẩm trong tài liệu

Bước 1: Chọn Database là database của Cửa hàng 2 (Hoặc database tương ứng) của Host B.



Bước 2: Dùng câu lệnh XQuery để tiến hành xóa sản phẩm trong tài liệu của Host B (Sản phẩm có MaSP = “H012”).

```

for $product in doc("SP-Hat.xml")//product
where $product[MaSP="H012"]
return
  xdmp:node-delete($product)

```

Bước 3: Kiểm tra trên Host B bằng cách dùng câu lệnh truy tài liệu nằm ở cơ sở dữ liệu trên Host B (cũng có thể thực hiện trên Host A) và kiểm tra sự tồn tại của (Sản phẩm có MaSP = “H012”)

Database: chinhanh3 **Explore** Server: App-Services

```

1 xquery version "1.0-m1";
2
3 doc("SP-Hat.xml")

```

Run **Result** Auto Raw Profile Explorer

Returned sequence of 1 item in 1.3498 ms. (~0.8845 ms. compared to previous run)

```

  <CN>CN2</CN>
</product>
  <product>
    <MaSP>H008</MaSP>
    <TenSP>Chapeau Casual</TenSP>
    <Gia>650000</Gia>
    <SoLuong>12</SoLuong>
    <CN>CN2</CN>
  </product>
  <product>
    <MaSP>H009</MaSP>
    <TenSP>Chapeau Chic</TenSP>
    <Gia>900000</Gia>
    <SoLuong>10</SoLuong>
    <CN>CN2</CN>
  </product>
  <product>
    <MaSP>H010</MaSP>
    <TenSP>Chapeau Vintage</TenSP>
    <Gia>850000</Gia>
    <SoLuong>8</SoLuong>
    <CN>CN2</CN>
  </product>
  <product>
    <MaSP>H011</MaSP>
    <TenSP>Chapeau Exclusive</TenSP>
    <Gia>1200000</Gia>
    <SoLuong>25</SoLuong>
    <CN>CN2</CN>
  </product>
</products>

```

Kết quả cho thấy trong tài liệu trên Host B không có sự hiện diện của Sản phẩm có MaSP = “H012” vì đã thực hiện thao tác xóa trên Host A.

Sửa một sản phẩm trong tài liệu

Bước 1: Chọn Database là database tương ứng của Host B.

Bước 2: Dùng câu lệnh XQuery để tiến hành sửa sản phẩm trong tài liệu của Host B.

The screenshot shows the Azure Cosmos DB Explorer interface. At the top, the 'Database' is set to 'chinhanh3' and the 'Server' is 'App-Services'. The 'Query Type' is set to 'XQuery'. The main area contains the following XQuery:

```
xquery version "1.0-ml";

for $product in doc("SP-Hat.xml")//product
where $product[MaSP="H001"]
return
  xdmp:node-replace($product/SoLuong, <SoLuong>300</SoLuong>)
```

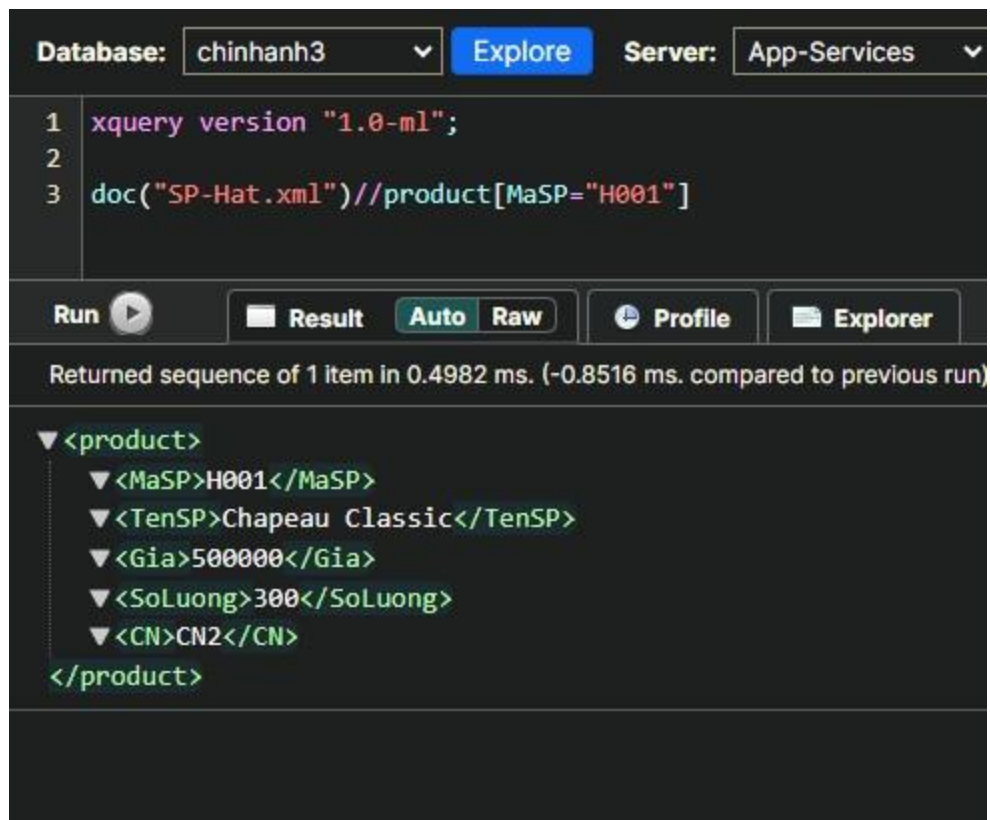
Below the query, there are buttons for 'Run', 'Result', 'Auto', 'Raw', 'Profile', and 'Explorer'. The 'Run' button is highlighted. Below the buttons, the document 'chinhanh3 (1 Documents)' is selected, and the file 'SP-Hat.xml' is shown. The document content is displayed in a tree view:

```
<?xml version="1.0" encoding="UTF-8"?>
<products>
  <product>
    <MaSP>H001</MaSP>
    <TenSP>Chapeau Classic</TenSP>
    <Gia>500000</Gia>
    <SoLuong>50</SoLuong>
    <CN>CN2</CN>
```

Câu lệnh trên dùng để sửa số lượng của sản phẩm có mã sản phẩm là “H001”.

Khi Explore tài liệu của Host B trên Host A, kết quả cho thấy ban đầu số lượng của sản phẩm này là 50.

Bước 3: Kiểm tra trên Host B bằng cách dùng câu lệnh truy vấn sản phẩm vừa thêm trên tài liệu nằm ở cơ sở dữ liệu trên Host B (cũng có thể thực hiện trên Host A).



Sau khi kiểm tra trên Host B, số lượng sản phẩm đó đã được thay đổi thành 300.

KẾT LUẬN

MarkLogic Server không chỉ là một hệ quản trị cơ sở dữ liệu mạnh mẽ mà còn là một giải pháp linh hoạt cho các hệ thống có yêu cầu về tính phân tán cao.

Một trong những điểm đáng chú ý là khả năng xử lý dữ liệu phân tán của MarkLogic Server đã giúp tối ưu hóa hiệu suất và đảm bảo sự ổn định trong quá trình thực thi các truy vấn và giao tiếp giữa các nút dữ liệu. Điều này làm tăng cường khả năng mở rộng của hệ thống, giúp nó linh hoạt đáp ứng được sự thay đổi trong quy mô và yêu cầu của hệ thống.

Tóm lại, MarkLogic Server không chỉ là một lựa chọn xuất sắc cho các ứng dụng cơ sở dữ liệu NoSQL mà còn là một giải pháp đáng tin cậy và hiệu quả cho các hệ thống đòi hỏi tính phân tán cao. Các khía cạnh tích cực trong nghiên cứu này mở ra cơ hội cho việc triển khai và mở rộng ứng dụng NoSQL phức tạp trong các môi trường phân tán và đòi hỏi sự linh hoạt.

TÀI LIỆU THAM KHẢO

[1] Inside-MarkLogic-Server

<https://www.marklogic.com/wp-content/uploads/2018/01/Inside-MarkLogic-Server.pdf>

[2] MarkLogic Server Document, Getting Started Guides

<https://docs.marklogic.com/guide/getting-started>