TRƯỜNG ĐẠI HỌC SƯ PHẠM THÀNH PHỐ HỒ CHÍ MINH

KHOA CÔNG NGHỆ THÔNG TIN

\_\_🕮\_\_

**ĐỒ ÁN MÔN CƠ SỞ DỮ LIỆU NÂNG CAO**

**ĐỀ TÀI: TÌM HIỂU VỀ MONGODB**

**Giảng viên hướng dẫn:**Ths. Lương Trần Hy Hiến

**Sinh viên thực hiện:**Nhóm 10

Nguyễn Cảnh Quyết - 43.01.104.146

Trương Chí Lâm - 43.01.104.089

Nguyễn Tấn Tài - 43.01.104.152

Lâm Kiết Tường - 43.01.104.202

Lê Văn Tiến - 43.01.104.180

Nguyễn Long Hồ - 43.01.104.053

*Thành phố Hồ Chí Minh – Năm 2019*

🙣 Mục lục 🙡

[*I. Tìm hiểu về NoSQL và MongoDB:* 3](#_Toc25181675)

[1. NoSQL : 3](#_Toc25181676)

[1.1. NoSQL là gì? 3](#_Toc25181677)

[1.2. NoSQL database là gì? 4](#_Toc25181678)

[1.3. Tại sao lại cần phải có NoSQL. 4](#_Toc25181679)

[2. MongoDB : 5](#_Toc25181680)

[2.1 Ưu điểm MongoDB : 6](#_Toc25181681)

[2.2 Nhược điểm MongoDB : 7](#_Toc25181682)

[2.3 Khi nào nên sử dụng MongoDB : 7](#_Toc25181683)

[2.4 Lịch sử phát triển MongoDB. 8](#_Toc25181684)

[*II. Các thuật ngữ hay sử dụng trong MongoDB:* 8](#_Toc25181685)

[*III. Đăng kí để sử dụng MongoDB Cloud:* 9](#_Toc25181686)

[1. Tạo tài khoản MongoDB và Login 9](#_Toc25181687)

[10](#_Toc25181689)

[2. Tạo Project MongoDB Cloud: 10](#_Toc25181690)

[16](#_Toc25181692)

[*IV. Sử dụng kết nối và truy vấn dữ liệu MongoDB trên NodeJS :* 16](#_Toc25181693)

[1. Connect tới MongoDB Cloud: 16](#_Toc25181694)

[17](#_Toc25181695)

[2. Lệnh dùng trong collection: 17](#_Toc25181696)

[*V. Phần Demo sản phẩm:* 20](#_Toc25181697)

[*1.* Tạo môi trường làm việc : 20](#_Toc25181698)

[22](#_Toc25181703)

[2. Giới thiệu sản phẩm Demo 24](#_Toc25181709)

[*VI. Tổng kết và đánh giá:* 26](#_Toc25181710)

[1. Tổng kết 26](#_Toc25181711)

[2. Đánh giá 26](#_Toc25181712)

[TÀI LIỆU KHAM KHẢO 27](#_Toc25181713)

I. Tìm hiểu về NoSQL và MongoDB:

- Trong những năm gần đây, với sự ra đời và phát triển mạnh mẽ của NoSQL thì MongoDB cũng đang nhận được nhiều sự chú ý trong cộng đồng công nghệ. Điểm mạnh của NoSQL nói chung và MongoDB nói riêng đó là tính linh hoạt trong việc cấu trúc dữ liệu do đó nó giúp đáp ứng tốt với những thay đổi hay việc mở rộng cơ sở dữ liệu. Nếu bạn là người mới nghiên cứu về NoSQL thì đây là một cơ sở dữ liệu đầu tiên nên tiếp cận bởi nó dễ dàng cài đặt và sử dụng để tiếp cận với những ý tưởng lưu trữ mới.

1. NoSQL :

1.1. NoSQL là gì?

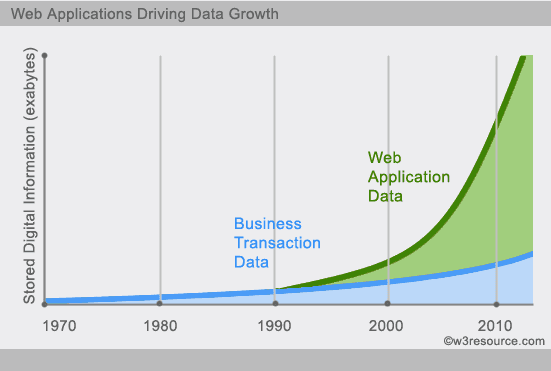
* NoSQL là 1 dạng CSDL mã nguồn mở và được viết tắt bởi: None-Relational SQL hay có nơi thường gọi là Not-Only SQL.
* NoSQL được phát triển trên Javascript Framework với kiểu dữ liệu là JSON và dạng dữ liệu theo kiểu key và value.
* NoSQL ra đời như là 1 mảnh vá cho những khuyết điểm và thiếu xót cũng như hạn chế của mô hình dữ liệu quan hệ RDBMS (Relational Database Management System - Hệ quản trị cơ sở dữ liệu quan hệ) về tốc độ, tính năng, khả năng mở rộng,...
* Với NoSQL bạn có thể mở rộng dữ liệu mà không lo tới những việc như tạo khóa ngoại, khóa chính, kiểm tra ràng buộc .v.v ...
* NoSQL bỏ qua tính toàn vẹn của dữ liệu và transaction để đổi lấy hiệu suất nhanh và khả năng mở rộng.
* NoSQL được sử dụng ở rất nhiều công ty, tập đoàn lớn, ví dụ như FaceBook sử dụng Cassandra do FaceBook phát triển, Google phát triển và sử dụng BigTable,...

1.2. NoSQL database là gì?

* Thuật ngữ NoSQL được giới thiệu lần đầu vào năm 1998 sử dụng làm tên gọi chung cho các lightweight open source relational database (cơ sở dữ liệu quan hệ nguồn mở nhỏ) nhưng không sử dụng SQL cho truy vấn. Vào năm 2009, Eric Evans, nhân viên của Rackspace giới thiệu lại thuật ngữ NoSQL trong một hội thảo về cơ sở dữ liệu nguồn mở phân tán. Thuật ngữ NoSQL đánh dấu bước phát triển của thế hệ database mới: distributed (phân tán) + non-relational (không ràng buộc). Đây là 2 đặc tính quan trọng nhất.

1.3. Tại sao lại cần phải có NoSQL.

* Sở dĩ người ta phát triển NoSQL xuất phát từ yêu cầu cần những database có khả năng lưu trữ dữ liệu với lượng cực lớn, truy vấn dữ liệu với tốc độ cao mà không đòi hỏi quá nhiều về năng lực phần cứng cũng như tài nguyên hệ thống và tăng khả năng chịu lỗi.
* Đây là những vấn đề mà các relational database không thể giải quyết được.
* Lượng dữ liệu mà các hệ thống cần phải xử lý giờ đây ngày 1 lớn. Ví dụ như Google, Facebook phải lưu trữ và xử lý một lượng dữ liệu cực lớn mỗi ngày.



Biểu đồ lưu trữ dữ liệu tăng theo thời gian

2. MongoDB :

- MongoDB (bắt nguồn từ “humongous”) là một hệ cơ sở dữ liệu NoSQL mã nguồn mở được thiết kế theo kiểu hướng đối tượng trong đó các bảng được cấu trúc một cách linh hoạt cho phép các dữ liệu lưu trên bảng không cần phải tuân theo một dạng cấu trúc nhất định nào. Chính do cấu trúc linh hoạt này nên MongoDB có thể đượng dùng để lưu trữ các dữ liệu có cấu trúc phức tạp và đa dạng và không cố định (hay còn gọi là Big Data)

[](https://bigsonata.files.wordpress.com/2014/06/logo.png).

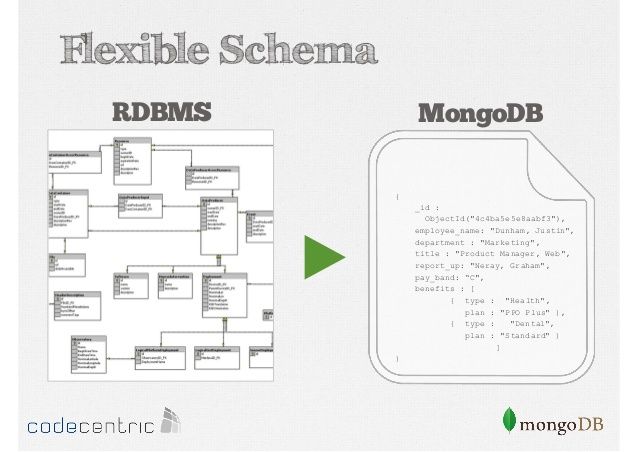
Hình 1.2.1: Logo của MongoDB

**- MongoDB** đứng ở tốp đầu trong danh sách các Hệ quản trị cơ sở dữ liệu NoSQL. Thay cho việc lưu trữ dữ liệu vào các bảng có quan hệ với nhau như truyền thống, MongoDB lưu các dữ liệu cấu trúc dưới dạng giống với JSON(JavaScript Object Notation) và gọi tên là BSON

- MongoDB có hai phiên bản Community (miễn phí) và Enterprise (trả phí). Mã nguồn, quản lý phiên bản đặt tại Github [1], phát hành theo giấy phép phần mềm mã nguồn mở GNU Affero General Public License (AGPL) hoặc Apache License tùy theo phiên bản. MongoDB được viết bằng ngôn ngữ lập trình C++ (80,1%), JavaScript (16,5%), Python (2,8%) và các ngôn ngữ lập trình khác (0,6%).  MongoDB là hệ quản trị cơ sở dữ liệu chạy trên nhiều hệ điều hành (Linux, Windows, MacOS), trên kiến trúc máy tính 32bit hoặc 64 bit (tuy nhiên không khuyến khích chạy trên máy tính 32bit do kích thước tập tin sẽ bị giới hạn bởi khả năng đánh địa chỉ ô nhớ của hệ điều hành).

2.1 Ưu điểm MongoDB :

* **Dữ liệu linh hoạt - MongoDB là document database, dữ liệu lưu dưới dạng JSON, không bị bó buộc về số lượng field, kiểu dữ liệu…**
* **Ngôn ngữ truy vấn phong phú -** MongoDB có sẵn các method để thực hiện create, read, update, delete dữ liệu(CRUD). Không có các fuction joins phức tạp như MySQL.
* Dữ liệu trong MongoDB không có sự ràng buộc lẫn nhau như trong Relational Database Management System (RDBMS), khi insert, xóa hay update nó không cần phải mất thời gian kiểm tra xem có thỏa mãn các bảng liên quan .
* Dữ liệu trong MongoDB được đánh chỉ mục(index) nên truy vấn nhanh gọn.
* Dễ dàng mở rộng hệ thống do có cluster là cụm các node chứa dữ liệu giao tiếp với nhau.
* Hỗ trợ replica set nhằm đảm bảo việc sao lưu và khôi phục dữ liệu.
* Không cần thiết conversion/mapping (chuyển đổi / ánh xạ) các đối tượng ứng dụng đến các đối tượng cơ sở dữ liệu.
* Sử dụng bộ nhớ trong để lưu trữ, cho phép truy cập dữ liệu nhanh hơn.
* Sử dụng MongoDB Cloud cho phép nhiều người dùng có thể cùng nhau xây dựng dự án
* Ít schema hơn: Vì schema được sinh ra là để nhóm các đối tượng vào 1 cụm, dễ quản lý. Ví dụ như tạo 1 schema tên là Students chẳng hạn thì chỉ có những gì liên quan đến student thì mới được cho vào schema này. Trong khi đó trong MongoDB thì chỉ 1 collection ta có thể chứa nhiều document khác nhau. Với mỗi document thì số trường, nội dung, kích thước lại có thể khác nhau.



Schema của RDBMS và MongoDB

2.2 Nhược điểm MongoDB :

* MongoDB không có các tính chất ràng buộc như trong RDBMS nên dễ bị làm sai dữ liệu
* Không hỗ trợ join giống như RDBMS nên khi viết function join trong code phải tự làm khiến cho tốc độ truy vấn bị giảm.
* Sử dụng nhiều bộ nhớ: do dữ liệu lưu dưới dạng key-value, các collection chỉ khác về value do đó key sẽ bị lặp lại. Không hỗ trợ join nên sẽ bị dư thừa dữ liệu.
* Bị giới hạn kích thước bản ghi: mỗi document không được có kích thước lớn hơn 16Mb.
* Dùng MongoDB Cloud thì lúc chạy ở local cần có kết nối Internet để truy cập tới DB Cloud

2.3 Khi nào nên sử dụng MongoDB :

* Danh mục sản phẩm thương mại điện tử
* Blog và quản lý các nội dung.
* Hệ thống realtime (thời gian thực) yêu cầu phản hồi nhanh.
* Lưu trữ dữ liệu không gian địa lý.
* Các trang web mạng di động và mạng xã hội.
* Các hệ thống bigdata với yêu cầu truy vấn nhanh.
* Làm search engine

2.4 Lịch sử phát triển MongoDB.

- Dự án MongoDB được bắt đầu triển khai vào tháng 10 năm 2007 bởi công ty phần mềm **10gen** trong khi công ty này đang xây dựng một nền tảng như là dịch vụ (Platform as a Service) giống như Google App Engine. Năm 2009, công ty chuyển sang mô hình phát triển nguồn mở, với công ty cung cấp hỗ trợ thương mại và các dịch vụ khác. Năm 2013, 10gen đổi tên thành MongoDB Inc. Nó được một số tổ chức sử dụng trong thực tế như:

* Caigslist : Công ty làm việc trong lịch vực môi giới quảng cáo trên các website khác (giống adMicro của Việt Nam). MongoDB giúp cho công ty này quản lý hàng tỉ các bản ghi quảng cáo thuận tiện và nhanh chóng.
* Foursquare là một mạng xã hội gắn các thông tin địa lý. Công ty này cần lưu dữ liệu của rất rất nhiều vị trí của các địa điểm như quán cafe, nhà hàng, điểm giải trí, lịch sử, … và ghi lại những nơi mà người sử dụng đã đi qua.
* CERN : Trung tâm nghiên cứu năng lượng nguyên tử của Châu Âu, sử dụng MongoDB để lưu trữ lại các kết quả, dữ liệu thí nghiệm của mình. Đây là một lượng dữ liệu khổng lồ sẽ dùng để sử dụng trong tương lai.

II. Các thuật ngữ hay sử dụng trong MongoDB:

**\_id**: Là trường bắt buộc có trong mỗi document. Trường \_id đại diện cho một giá trị duy nhất trong document MongoDB. Trường \_id cũng có thể được hiểu là khóa chính trong document. Nếu bạn thêm mới một document thì MongoDB sẽ tự động sinh ra một \_id đại diện cho document đó và là duy nhất trong cơ sở dữ liệu MongoDB.

**Collection**: Là nhóm của nhiều document trong MongoDB. Collection có thể được hiểu là một bảng tương ứng trong cơ sở dữ liệu RDBMS (Relational Database Management System). Collection nằm trong một cơ sở dữ liệu duy nhất. Các collection không phải định nghĩa các cột, các hàng hay kiểu dữ liệu trước.

**Cursor**: Đây là một con trỏ đến tập kết quả của một truy vấn. Máy khách có thể lặp qua một con trỏ để lấy kết quả.

**Database**: Nơi chứa các Collection, giống với cơ sở dữ liệu RDMS chúng chứa các bảng. Mỗi Database có một tập tin riêng lưu trữ trên bộ nhớ vật lý. Một mấy chủ MongoDB có thể chứa nhiều Database.

**Document**: Một bản ghi thuộc một Collection thì được gọi là một Document. Các Document lần lượt bao gồm các trường tên và giá trị.

**Field**: Là một cặp name – value trong một document. Một document có thể có không hoặc nhiều trường. Các trường giống các cột ở cơ sở dữ liệu quan hệ.

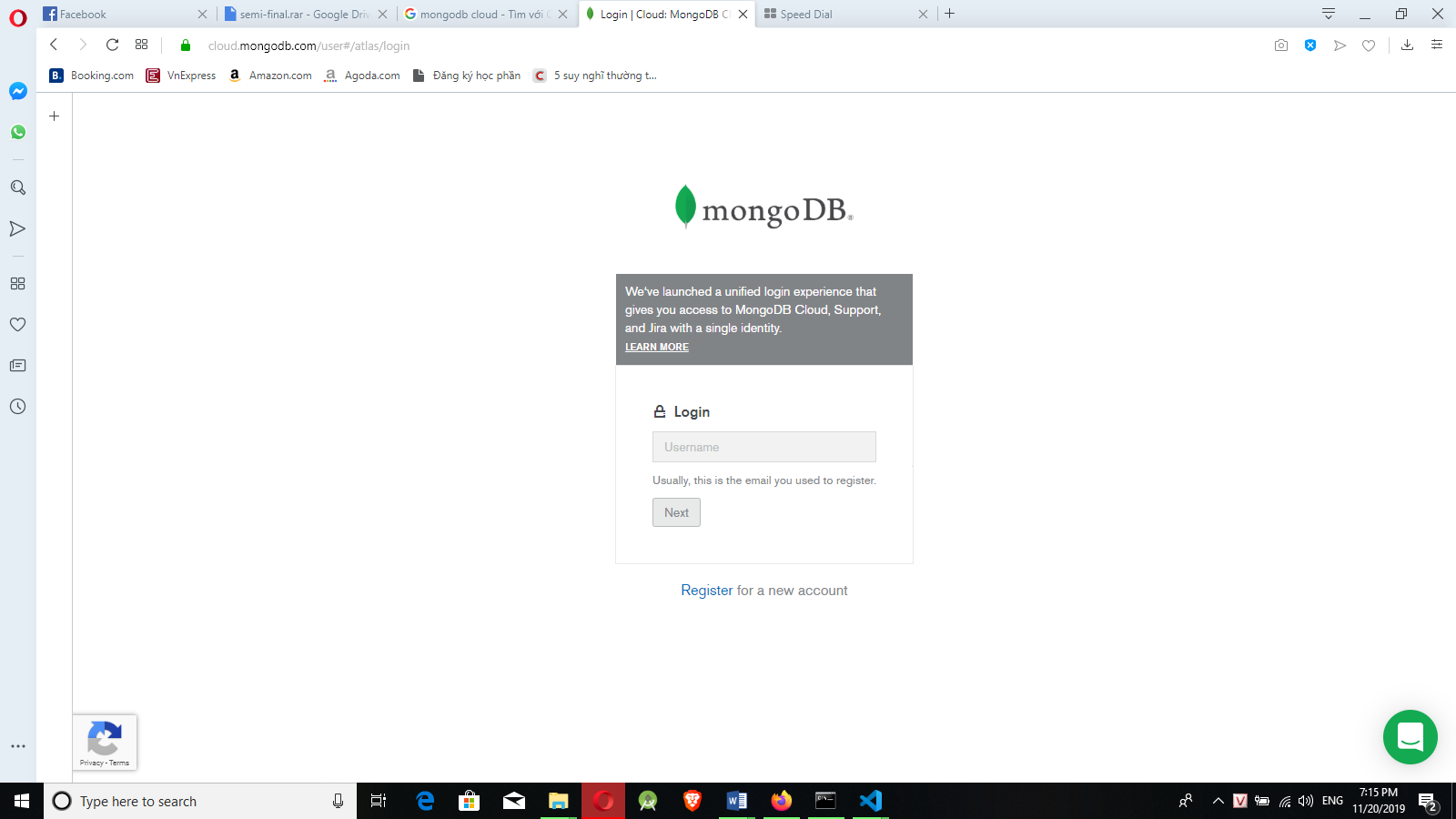
**JSON**: Viết tắt của JavaScript Object Notation. Con người có thể đọc được ở định dạng văn bản đơn giản thể hiện cho các dữ liệu có cấu trúc. Hiện tại JSON đang hỗ trợ rất nhiều ngôn ngữ lập trình.

**Index**: Là những cấu trúc dữ liệu đặc biệt, dùng để chứa một phần nhỏ của các tập dữ liệu một cách dễ dàng để quét. Chỉ số lưu trữ giá trị của một fields cụ thể hoặc thiết lập các fields, sắp xếp theo giá trị của các fields này. Index hỗ trợ độ phân tích một cách hiệu quả các truy vấn. Nếu không có chỉ mục, MongoDB sẽ phải quét tất cả các documents của collection để chọn ra những document phù hợp với câu truy vấn. Quá trình quét này là không hiệu quả và yêu cầu MongoDB để xử lý một khối lượng lớn dữ liệu.

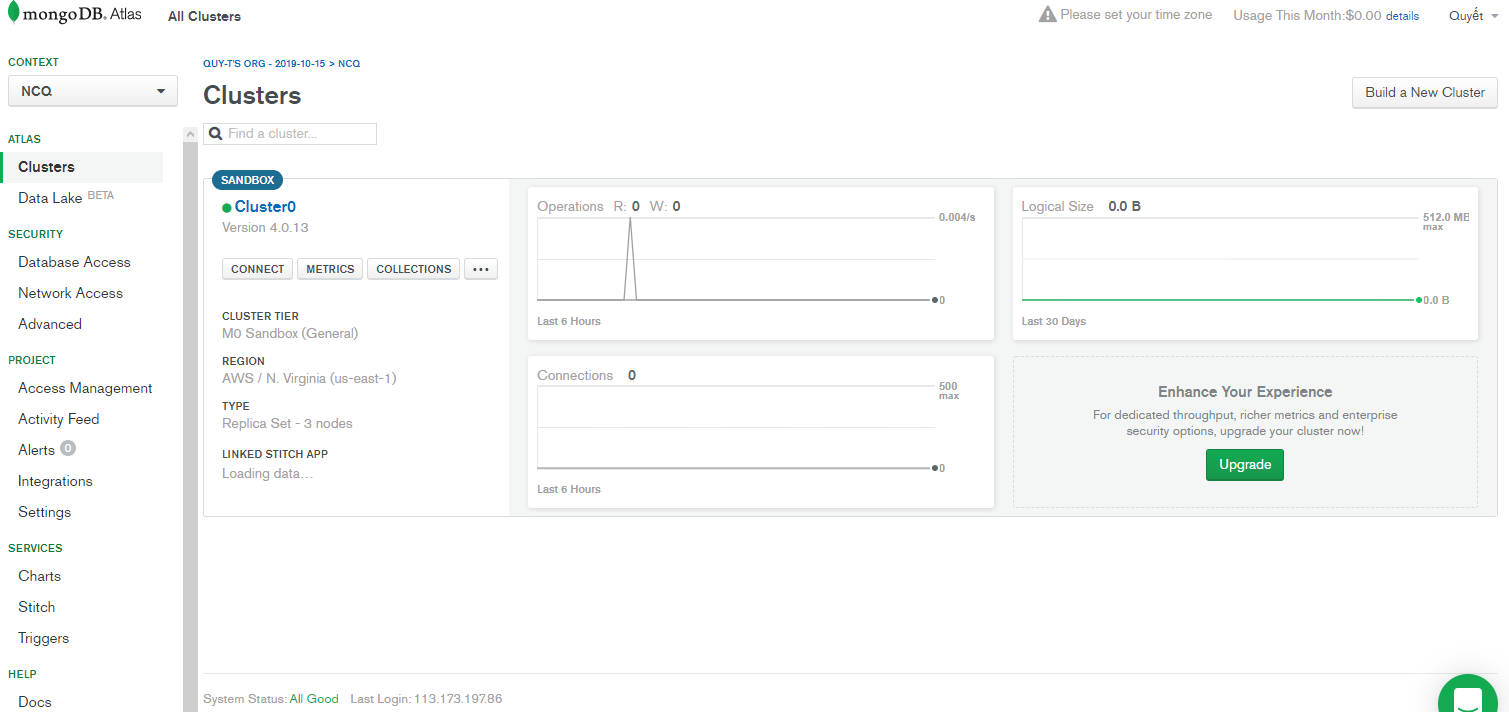
III. Đăng kí để sử dụng MongoDB Cloud:

1. Tạo tài khoản MongoDB và Login

Truy cập : https://cloud.mongodb.com/user

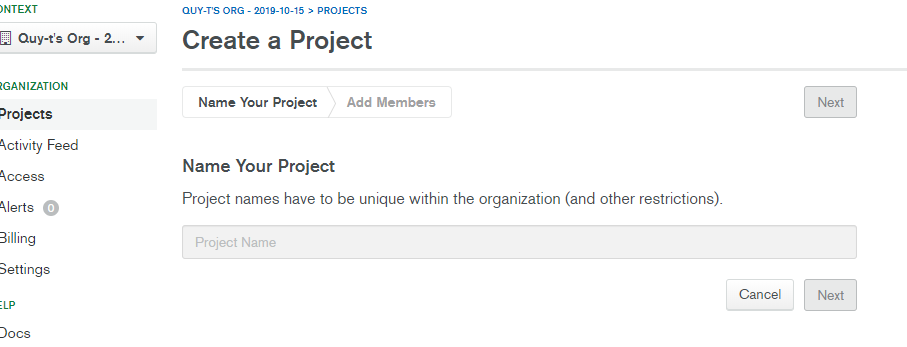


Hình 3.1.1 : Login MongoDB Cloud

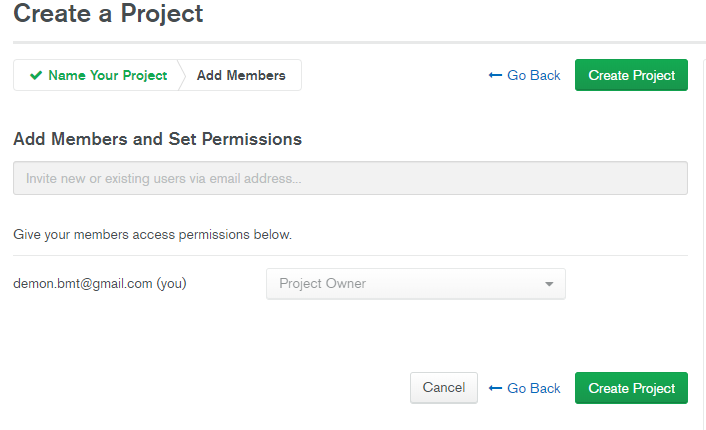


Hình 3.1.2 : Giao diện MongoDB Cloud

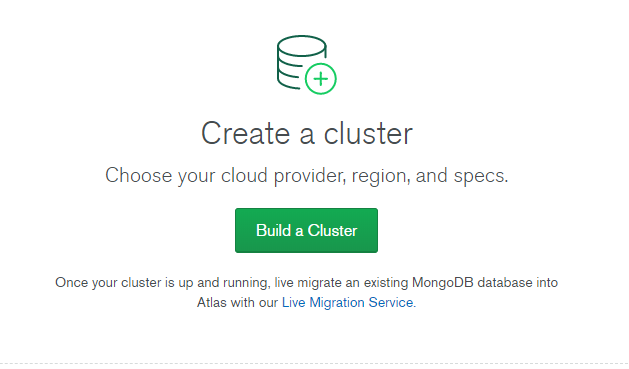
2. Tạo Project MongoDB Cloud:



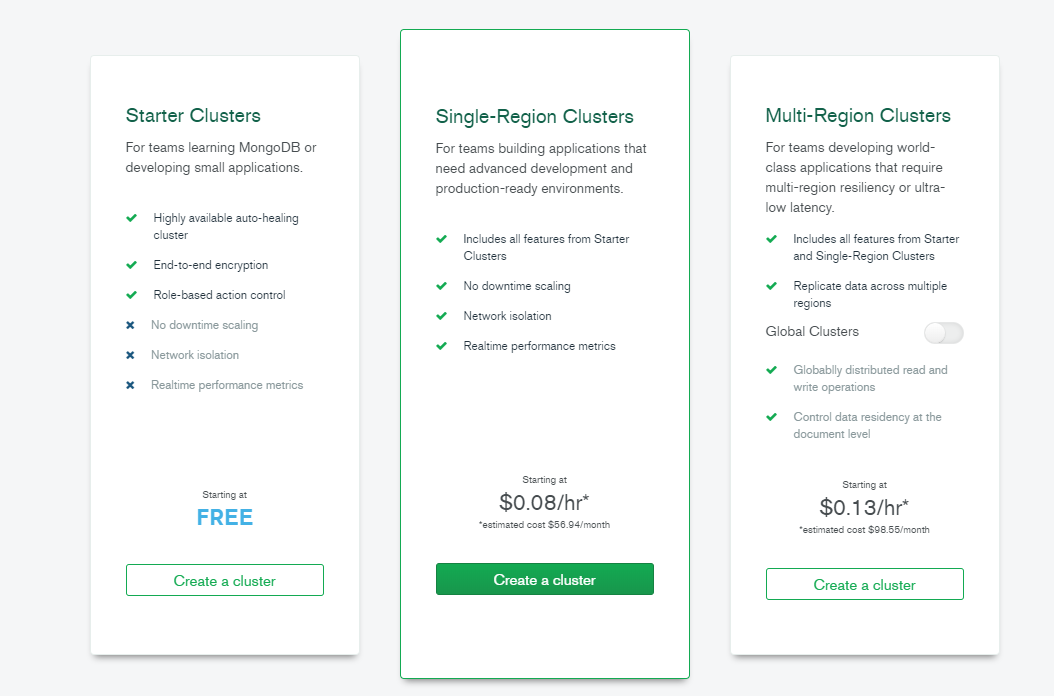
Hình 3.2.1 : Đặt tên Project



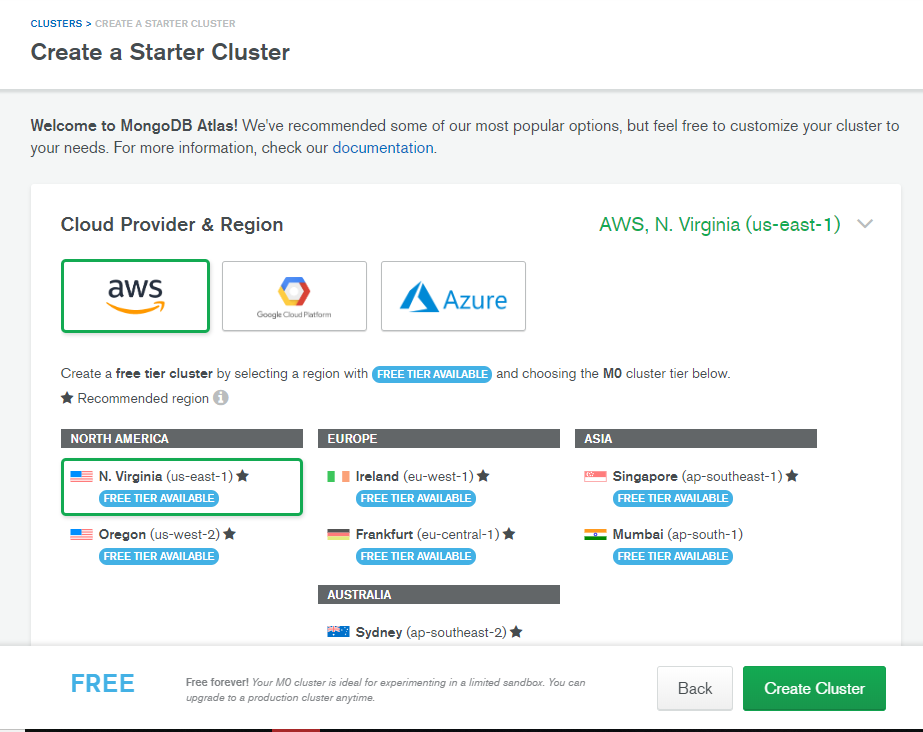
Hình 3.2.2 : Thêm thành viên cùng quản lý ( Có thể bỏ qua )



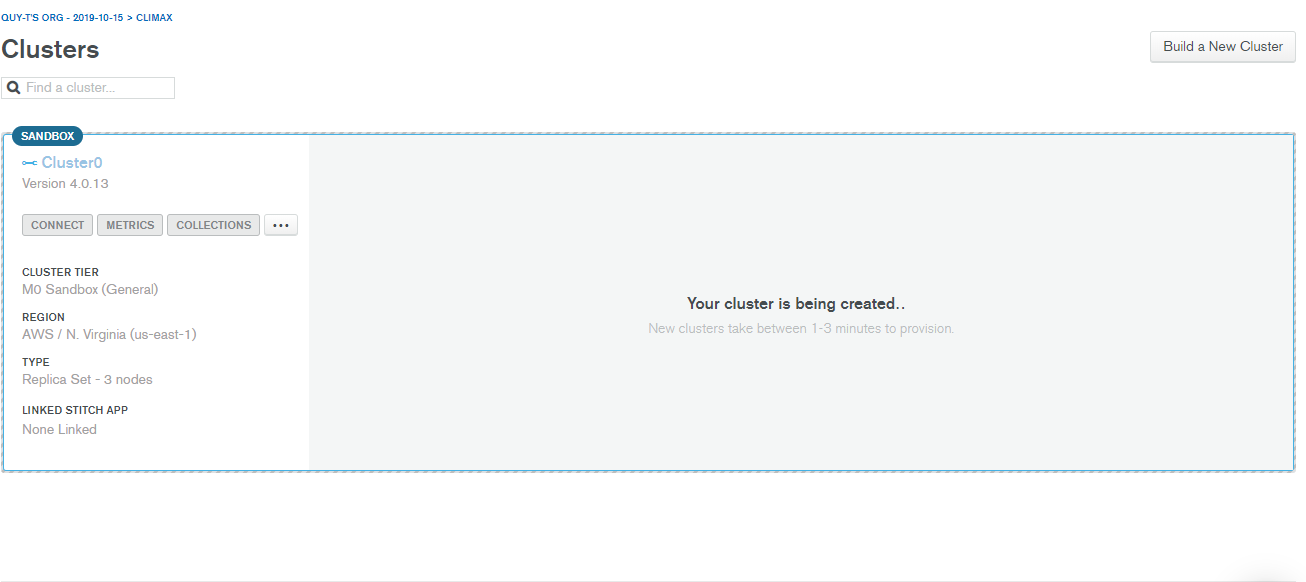
Hình 3.2.3 : Thêm Cluster



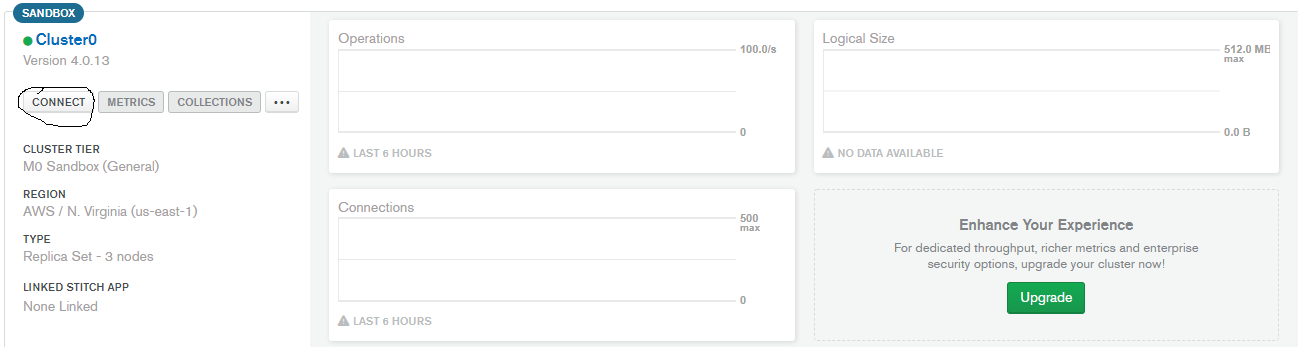
Hình 3.2.4 : Chọn gói sử dụng Clusters



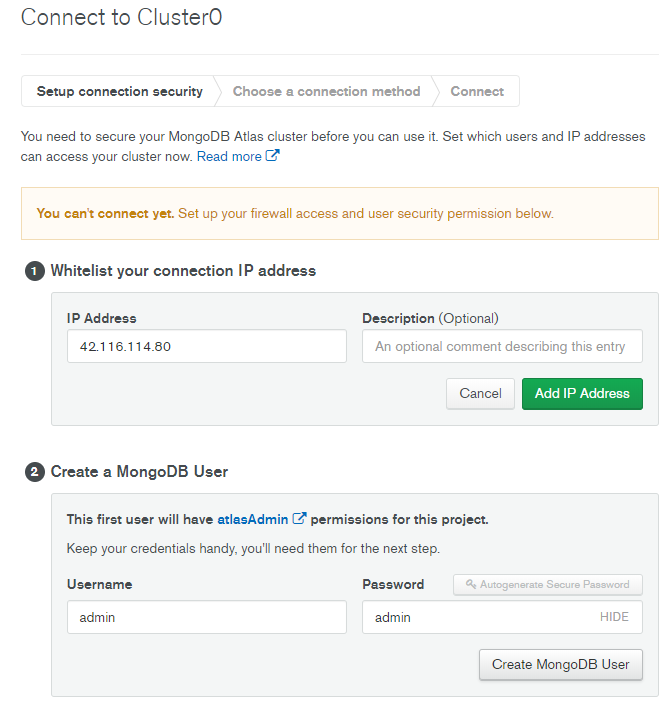
Hình 3.2.5 : Lựa chọn server cloud



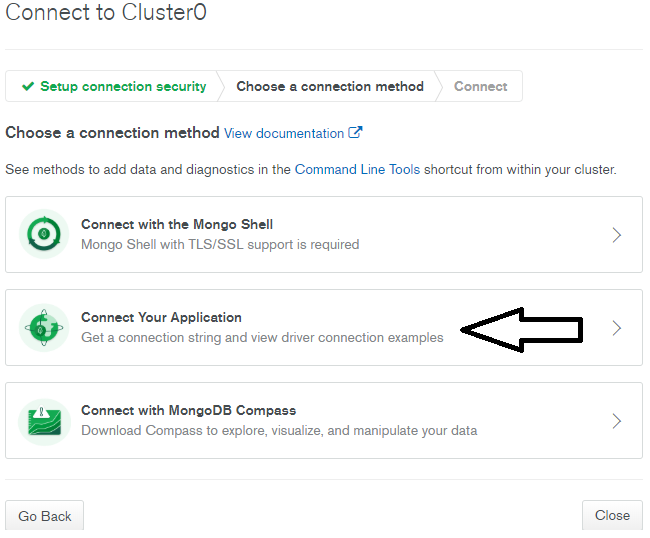
Hình 3.2.6 : Chờ khởi tạo Clusters (1 – 3p)



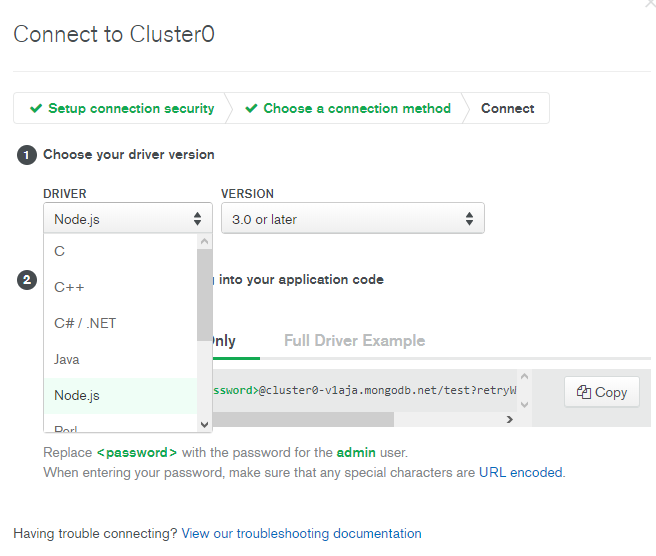
Hình 3.2.7 : Thực hiện Connect



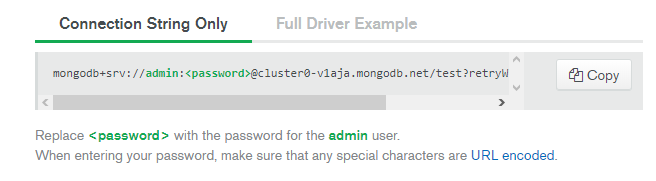
Hình 3.2.8 : Thực hiện Setup Connection Security



Hình 3.2.9 : Choose A Connection Method



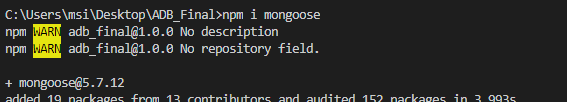
Hình 3.2.10 : Chọn ngôn ngữ thực hiện



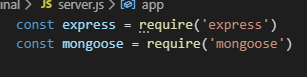
Hình 3.2.10 : Chuỗi kết nối tới APP

IV. Sử dụng kết nối và truy vấn dữ liệu MongoDB trên NodeJS :

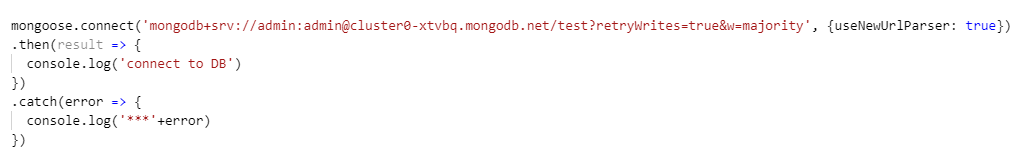
1. Connect tới MongoDB Cloud:

**

Hình 5.1.1.1 : Thêm Express Mongoose Cho NodeJS

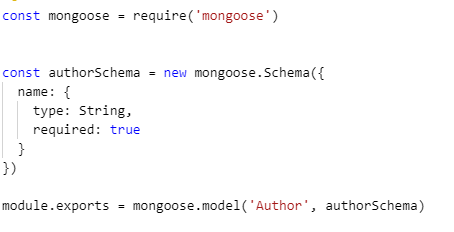


Hình 5.1.1.2 : Thêm Express Mongoose Cho NodeJS



Hình 5.1.2 : Dùng chuỗi kết nối được cung cấp để kết nối tới MongoDB Cloud

2. Lệnh dùng trong collection:

**

*Hình 4.2.1 : Cú pháp tạo Schema bằng mongoose*

**

*Hình 4.2.2 : Khai báo khi sử dụng model*

**

*Hình 4.2.3 : Cú pháp them mới model*

**

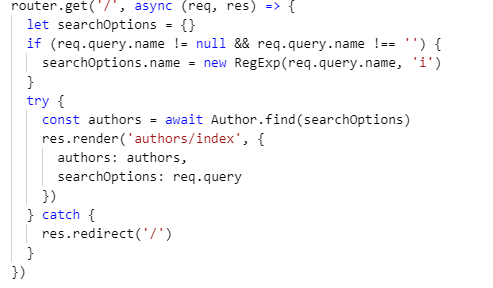
*Hình 4.2.4 : Cú pháp edit model*

**

*Hình 4.2.5 : Cú pháp update model*

**

*Hình 4.2.6 : Cú pháp xóa model*

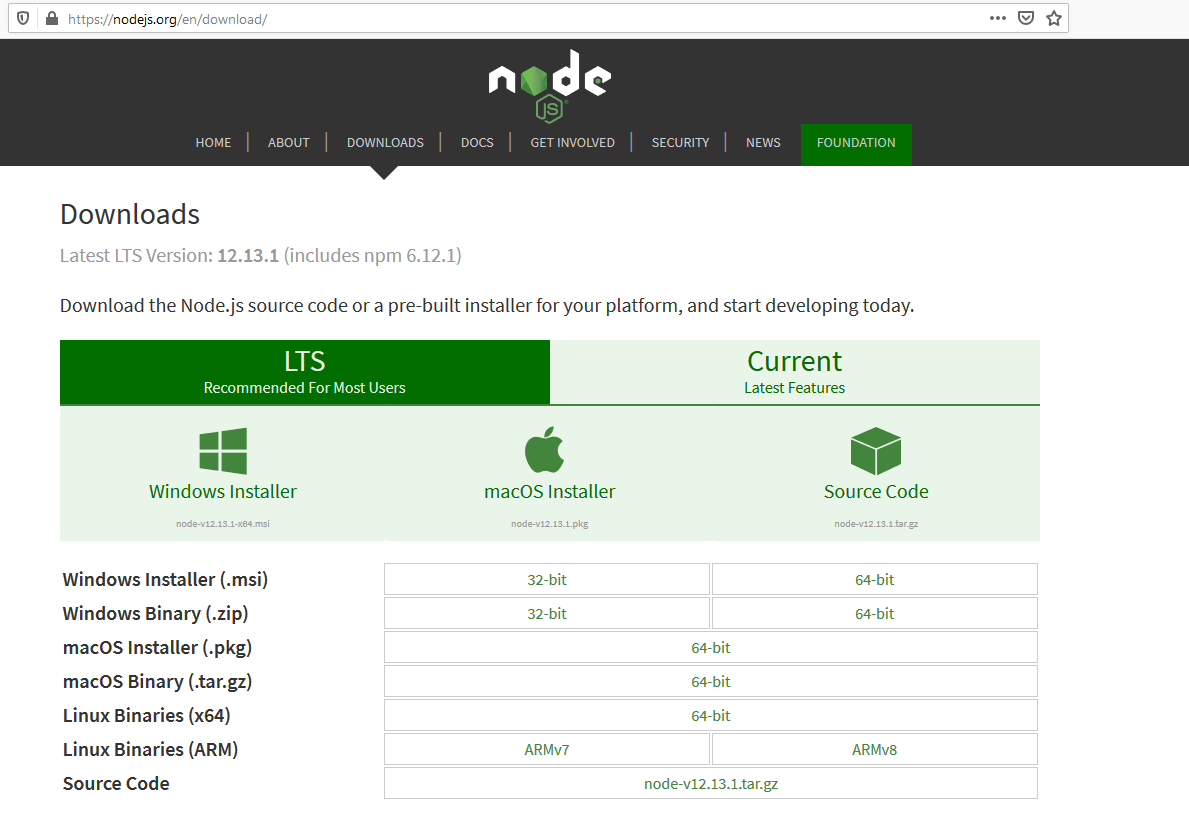
**

*Hình 4.2.7 : Cú pháp tìm kiếm*

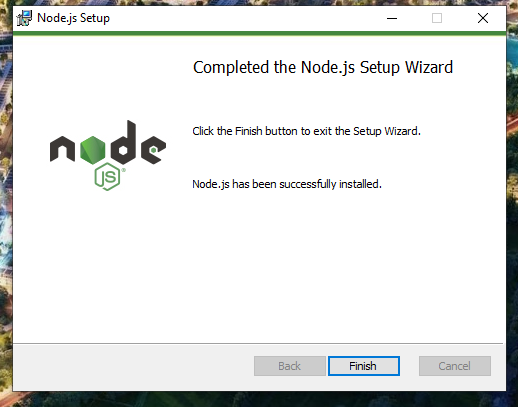
V. Phần Demo sản phẩm:

*Sản phẩm demo của nhóm sử dụng NoSQL MongoDB Cloud*

1. Tạo môi trường làm việc :

**

Hình 7.1.1 : Download NodeJS



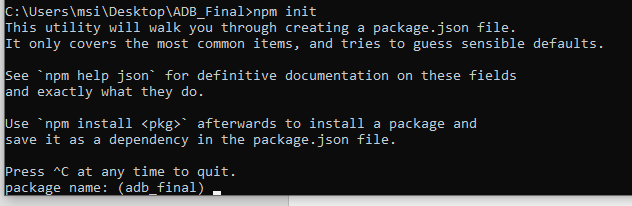
Hình 7.1.2 : Cài đặt vào máy



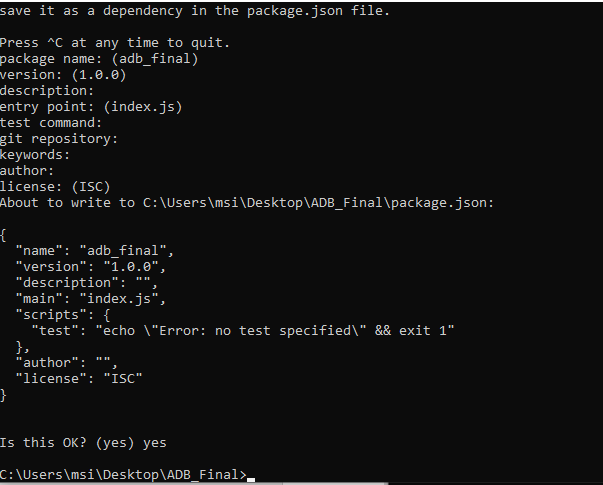
Hình 7.1.3 : Kiếm tra version nodeJS : Nếu hiện thì tức là cài thành công

**

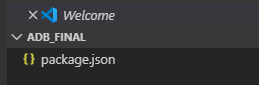
Hình 7.1.4.1 : Khởi tạo project



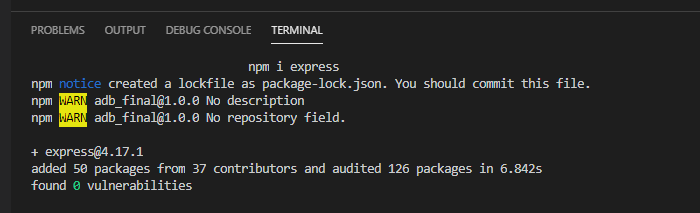
Hình 7.1.4.2 : Khởi tạo điền các thông tin project



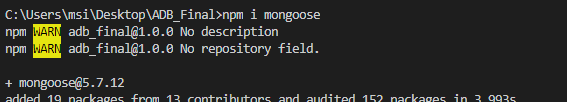
Hình 7.1.4.3 : Xác nhân khởi tạo project

**

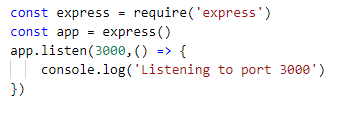
Hình 7.1.5 : Project JS được khởi tạo

**

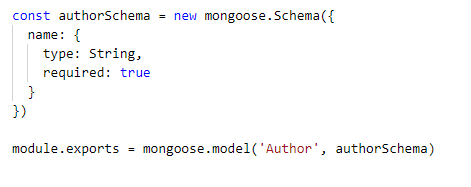
Hình 7.1.6.1 : Import các Package cần thiết

**

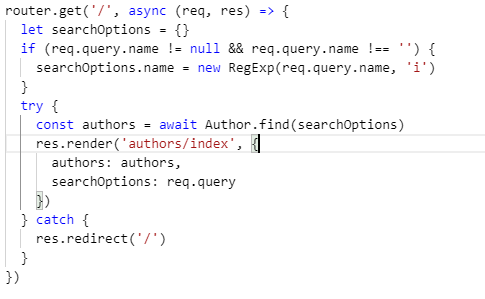
Hình 7.1.6.2 : Import các Package cần thiết



Hình 7.1.7 : Khai báo và bắt đầu code

**

Hình 7.1.8 : Khởi tạo Schema (Model)

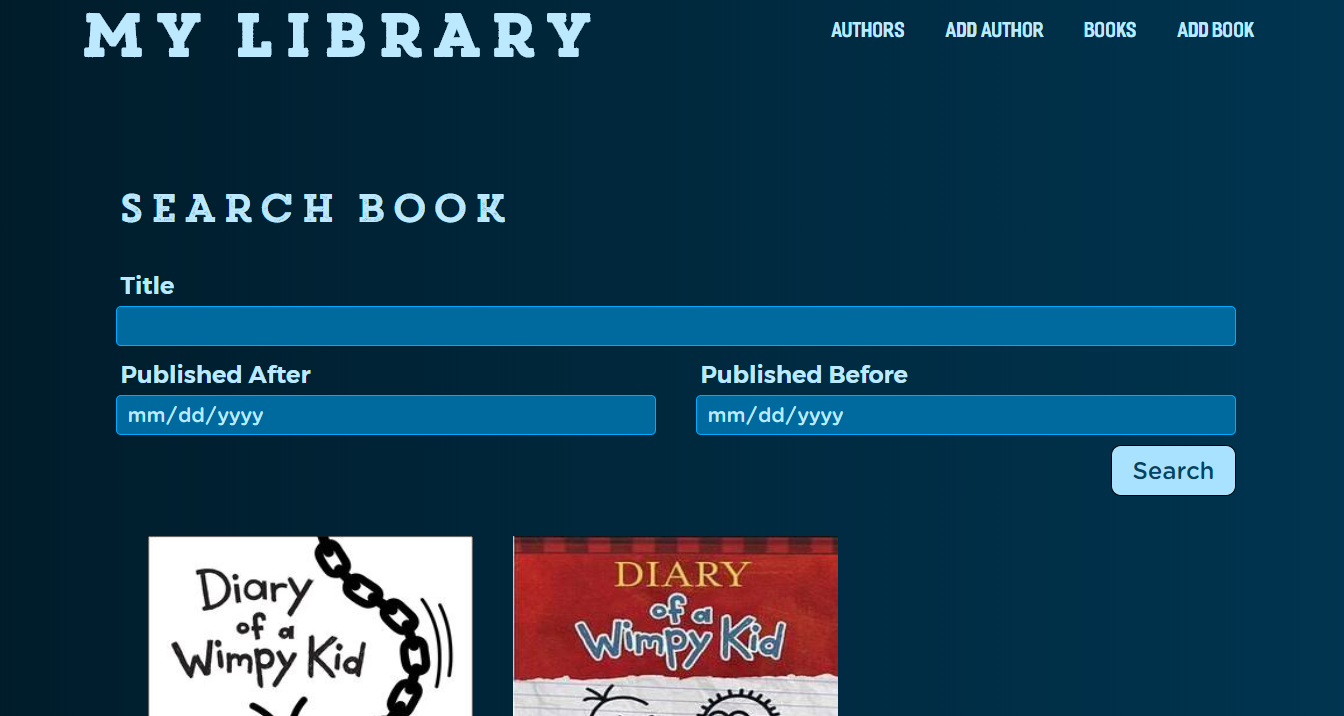
**

Hình 7.1.9 : Chia nhánh xử lí

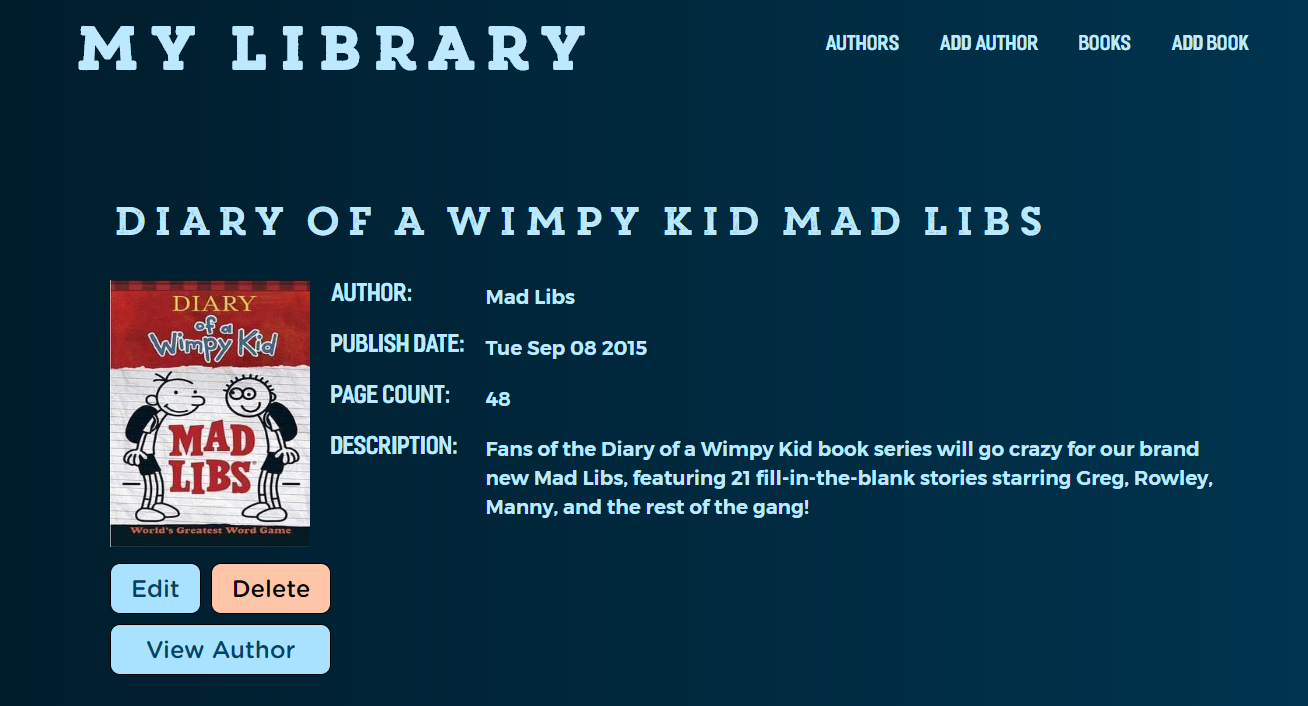
1. Giới thiệu sản phẩm Demo



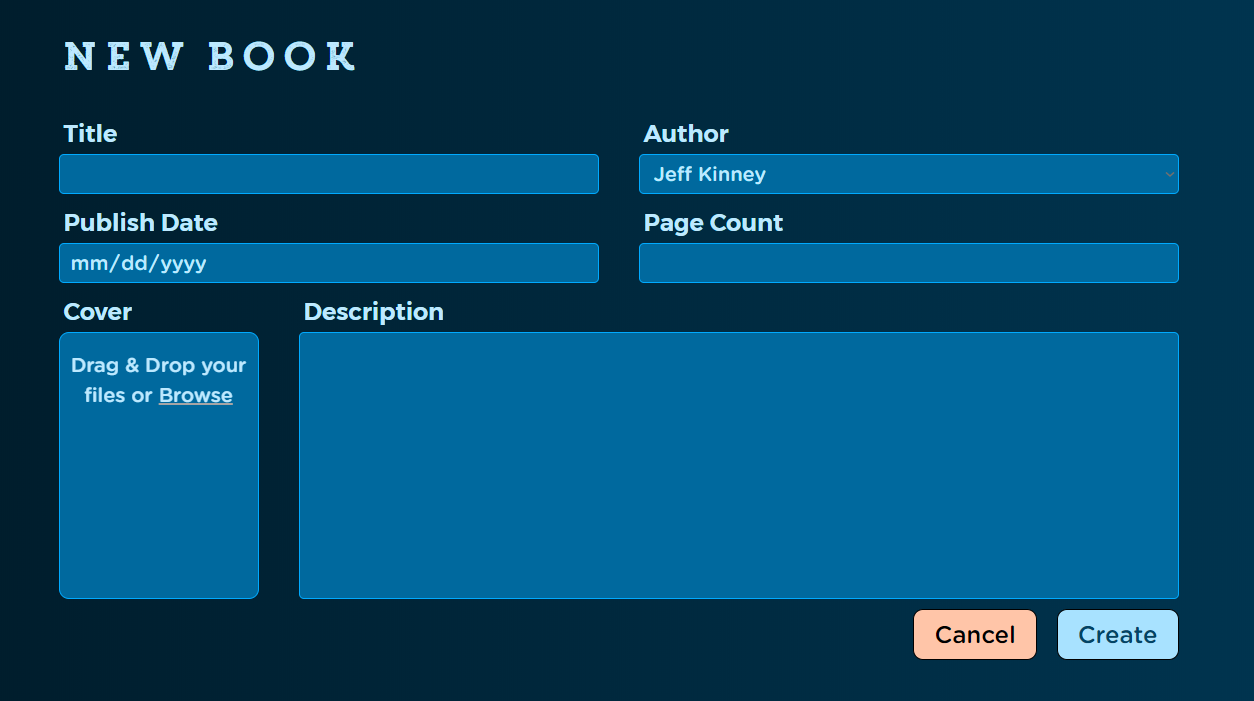
Hình 6.2.1 : Trang chủ DEMO



Hình 6.2.2 : Trang tìm kiếm sách



Hình 6.2.3 : Trang xem thông tin chi tiết sách



Hình 6.2.4 : Trang thêm sách mới

VI. Tổng kết và đánh giá:

1. Tổng kết

* MongoDB không sử dụng mô hình quan hệ ràng buộc(non-relational)
* Mô hình Lưu trữ dạng {“key : value”}
* MongoDB là 1 CSDL phân tán lưu chữ dữ liệu trên nhiều server khác nhau, đa dạng hóa dữ liệu, là dạng dữ liệu lớn( BigData).
* Độ chịu tải cao ,chịu lỗi cao
* Có khả năng nâng cấp theo chiều ngang, dễ dàng bảo trì, thay đổi,sửa chữa , phát triển

1. Đánh giá

* Chính vì những đặc điểm đó ta nên dùng nó cho các hệ thống, ứng dụng đòi hỏi lượng người dùng và dữ liệu lưu trữ đa dạng, lớn đến cực lớn.Ví dụ như mạng xã hội(facebook, Instagram, twitter…), các trang lưu trữ dữ liệu lớn như google, github, …..

TÀI LIỆU KHAM KHẢO

[1]. **Ebook NoSQL** - Nhữ Đình Thuận

[2]. **Kristina Chodorow - Micheal Dirolf, MongoDB**: The Definitive Guide, O'reilly, 2010

[3]. Eeclo Plugge, Peter Membrey and Tim Hawkins, **The Definitive Guide to MongoDB the NoSQL Database for Cloud and Desktop Computing.**

[4]. **NoSQL resources** - *http://nosql-database.org*

[5]. **MongoDB wikipedia** *- https://en.wikipedia.org/wiki/MongoDB*

[6]. *https://www.mongodb.com*

[7]. *https://cloud.mongodb.com*

[8]. *http://kienthucweb.net/gioi-thieu-ve-co-so-du-lieu-mongodb.html*

[9]. *https://vietjack.com/mongodb/*

[10]. *https://freetuts.net/tong-quan-ve-mongodb-203.html*