

# NHẬP MÔN CÔNG NGHỆ PHẦN MỀM

## Project Proposal

Yêu cầu nhóm sinh viên hoàn thành tài liệu Project Proposal cho đề án đã được giao theo biểu mẫu đính kèm.



Bộ môn Công nghệ phần mềm  
Khoa Công nghệ thông tin  
Đại học Khoa học tự nhiên TP HCM

# MỤC LỤC

Các nội dung chính .....	1
1 Bảng đánh giá thành viên.....	2
2 Phát biểu bài toán sơ lược.....	3
3 Giải pháp đề xuất.....	3
4 Kế hoạch phát triển .....	4
5 Kế hoạch nhân sự & chi phí.....	16

# Project Proposal

## Các nội dung chính

Mục tiêu tài liệu tập trung vào các chủ đề:

- ✓ Tạo ra tài liệu Project Proposal.
- ✓ Hoàn chỉnh tài liệu Project Proposal với các nội dung:
  - Phát biểu bài toán sơ lược
  - Giải pháp đề xuất
  - Kế hoạch phát triển
  - Kế hoạch nhân sự & chi phí
- ✓ Đọc hiểu tài liệu Project Proposal.

# 1

## Bảng đánh giá thành viên

MSSV	Họ Tên	% đóng góp (tối đa 100%)	Chữ ký
22810211	Nguyễn Ba Hoài Nhựt	100%	

## 2 Phát biểu bài toán sơ lược

- Mục tiêu chính của dự án này là xây dựng một hệ thống crawler dữ liệu để thu thập thông tin về các hội nghị uy tín trên toàn cầu. Các thông tin này có thể bao gồm tên hội nghị, các mốc deadline, ngày diễn ra, địa điểm, chủ đề và các thông mở rộng khác như: danh sách diễn giả, các bài báo đã được chấp nhận, và thông tin liên quan đến đăng ký tham dự,...
- Khó khăn và thách thức:
  - Một số trang web có chống crawl hoặc bị giới hạn về tốc độ truy cập.
  - Dữ liệu trên các trang web có thể thay đổi thường xuyên, đòi hỏi cập nhật định kỳ.
  - Dữ liệu có thể không được trình bày trong định dạng dễ phân tích.

## 3 Giải pháp đề xuất

### 3.1 Phần mềm: web app cào hội nghị

#### 3.1.1. Danh sách các chức năng phần mềm

Nhu cầu	Yêu cầu
Là người dùng tôi muốn tìm kiếm thông tin các hội nghị	Tìm kiếm hội nghị
Là người dùng tôi muốn lọc hội nghị theo quốc gia/ thời gian sắp diễn ra.	Lọc theo quốc gia & sắp xếp theo thời gian
Là admin tôi muốn chỉnh sửa thời gian cào tự động hoặc trigger crawl thủ công	Cài đặt thời gian chạy tự động và crawl dữ liệu

### 3.2 Phần cứng

*Các yêu cầu về máy móc: Hầu hết các cấu hình hiện này đều hỗ trợ NodeJs*

## 4 Kế hoạch phát triển

### 4.1 Phân tích yêu cầu

- ❖ **Mục tiêu của đồ án:** Xây dựng một trang web **IT Conferences Crawler** để giúp người dùng tìm kiếm, theo dõi và đăng ký tham gia các hội thảo trong lĩnh vực công nghệ thông tin
- ❖ **Yêu cầu chức năng:**

Tìm kiếm hội thảo: người dùng có thể tìm kiếm hội thảo theo tên hội thảo, thời gian, chủ đề ,...v...v

Xem danh sách hội thảo: hiển thị danh sách hội thảo dựa trên tìm kiếm hoặc danh mục

Xử lý dữ liệu: Sau khi thu thập dữ liệu, cần xử lý và lọc thông tin để loại bỏ dữ liệu không cần thiết và đảm bảo độ chính xác của dữ liệu.

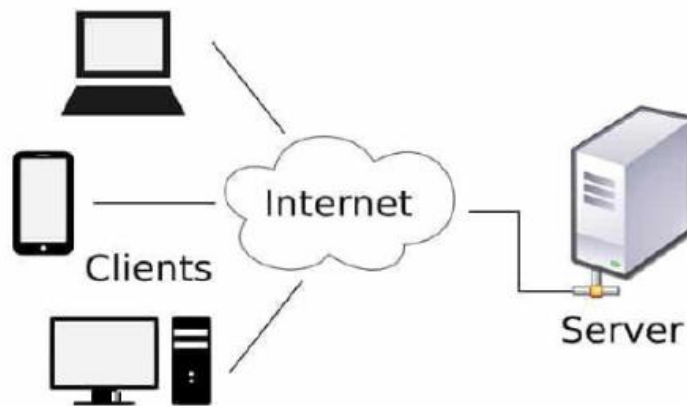
Lưu trữ dữ liệu: Dữ liệu thu thập được cần được lưu trữ trong một cơ sở dữ liệu hoặc tệp để có thể sử dụng sau này.

- ❖ **Yêu cầu phi chức năng:** Dữ liệu phải được cập nhật thường xuyên để đảm bảo tính chính xác, trình bày theo định dạng dễ phân tích
- ❖ **Yêu cầu kĩ thuật:**
  - Ngôn ngữ lập trình: Dùng ngôn ngữ lập trình để xây dựng website
  - Cơ sở dữ liệu: Lưu trữ thông tin về hội thảo
- ❖ **Yêu cầu giao diện:** xây dựng một giao diện người dùng để tương tác với dữ liệu thu thập được, thân thiện và dễ dùng đảm bảo cho việc người dùng dễ dàng tìm kiếm thông tin hội thảo
- ❖ **Yêu cầu kiểm thử:** Đảm bảo tích hợp các chức năng hoạt động ổn định
- ❖ **Yêu cầu quản lý đồ án:** Đảm bảo được triển khai đúng hạn
- ❖ Danh sách các trang web/ nguồn dữ liệu thu thập thông tin:  
<https://dev.events/tech?sorting=startDate>

### 4.2 Thiết kế phần mềm

**Mô tả kiến trúc hệ thống.**

Trang web được xây dựng dựa trên kiến trúc client - server, framework NextJs: Người dùng truy cập thông qua giao diện người dùng (client) - máy chủ (server) xử lý yêu cầu và trả về dữ liệu.



Các thành phần chính:

Front-end: Sử dụng công nghệ ReactJs để hiển thị giao diện cho người dùng (framework NextJs)

Back-end: Sử dụng ngôn ngữ lập trình Nodejs để xử lý logic và dữ liệu (framework NextJs)

Cơ sở dữ liệu: Lưu trữ thông tin về hội thảo và các dữ liệu liên quan khác với Google FireBase để sử dụng sau này.

Các dịch vụ bên ngoài: MongoDB để lưu trữ dữ liệu, Hosting Vercel

### Phân tích thiết kế hệ thống:

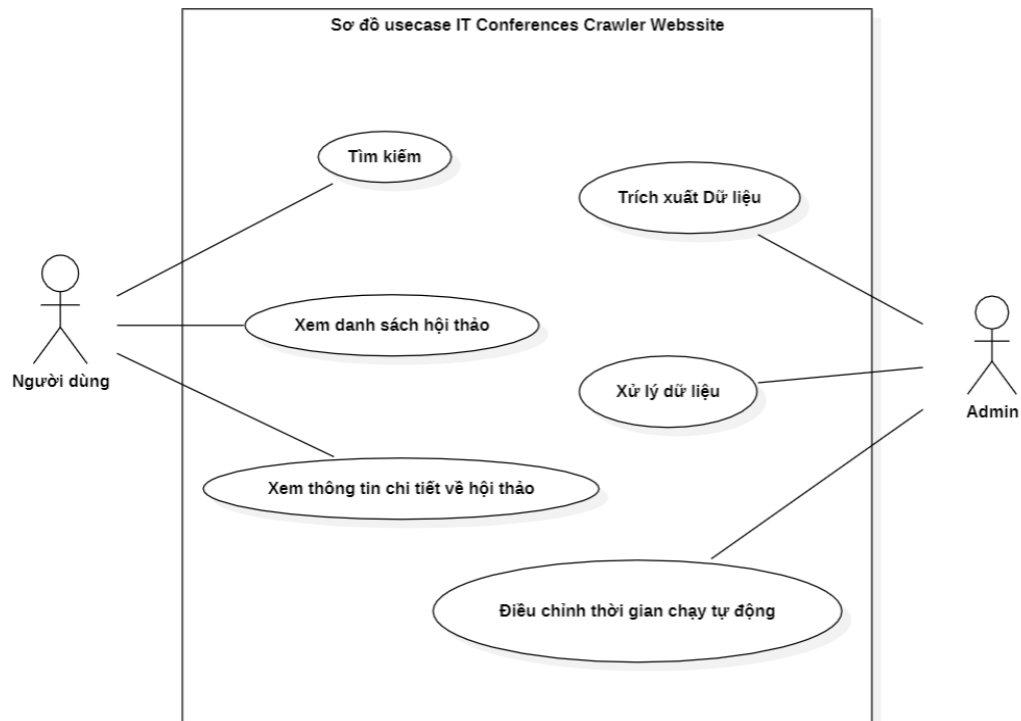
*Các tác nhân chính tham gia vào hệ thống:*

- **Người dùng:** Là người sử dụng trang web có nhu cầu tìm kiếm thông tin về các hội thảo, xem danh sách các hội thảo
- **Admin:** là người quản lý hệ thống, có quyền truy cập và thực hiện các chức năng quản lý (trích xuất dữ liệu, xử lý dữ liệu, điều chỉnh thời gian chạy tự động)

*Các UseCase chính tham gia hệ thống:*

- **Người dùng:**
  - Tìm kiếm hội thảo
  - Xem danh sách hội thảo
  - Xem thông tin chi tiết về hội thảo
- **Admin:**
  - Trích xuất dữ liệu
  - Xử lý dữ liệu
  - Điều chỉnh thời gian chạy tự động

*Lược đồ UseCase tổng quát:*



Đặc tả các UC từng nghiệp vụ bằng văn bản:

Đặc tả UC tìm kiếm

Tên UC	Tìm kiếm
Mô tả	UC bắt đầu khi guest có nhu cầu tìm kiếm hội nghị
Dòng cơ bản	<ol style="list-style-type: none"> <li>Guest nhập keyword tìm kiếm và bấm enter</li> <li>Hệ thống trả về kết quả tìm kiếm với từ khóa được nhập.</li> <li>Kết Thúc UC</li> </ol>
Dòng thay thế	Tại bước 2 nếu không có kết quả thì trả về blank



**Đặc tả UC xem danh sách hội thảo**

<b>Tên UC</b>	<b>Xem danh sách hội thảo</b>
<b>Mô tả</b>	UC bắt đầu khi người dùng có nhu cầu tìm kiếm hội nghị
<b>Dòng cơ bản</b>	<ol style="list-style-type: none"> <li>1. Người dùng truy cập website</li> <li>2. Hệ thống sẽ hiển thị danh sách hội nghị với tìm kiếm và lọc mặc định ( không lọc không tìm kiếm)</li> <li>3. Kết thúc UC</li> </ol>

**Đặc tả UC xem thông tin chi tiết hội thảo**

<b>Tên UC</b>	<b>Xem thông tin chi tiết hội thảo</b>
<b>Mô tả</b>	UC bắt đầu khi người dùng muốn tìm hiểu về 1 hội nghị cụ thể trong danh sách
<b>Dòng cơ bản</b>	<ol style="list-style-type: none"> <li>1. Người dùng thực hiện UC tìm kiếm hội nghị</li> <li>2. Người dùng chọn vào link hội nghị cụ thể</li> <li>3. Kết thúc UC</li> </ol>

**Đặc tả UC trích xuất dữ liệu**

<b>Tên UC</b>	<b>Trích xuất dữ liệu</b>
<b>Mô tả</b>	UC bắt đầu khi người quản trị thực hiện thu thập dữ liệu
<b>Dòng cơ bản</b>	<ol style="list-style-type: none"> <li>1. Người quản trị tiến hành vào trang hệ thống chọn chức năng cào data</li> <li>2. Kết quả trả về là số dòng dữ liệu và thời gian thực hiện cào dữ liệu.</li> <li>3. Kết thúc UC</li> </ol>
<b>Dòng thay thế</b>	Trường hợp có lỗi sẽ hiển thị lỗi trên màn hình

**Đặc tả UC Trích xuất dữ liệu**

Tên UC	Trích xuất dữ liệu
Mô tả	UC bắt đầu khi người quản trị thực hiện thu thập dữ liệu
Dòng cơ bản	<ol style="list-style-type: none"> <li>1. Người quản trị tiến hành vào trang hệ thống chọn chức năng cào data</li> <li>2. Kết quả trả về là số dòng dữ liệu và thời gian thực hiện cào dữ liệu.</li> <li>3. Kết thúc UC</li> </ol>
Dòng thay thế	Trường hợp có lỗi sẽ hiển thị lỗi trên màn hình

**Thiết kế giao diện:**

Chức năng	UI
Tìm kiếm	SearchBox + Button Search + Icon search
Filter Country	Multiple Select Chip Country
Sort Time	Customize select label Time
Thông tin hội nghị	Card Conferences
Phân trang	Pagination

Hình Giao diện web app IT conferences crawler

### 4.3 Cài đặt (implement) phần mềm

Cài đặt Node từ trang : <https://nodejs.org/en>

Giải nén source code, tại thư mục gốc của source code chạy lệnh “npm run dev” để vào chế độ develop.

*Chạy lệnh ‘npm run build’ sau đó ‘npm run start’ để chạy app ở chế độ production.*

Truy cập: <http://localhost:3000/api> để cào dữ liệu.

Truy cập: <http://localhost:3000> để xem danh sách hội nghị.

### 4.4 Kiểm thử phần mềm

Prerequisites				
1	Access to Chrome Browser			
2	Access to FireFox Browser			
3	Access to Cốc Cốc Browser			
Test function				
ID	Test Data			
1	Tìm kiếm hội nghị			
2	Filter hội nghị			
3	Sort hội nghị			
Test UI				
	Test Data	Pass/Fail/Note Executed/ Suspended		
1	UI nên show textbox để tìm kiếm	Pass		
2	Có selection để filter	Pass		
3	Có selection để sort	Pass		
Test Case ID	C001	Test Case Description: Test The Searching Conferences by Guest		
Date Tested	21-05-2024	Test Case (Pass/Fail/Not): Pass		
Step	Step Detail	Expected Results	Actual Results	Pass/Fail/Note Executed/ Suspended
1	Chọn filter theo 1 Country bất kì	Danh sách hội nghị chỉ hiển thị những hội nghị được tổ chức ở country đã chọn	As Expected	Pass
2	Chọn sort Latest	Danh sách hiển thị từ ngày gần nhất trở về sau	As Expected	Pass
3	Chọn sort Upcoming	Danh sách hiển thị các hội nghị sắp diễn ra	As Expected	Pass
4	Vừa chọn Filter theo Country vừa sort	Hiển thị các hội nghị thỏa filter và sort vừa chọn	As Expected	Pass
5	Reload web	Không mất các sort và filter vừa chọn	As Expected	Pass
6	Bỏ chọn filter theo Country	Danh sách Không filter theo Country nữa	As Expected	Pass
7	Bỏ chọn sort	Danh sách Không sort	As Expected	Pass
8	Chọn nút home	Hiển thị danh sách không sort không filter, không kết quả tìm kiếm	As Expected	Pass

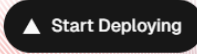
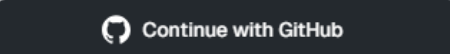
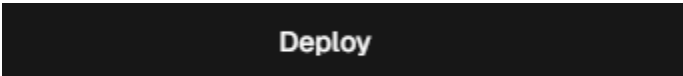
Test Case ID	C001	Test Case Description: Test The Sort Conferences by Guest		
Date Tested	21-05-2024	Test Case (Pass/Fail/Not): Pass		
1	Nhập khoảng trống	Danh sách hội nghị trống		
2	Nhập country bất kì	Danh sách hiển thị những hội nghị có keyword search	As Expected	Pass
3	Chọn sort Latest	Danh sách hiển thị từ ngày gần nhất trở về sau	As Expected	Pass
4	Chọn sort Upcoming	Danh sách hiển thị các hội nghị sắp diễn ra	As Expected	Pass
5	Vừa chọn Filter theo Country vừa s	Hiển thị các hội nghị thỏa filter và sort vừa chọn	As Expected	Pass
6	Reload web	Không mất các sort và filter vừa chọn	As Expected	Pass
7	Bỏ chọn filter theo Country	Danh sách Không filter theo Country nữa	As Expected	Pass
8	Bỏ chọn sort	Danh sách Không sort	As Expected	Pass
9	Chọn nút home	Hiển thị danh sách không sort không filter, không kết quả tìm kiếm	As Expected	Pass

Test Case ID	C001	Test Case Description: Test The Filter Conferences by Guest		
Date Tested	21-05-2024	Test Case (Pass/Fail/Not): Pass		
1	Chọn sort Latest	Danh sách hiển thị từ ngày gần nhất trở về sau	As Expected	Pass
2	Chọn sort Upcoming	Danh sách hiển thị các hội nghị sắp diễn ra	As Expected	Pass
3	Nhập từ khóa search bất kì	Danh sách hiển thị các hội nghị được filter và sort theo keyword tìm kiếm	As Expected	Pass
4	Reload app	Giữ nguyên kết quả	As Expected	Pass
5	Chọn Home	Quay về giao diện ban đầu	As Expected	Pass

## 4.5 Triển khai, bảo trì

Triển khai:

- Create folder Github để deploy toàn bộ source lên Github
- Truy cập <https://vercel.com/> và chọn  để thực hiện triển khai web app.
- Chọn  để Import Git Repository.
- Chọn import ở folder source code của web app, sau đó click  để triển khai web app.

### Import Git Repository

RicaNguyen

Search...

IT-Conferences-Crawler · 8d ago

Import

DTTX-HKI-2024 · 119d ago

Import

Python-Albase · 155d ago

Import

dbms-2024 · 136d ago

Import

a- · 141d ago

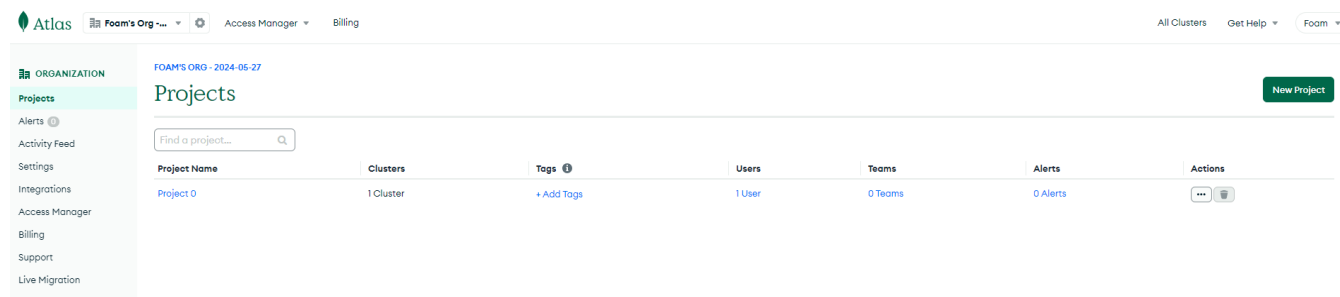
Import

[Import Third-Party Git Repository →](#)

Kết nối database MongoDB:

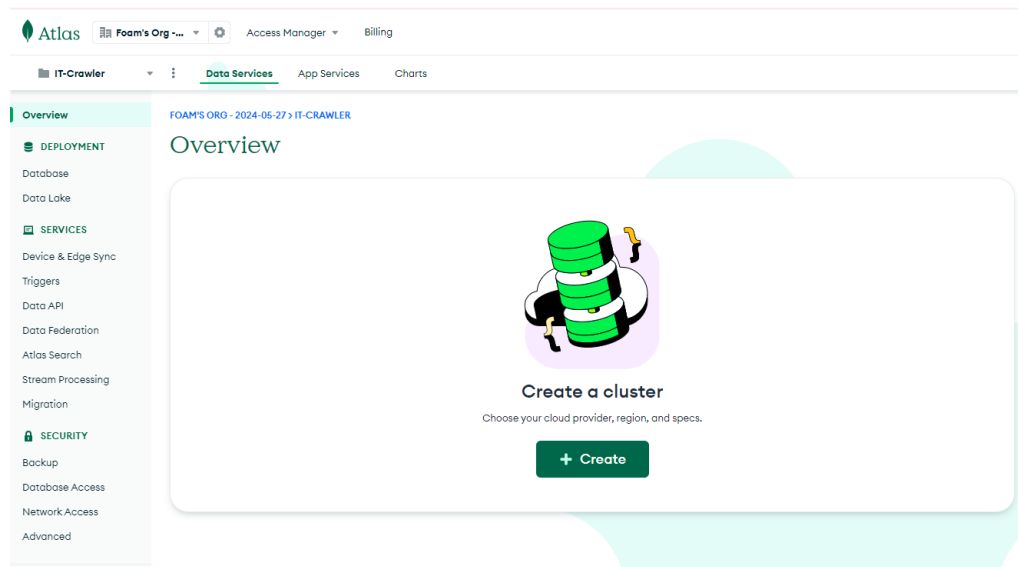
Step 1: Đăng ký tài khoản: Truy cập <https://www.mongodb.com/cloud/atlas/register> đăng kí sau đó đăng nhập.

Step 2: Tạo mới project: Click Project và Click New Project

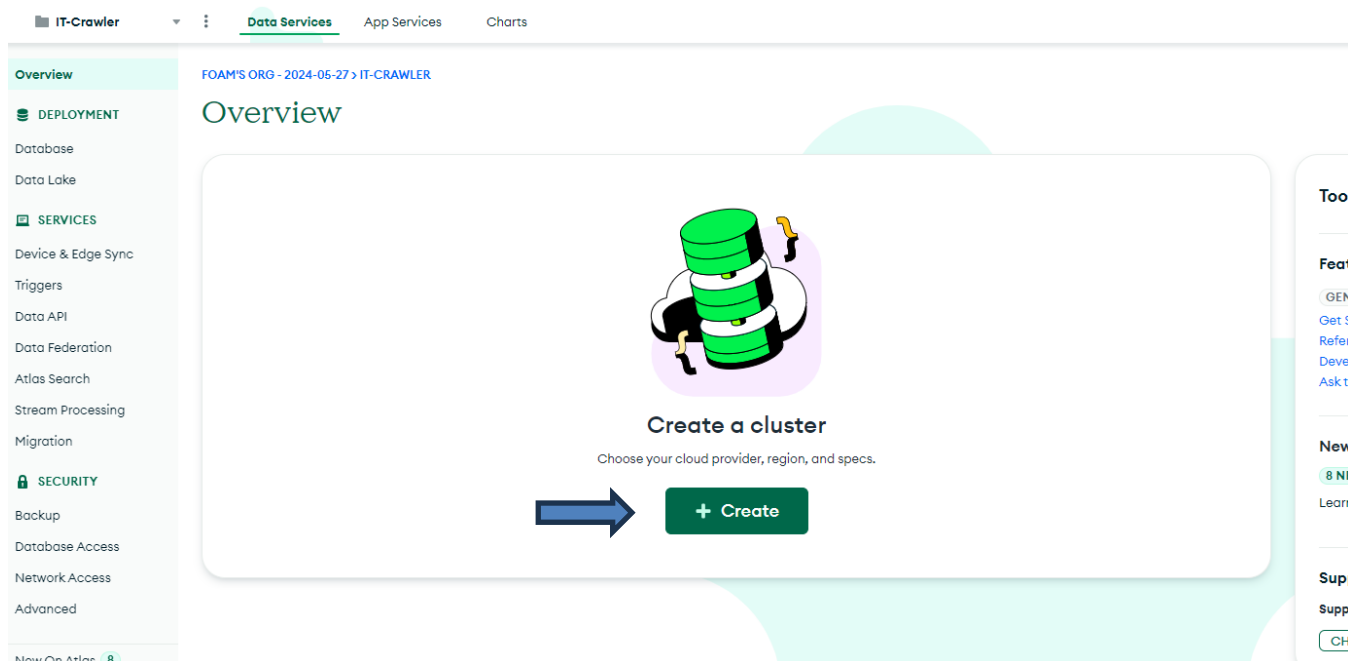


Step 3: Đặt tên Project (tên Project phải là duy nhất). Nếu tên không bị trùng thì Click vào Next

Step 4: Thêm thành viên và set quyền truy cập. Hiện tại chưa làm chung với ai nên Click vào Create Project. Đợi một lúc ta có Project như hình:



Step 5: Tạo Database: Click vào Create



Step 6: Chọn Cloud Database: Cái nào miễn phí thì chọn.

## Deploy your cluster

Use a template below or set up advanced configuration options. You can also edit these configuration options once the cluster is created.

Cluster Type	Price	Description	STORAGE	RAM	vCPU
<input type="radio"/> M10	\$0.10/hour	For production applications with sophisticated workload requirements.	10 GB	2 GB	2 vCPUs
<input type="radio"/> Serverless		For application development and testing, or workloads with variable traffic.	Up to 1 TB	Auto-scale	Auto-scale
<input checked="" type="radio"/> M0	Free	For learning and exploring MongoDB in a cloud environment.	512 MB	Shared	Shared

**Free forever!** Your M0 cluster is ideal for experimenting in a limited sandbox. You can upgrade to a production cluster anytime.

Step 6: Tạo Cluster:

Provider & Region: AWS và Singapore (chọn vị trí gần bạn)

Cluster Name: đặt tên là cluster-mongo-IT-Conferences

Sau đó click vào Create Deployment

Lưu ý:

☐ Automate security setup i bỏ check mục này để cho phép bất kì IP truy cập database(whitelist IP)

Use a template below or set up advanced configuration options. You can also edit these configuration options once the cluster is created.

☐ **M10** \$0.09/hour

For production applications with sophisticated workload requirements.

STORAGE	RAM	vCPU
10 GB	2 GB	2 vCPUs

☐ **Serverless** \$0.12/1M reads

For application development and testing, or workloads with variable traffic.

STORAGE	RAM	vCPU
Up to 1 TB	Auto-scale	Auto-scale

☒ **M0** Free

For learning and exploring MongoDB in a cloud environment.

STORAGE	RAM	vCPU
512 MB	Shared	Shared

✔ **Free forever!** Your M0 cluster is ideal for experimenting in a limited sandbox. You can upgrade to a production cluster anytime.

**Name**  
You cannot change the name once the cluster is created.

☐ Automate security setup i

☐ Preload sample dataset i

**Provider**

aws

Google Cloud

Azure

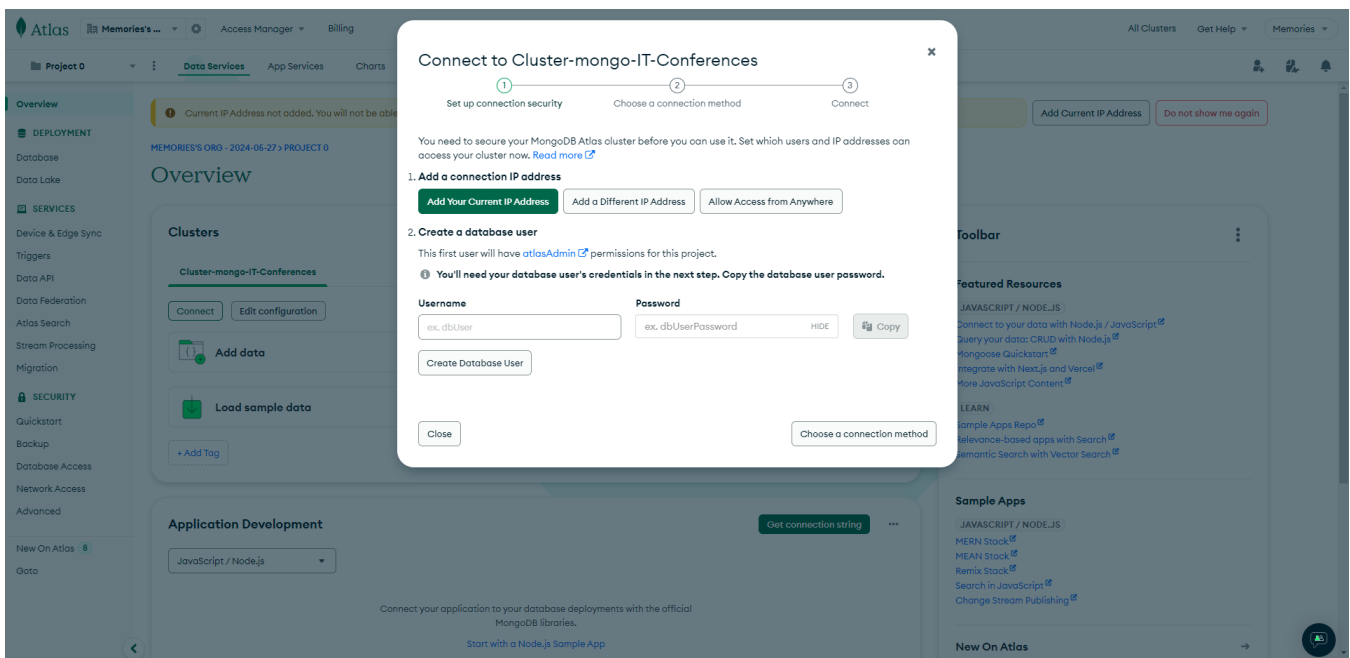
**Region**  
Singapore (ap-southeast-1) ★

★ Recommended i ✔ Low carbon emissions i

**Tag** (optional)  
Create your first tag to categorize and label your resources. [Learn more](#)

I'll do this later
Go to Advanced Configuration
Create Deployment

Giao diện khi hoàn thành:



Step 9: Thêm IP và Database User

## Phần thêm IP

Add Your Current IP Address

Add a Different IP Address

Allow Access from Anywhere

⇒ Chọn luôn là Allow Access from Anywhere và click luôn vào Add IP Address

Phần thêm Database User: Đặt tên là mongo-crawler và password 123456!. Tiếp theo là click vào Create Database User

Kết quả:



Step 10: Chọn phương thức kết nối:

Click vào Choose a connection memthod

**Connect to Cluster-mongo-IT-Conferences**

1 Set up connection security 2 Choose a connection method 3 Connect

You need to secure your MongoDB Atlas cluster before you can use it. Set which users and IP addresses can access your cluster now. [Read more](#)

1. Add a connection IP address

Add Your Current IP Address Add a Different IP Address Allow Access from Anywhere

2. Create a database user

✓ A database user has been added to this project. Create another user later in [Database Access](#).  
You'll need your database user's credentials in the next step.

Close Choose a connection method

Step 11: Tạo mã kết nối:

Chọn driver and version và Copy đoạn mã kết nối tự động được sinh ra:

## Connect to Cluster-mongo-IT-Conferences



### Connecting with MongoDB Driver

#### 1. Select your driver and version

We recommend installing and using the latest driver version.

Driver	Version
Node.js	5.5 or later

#### 2. Install your driver

Run the following on the command line

```
npm install mongodb
```

[View MongoDB Node.js Driver installation instructions.](#)

#### 3. Add your connection string into your application code

☐ View full code sample

```
mongodb+srv://mongo-crawler:<password>@cluster-mongo-it-confer.qutyf4r.mongodb.net/?
retryWrites=true&w=majority&appName=Cluster-mongo-IT-Conferences
```

Replace **<password>** with the password for the **mongo-crawler** user. Ensure any option params are [URL encoded](#).

#### RESOURCES

[Get started with the Node.js Driver](#)

[Node.js Starter Sample App](#)

[Access your Database Users](#)

[Troubleshoot Connections](#)

[Go Back](#)

[Done](#)

Sau đó run lệnh `npm install mongodb` ở command line.

Tạo thư mục db trong folder app -> tạo file mới đặt tên: `mongo.ts` và paste đoạn Code đã copy ở Step 11 vào.

# 5

## Kế hoạch nhân sự & chi phí

- Cấu trúc nhân sự cho toàn dự án: 1 nhân sự
- Chi phí cho toán bộ dự án: không mất phí.