\*\*\*\*\*Lịch sử hình thành

MongoDB (bắt nguồn từ “humongous”) là một hệ cơ sở dữ liệu NoSQL mã nguồn mở.

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

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. Dự án được bắt đầu triển khai vào tháng 10 năm 2007 bởi 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. Phải đến năm 2009, dự án này được tách độc lập. Hệ thống có thể chạy trên Windows, Linux, OS X và Solaris. 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.
* MTV Networks, Disney Interactive Media Group, bit.ly, The New York Times, The Guardian, SourceForge, Barclays, …

\*\*\*\*\*Khái niệm MongoDB

MongoDB là một chương trình cơ sở dữ liệu 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).

Trên thực tế thì MongoDB sử dụng thuật ngữ khác là **collection** hay bộ sưu tập thay vì bảng. Các cơ sở dữ liệu quan hệ (như MySQL hay SQL Server...) sử dụng các bảng để lưu dữ liệu. Các bảng này được cấu trúc với một số lượng cột (column) nhất định và các cột này cũng được định nghĩa với kiểu dữ liệu nhất định. Ngược lại MonoDB lưu **document** (hay tài liệu tương tự như các record trong MySQL hay SQL Server) vào các **collection** với định dạng JSON hay Javascript Object Notation. Do đó khi truy vấn dữ liệu hay cập nhật dữ liệu của **document** trong MongoDB chúng ta sử dụng cú pháp theo kiểu hươngs đối tượng.

### ƯU VÀ NHƯỢC ĐIỂM CỦA MONGODB LÀ GÌ?

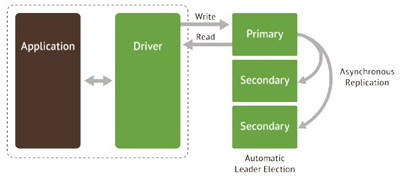
### ***a. Ưu điểm***

* Dữ liệu lưu trữ phi cấu trúc, không có tính ràng buộc, toàn vẹn nên tính sẵn sàng cao, hiệu suất lớn và dễ dàng mở rộng lưu trữ
* Dữ liệu được caching (ghi đệm) lên RAM, hạn chế truy cập vào ổ cứng nên tốc độ đọc và ghi cao

***b. Nhược điểm***

* Không ràng buộc, toàn vẹn nên không ứng dụng được cho các mô hình giao dịch yêu cầu độ chính xác cao
* Không có cơ chế transaction (giao dịch) để phục vụ các ứng dụng ngân hàng
* Dữ liệu được caching, lấy RAM làm trọng tâm hoạt động vì vậy khi hoạt động yêu cầu một bộ nhớ RAM lớn
* Như đã giới thiệu ở trên, mọi thay đổi về dữ liệu mặc định đều chưa được ghi xuống ổ cứng ngay lập tức vì vậy khả năng bị mất dữ liệu từ nguyên nhân mất điện đột xuất là rất cao.

### MONGODB HOẠT ĐỘNG NHƯ THẾ NÀO?



Mô hình hoạt động của MongoDB

* MongoDB hoạt động dưới một tiến trình ngầm service luôn mở một cổng (Cổng mặc định là 27017) để lắng nghe các yêu cầu truy vấn, thao tác từ các ứng dụng gửi vào sau đó mới tiến hành xử lý.
* Mỗi một bản ghi của MongoDB được tự động gắn thêm một field có tên “\_id” thuộc kiểu dữ liệu ObjectId mà nó quy định để xác định được tính duy nhất của bản ghi này so với bản ghi khác, cũng như phục vụ các thao tác tìm kiếm và truy vấn thông tin về sau.
* Trường dữ liệu “\_id” luôn được tự động đánh index (chỉ mục) để tốc độ truy vấn thông tin đạt hiệu suất cao nhất.
* Mỗi khi có một truy vấn dữ liệu, bản ghi được cache (ghi đệm) lên bộ nhớ Ram, để phục vụ lượt truy vấn sau diễn ra nhanh hơn mà không cần phải đọc từ ổ cứng.
* Khi có yêu cầu thêm/sửa/xóa bản ghi, để đảm bảo hiệu suất của ứng dụng mặc định MongoDB sẽ chưa cập nhật xuống ổ cứng ngay, mà sau 60 giây MongoDB mới thực hiện ghi toàn bộ dữ liệu thay đổi từ RAM xuống ổ cứng.

KHÁI NIỆM CƠ SỞ DỮ LIỆU

**Database**  
Dịch ra là cơ sở **dữ liệu**.  
Nó là nơi tạo ra, cập nhật và lưu trữ mọi thông tin (dữ liệu) mà chúng ta sử dụng trên mạng, mọi thao tác chúng ta thực hiện với game và trang quản lý tài khoản... chỉ là một hình thức gián tiếp tác động vào cơ sở dữ liệu này.  
  
Nói nôm na thì bộ hồ sơ gồm danh sách học sinh trong lớp, thông tin từng học sinh... chính là database (cơ sở dữ liệu) của lớp học. Database cũng có thể ví như một thư viện với vô vàn sách được sắp xếp gọn gàng theo từng chủ đề và luôn được cập nhật.  
  
Mọi thông tin tài khoản của bạn đều được lưu ở database giống như những ghi chép trong hồ sơ mà bạn có thể thấy trong đời thực, nếu tiếp cận được với database thì bạn có thể đọc được một số thông tin. Tuy nhiên, tiếp cận được với database không có nghĩa là bạn có thể đọc được mọi thông tin trong nó vì một số thông tin hoặc thậm chí toàn bộ database có thể đã được mã hóa hoặc cần được tác động thông qua quyền và chức năng điều khiển khác.