

Bài tập tuần 3:

1 Tên hệ quản trị Cơ sở dữ liệu

Mongo DB

2 Lịch sử hình thành:

Dự án Mongo DB được hình thành phát triển vào năm 2007 bởi công ty 10gen (hiện nay đã đổi tên thành MongoDB Inc.) Ban đầu nó được thiết kế như một phiên của một sản phẩm nền tảng dịch vụ (Platform as Service) Tuy nhiên đến năm 2009, công ty quyết định tách nó ra và phát hành dưới dạng mã nguồn mở. Từ đó đến nay, MongoDB đã trở thành hệ quản trị CSDL NoSQL phổ biến nhất thế giới.

3 Các giá và tổ chức quản lý

Các sáng lập Eliot Horowitz và Dwight Merriman.

Tổ chức quản lý hiện tại MongoDB Inc

4 Mô hình lưu trữ

Mongo DB loại Cơ sở dữ liệu Document Store (Hướng tài liệu). Thay vì lưu dữ

liều dưới dạng bảng và các hàng cố định như trong SQL truyền thống. MongoDB liều dữ liệu dưới dạng các tài liệu BSON. Các tài liệu các hàng cố cấu trúc linh hoạt, nghĩa là các bản ghi trong cùng một bộ sưu tập không bắt buộc phải có cùng các trường dữ liệu giống hệt nhau.

5 Ngôn ngữ thao tác dữ liệu

MongoDB không sử dụng SQL thay vào đó nó sử dụng MongoDB Query Language (MQL). Đây là ngôn ngữ truy vấn trên các đối tượng và phương thức cho phép thực hiện CRUD và các phép tổng hợp dữ liệu.

6 Cơ chế phân tán:

Cơ chế phân tán của MongoDB dựa trên kỹ thuật Phân nhánh

Nguyên lý khi dữ liệu quá lớn vượt quá khả năng lưu trữ 1 máy chủ vật lý MongoDB sẽ tự động chia nhỏ tập dữ liệu đó và phân phối chúng trên nhiều máy chủ khác nhau. Hệ thống có một tên chính

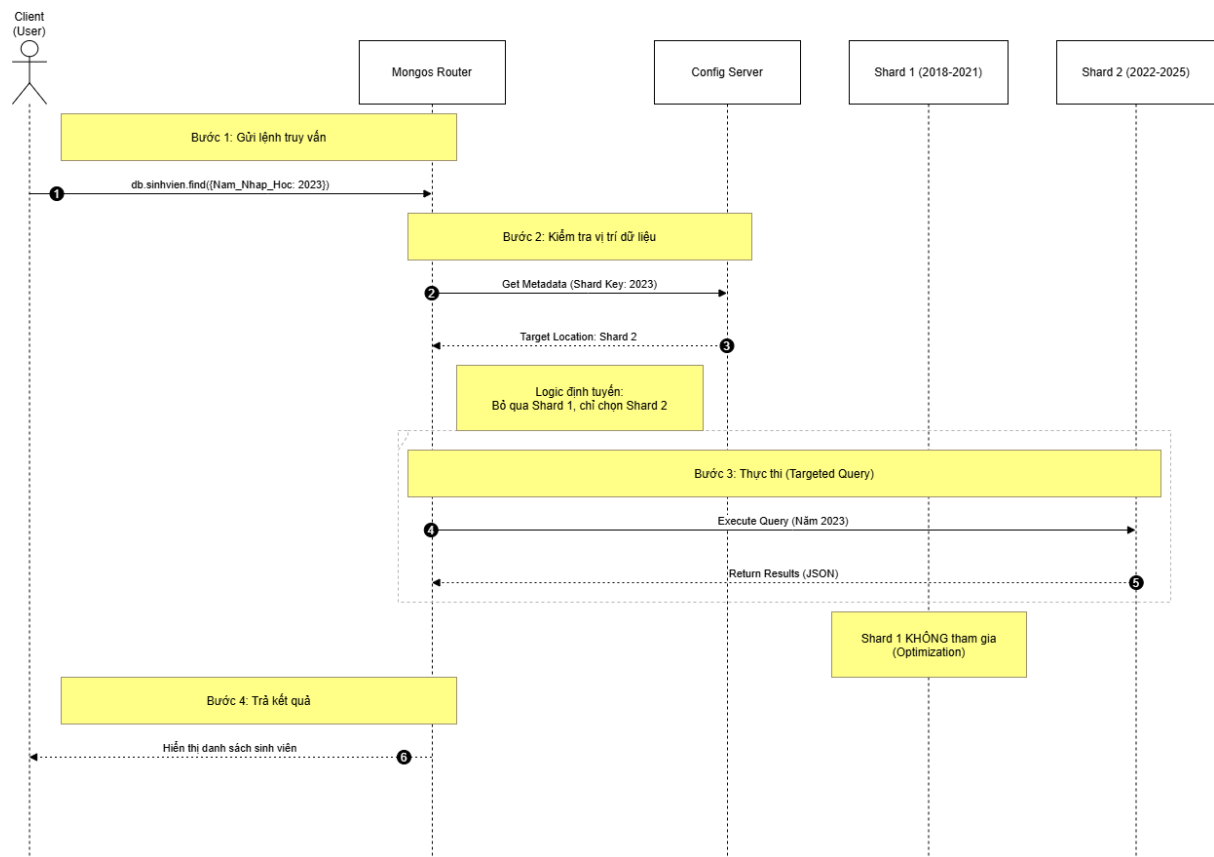
Hệ Chạy ngầm (Bastlancer) giúp từ động di chuyển các khối dữ liệu giữa các shard để đảm bảo không có máy chủ nào bị quá tải giúp hệ thống luôn ở tình trạng ổn định.

Quy trình mô phỏng xử lý truy vấn
Bước 1 Tiếp nhận yêu cầu truy vấn chuyển đổi thành câu lệnh Find đến cổng giao tiếp của server

Bước 2 Phân tích và Định tuyến
Mongo Router phân tích cách thức làm việc và phân bổ truy vấn đang tìm kiếm dữ liệu dựa trên địa chỉ IP - Name - Namespace

Router lập tức liên hệ config server để tra tìm năm 2023

Bước 3 Xác định mục tiêu (Targeting)
B



use TruongDaiHoc;

sh.enableSharding("TruongDaiHoc");

db.sinhvien.createIndex({ "Nam_Nhap_Hoc": 1 });

sh.shardCollection("TruongDaiHoc.sinhvien", { "Nam_Nhap_Hoc": 1 });

print("\n>>> BẮT ĐẦU NẠP DỮ LIỆU...");

var dataMau = [

 { "MSSV": "SV101", "Ten": "Nguyen Van A", "Khoa": "CNTT",
 "Nam_Nhap_Hoc": 2019 },

 { "MSSV": "SV102", "Ten": "Tran Thi B", "Khoa": "KT", "Nam_Nhap_Hoc":
2020 },

```
    { "MSSV": "SV201", "Ten": "Le Van C", "Khoa": "CNTT", "Nam_Nhap_Hoc":  
2023 },  
    { "MSSV": "SV202", "Ten": "Pham Thi D", "Khoa": "NN", "Nam_Nhap_Hoc":  
2023 },  
    { "MSSV": "SV203", "Ten": "Hoang Van E", "Khoa": "CNTT",  
"Nam_Nhap_Hoc": 2024 }  
];
```

```
db.sinhvien.insertMany(dataMau);  
print(">>> Đã nạp thành công 5 hồ sơ sinh viên.");  
print(">>> Dữ liệu đã được tự động phân tán sang Shard 1 và Shard 2.");
```

```
print("\n>>> BẮT ĐẦU TRUY VẤN...");  
print(">>> Yêu cầu: Tìm sinh viên khoa CNTT nhập học năm 2023");
```

```
var query = {  
    "Khoa": "CNTT",  
    "Nam_Nhap_Hoc": 2023  
};
```

```
var ketqua = db.sinhvien.find(query);
```

```
print("-----");  
print("KẾT QUẢ TÌM THẤY:");  
ketqua.forEach(doc => {  
    print(` - Sinh viên: ${doc.Ten} | MSSV: ${doc.MSSV} | Năm:  
${doc.Nam_Nhap_Hoc} `);  
});
```

```
print("-----");
```

```
print("\n>>> KIỂM TRA LUỒNG DỮ LIỆU (EXPLAIN PLAN)...");
```

```
var plan = db.sinhvien.find(query).explain("executionStats");
```

```
var chienLuoc = plan.queryPlanner.winningPlan.stage;
```

```
var shardXuLy = plan.queryPlanner.winningPlan.shards[0].shardName;
```

```
print(`1. Chiến lược thực thi: ${chienLuoc}`);
```

```
// Giải thích: SINGLE_SHARD nghĩa là Router chỉ gửi lệnh đến 1 máy (Tốt)
```

```
print(`2. Máy chủ xử lý: ${shardXuLy}`);
```

```
// Giải thích: Shard_2 nghĩa là hệ thống đã định tuyến đúng đến nơi chứa dữ liệu năm  
2023
```

```
print("\n>>> KẾT THÚC MÔ PHỎNG.");
```