

## Bài tập lý thuyết tuần 5

Nguyễn Thành Viên CNTT5, 1252080231

- Tên họ & số: Apache Cassandra

Thực hiện wide Column Store, quản lý kẽ hở phân  
loại trên website NoSQL database - From Basics to Intermediate  
use

- Lịch sử hình thành

+ Apache Cassandra chính thức ra mắt vào năm

2008 bởi Facebook

+ Mục đích ban đầu là phục vụ hệ thống Inbox Search  
với yêu cầu dữ liệu lớn, phân tán và chủ động.

+ Năm 2009, Cassandra chính thức chuyển giao cho Apache Software Foundation quản lý.

+ Hiện nay, Cassandra là một trong những hệ quản trị NoSQL phổ biến nhất trong danh nghiệp

- Tác giả, tổ chức quản lý

+ Tác giả ban đầu: Facebook

+ Tổ chức quản lý hiện tại: Apache Software Foundation

+ Giấy phép: Apache License

- Mô hình lưu trữ

Lưu trữ NoSQL, wide column store

- Đặc điểm

→ Dữ liệu được lưu trữ theo bảng nhưng không gộp, POBNS  
+ Nội dung gồm:

partition key, clustering key

+ Chia đều dữ liệu

+ Schwartzian Transform để lưu trữ.

- Ngôn ngữ thao tác dữ liệu

Cassandra sử dụng CQL (Cassandra Query Language)

Cú pháp gần giống SQL truyền thống

Cách phân tán dữ liệu

Kết cấu phân tán:

+ Cassandra sử dụng mô hình Peer-to-Peer

+ Không có node Master

+ Mỗi node có vai trò ngang nhau

phân mảnh dữ liệu

+ Dữ liệu được phân tán trên Partition key

+ Áp dụng thuật toán Consistent Hashing

Sao chép dữ liệu

Mỗi bản ghi có thể tồn tại trên nhiều node

Grub hệ thố:

+ Chia bài toán

+ Chia giao diện chỉ đọc gấp 100 lần

- MS phay luu tru cua tuy van phan de

Bai oanh muc phay:

Xem hinh thong quan ly otay ky he phan cuu sinh  
vien hoc.

Du lieu sinh vien duoc luu tru va phan tich theo chuc  
nhom. Trong ban phan nay moi bang va trang san sang  
coo dia hoi thuong.

- Thieu loi ben, du lieu mang Cassandra

Create table daugke Hoppon (

chon text,

masinhvien text,

mamnhvien text,

hocky text,

primary key (chon, masinhvien)

);

- MS phay luu tru du lieu phan tich

(o g vdi) He thong Cassandra gom 3 node:

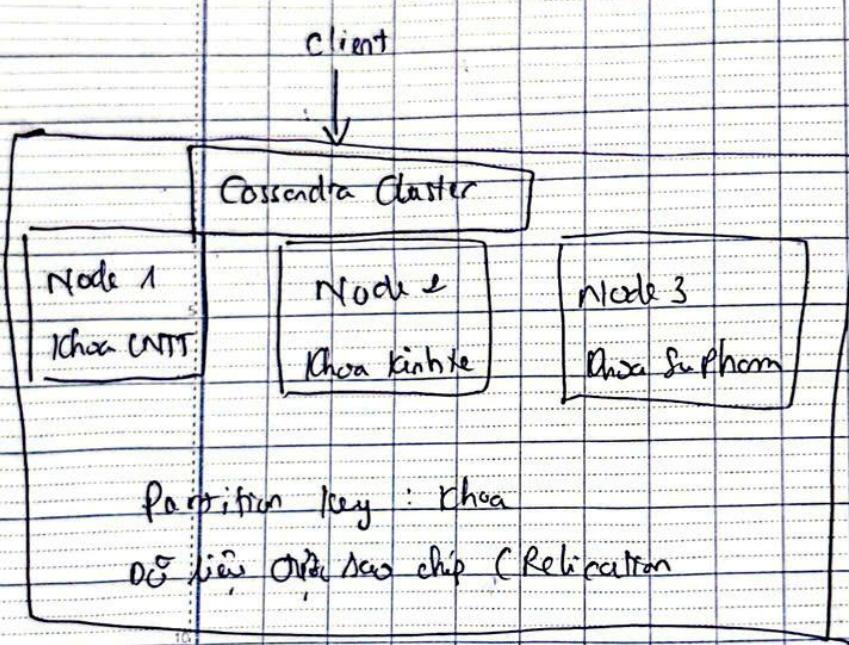
Node1: Du lieu sinh vien chua CNTT

Node2: Du lieu sinh vien khac sinh vien

Node3: Du lieu sinh vien phan giao Phan

- Mở phím Thao tác chiết Xuất phân tán  
insert into donglyhocphan (chien, masinhien, manmanhoc,  
hocky)  
values ('xanhte'), ('socol'), ('kt201'), ('2024-2'));
- Cố định  
+ Khoảng cm<sup>3</sup> đứng ngang  
+ Phù hợp với kí tự thứ 4 là kí tự tiếp theo  
+ Kím backslash \& \n vào trong chuỗi thông phân tán
- Nhắc nhở  
+ Khiết kế phép nối join  
+ Khiết kế phù hợp với kết quả cách phân tách phím tắt phím tắt (ctrl quen hệ).
- Try vào phân tán khay có Partition key  
Select \* from donglyhocphan  
where manmanhoc = 'cs101';
- Try vào phân tán có Partition key  
Select \* from donglyhocphan  
where faculty = 'CNTT');

- mảng khub lưu trữ phân tán Cassandra



- Dữ liệu được phân tán trên nhiều node trong Cassandra Cluster  
dựa trên Partition Key (Khoa)

- Dữ liệu chia sao chép giữa các node nhằm đảm bảo tính  
tồn tại và khả năng chịu lỗi