**TRƯỜNG ĐẠI HỌC TÀI NGUYÊN VÀ MÔI TRƯỜNG THÀNH PHỐ HỒ CHÍ MINH**

**KHOA HỆ THỐNG THÔNG TIN VÀ VIỄN THÁM**



**TIỂU LUẬN**

**HỌC PHẦN: HỆ QUẢN TRỊ CƠ SỠ DỮ LIỆU**

**ĐỒ ÁN**

Sinh viên thực hiện **: HUỲNH TẤN ĐẠT, NGUYỄN NGỌC TÚ, NGUYỄN THANH BAO**

Lớp **: 09\_ĐH\_cntt03**

Giảng viên hướng dẫn**:** HUỲNH PHAM TRỌNG

**TP. Hồ Chí Minh, tháng 11 năm 2022**

Logo

Description automatically generated**Volume 2, Issue 8, August 2012 ISSN: 2277 128X**

**International Journal of Advanced Research in Computer Science and Software Engineering**

**Research Paper**

**Available online at: [www.ijarcsse.com](http://www.ijarcsse.com/)**

# SQL and NoSQL Databases

**Vatika Sharma1, Meenu Dave2**

*1M.Tech. Scholar,Department of CSE, Jagan Nath University, Jaipur, India vatika.sharma15@gmail.com*

*2Assistant Professor, Department of CSE, Jagan Nath University, Jaipur, India meenu.dave@jagannathuniversity.org*

Tóm tắt: NoSQL (Not only SQL) là một cơ sở dữ liệu được sử dụng để lưu trữ một lượng lớn dữ liệu. Cơ sở dữ liệu NoSQL được phân phối không quan hệ với mã nguồn mở và có thể mở rộng theo chiều ngang (theo cách tuyến tính). NoSQL không tuân theo thuộc tính của ACID như trong SQL. Trong tài liệu nghiên cứu này, chúng tôi đang khảo sát về NoSQL, nền tảng của nó, các nguyên tắc cơ bản như ACID, BASE và định lý CAP. trên cơ sở định lý CAP, nghiên cứu được thực hiện về các loại kho dữ liệu NoSQL khác nhau với những ví dụ của nó, đặc điểm, và ưu nhược điểm của NoSQL.

Keywords: ACID, BASE, SQL, NoSQL, CAP, CURD

# I. khái niệm

NoSQL viết tắt của Not only SQL. Nó được phát âm là noseequel.Đây là một trong những loại lưu trữ dữ liệu khác ngoài cơ sở dữ liệu (đã được sử dụng trước đó) được sử dụng để lưu trữ lượng lưu trữ dữ liệu khổng lồ như dữ liệu trong Facebook,...quan hệ (đôi khi được gọi là bắt nguồn từ quan hệ cơ sở dữ liệu) , cơ sở dữ liệu truy xuất thông tin nhanh và di động. NoSQL về cơ bản bắt nguồn từ hệ thống cơ sở dữ liệu RDB. Này cơ sở dữ liệu thường tương tác với hệ điều hành UNIX. Cơ sở dữ liệu NoSQL là những cơ sở dữ liệu phi quan hệ, mã nguồn mở, phân bố trong tự nhiên cũng như nó có hiệu suất cao theo cách tuyến tính có thể mở rộng theo chiều ngang Cơ sở dữ liệu phi quan hệ không sắp xếp dữ liệu của nó trong các bảng liên quan (tức là dữ liệu được lưu trữ theo cách không chuẩn hóa). Cơ sở dữ liệu NoSQL là mã nguồn mở; Do đó, mọi người có thể xem xét mã của nó một cách tự do, cập nhật nó theo nhu cầu của mình và biên dịch nó. Phân tán có nghĩa là dữ liệu được lan truyền đến các máy khác nhau và được quản lý bởi các máy khác nhau nên ở đây nó sử dụng khái niệm dữ liệu nhân rộng. NoSQL có thể được biểu diễn một cách tượng trưng như trong hình 1:

Icon

Description automatically generated

hình 1: Symbolic representation of NoSQL

Hình 1 cho biết việc truy vấn cơ sở dữ liệu mà không có bất kỳ tương tác hoặc giao diện nào của ngôn ngữ SQL. Đường nghiêng trong hình hiển thị việc sử dụng cơ sở dữ liệu mà không cần sử dụng SQL (Ngôn ngữ truy vấn có cấu trúc). Vì vậy, để truy cập các cơ sở dữ liệu này, chúng ta có thể sử dụng một số cơ sở dữ liệu khác các định dạng như XML để lưu trữ và truy xuất thông tin từ cơ sở dữ liệu.

Với sự ra đời của các trang mạng xã hội như facebook và twitter, nhu cầu về công nghệ mới có thể xử lý rất lớn lượng dữ liệu đã dẫn đến sự xuất hiện của nhiều công nghệ mới khác nhau và một trong những điểm nổi bật là NoSQL khá hữu ích trong việc lưu kho dữ liệu. NoSQL (phi quan hệ) tương đối nhanh hơn cơ sở dữ liệu quan hệ. Trước đây, trong SQL, chúng tôi đã sử dụng ngôn ngữ Truy vấn để tìm nạp cũng như lưu trữ dữ liệu; đối với NoSQL, chúng tôi lưu trữ các thực thể dữ liệu lớn bằng cách sử dụng tài liệu trong Định dạng XML (Ngôn ngữ đánh dấu eXtensible). Ngôn ngữ XML về cơ bản được sử dụng để lưu trữ dữ liệu có cấu trúc trong một con người có thể đọc được dạng. Điều này được thảo luận thêm trong các phần sau.

# II. BACKGROUND

Vào năm 1960, cơ sở dữ liệu Đa giá trị đã được phát triển để thực hiện một số thao tác cơ sở dữ liệu phân cấp bằng cách sử dụng lưu trữ cây B+.

Theo Scott Jones M[umps] được phát triển vào năm 1966 để hoạt động dựa trên các loại cơ sở dữ liệu (phân cấp) như vậy.Năm 1977, M[umps] đã được phê duyệt là ngôn ngữ tiêu chuẩn ANSI. Năm 1979, Ken Thompson đã phát triển DBM (Trình quản lý cơ sở dữ liệu)sử dụng các kỹ thuật băm để cung cấp cho người dùng khả năng truy xuất dữ liệu nhanh hơn. Vào năm 1980, số lượng lớn những người kế thừa DBM là phát triển như sau:

TDBM là cơ sở dữ liệu loại DBM hỗ trợ các giao dịch nguyên tử (tuân theo các thuộc tính ACID A: Atomicity, C:Tính nhất quán, I: Cách ly, D: Độ bền). Ưu điểm của TDBM bao gồm:

(1) Nếu ứng dụng đang sử dụng TDBM, trong trường hợp bị hư hỏng hoặc treo, nội dung được lưu trữ trong cơ sở dữ liệu sẽ được an toàn.

(2) TDBM cũng hỗ trợ các giao dịch nguyên tử lồng nhau sử dụng các hoạt động đa luồng.SDBM là cơ sở dữ liệu loại DBM được sử dụng vì một số lý do cấp phép. GT.M là một sản phẩm kế thừa khác của DBM tập trung vào xử lý giao dịch có hiệu suất cao. Năm 2000, DBM này mã nguồn mở [4].

Sự phát triển của cơ sở dữ liệu SQL bắt đầu vào cuối những năm 1990. Sau một vài năm, nó đã trở thành đối thủ nặng ký của RDBMS(Relational Database Management Systems) [4]. Trong năm 2009 và 2010 đã có các hội nghị NoSQL được tổ chức như NoSQL live, NoSQL eu, NoSQL East, v.v. Tên NoSQL này lần đầu tiên được Carlo Strozzi sử dụng vào năm 1998 làm tên của tệp anh ấy đang phát triển cho cơ sở dữ liệu của mình.

Ngày nay, NoSQL đang trở nên quá phổ biến do dung lượng lưu trữ cao và cũng bởi vì các thuộc tính của nó tránh được các tính năng cơ bản của SQL.

Cơ sở dữ liệu quan hệ được thiết kế để chạy trên một máy duy nhất, vì vậy để làm được điều này, chúng ta cần một máy lớn để mở rộng quy mô. Một Giải pháp khắc phục là mua thật nhiều máy (và gom thành một cụm) để lưu trữ dữ liệu. Nó sẽ là một quá trình rẻ hơn cũng như khả năng mở rộng theo chiều ngang (đường tuyến tính) cũng được. Ngay cả khi một máy bị hỏng, độ tin cậy chung của cụm là khá cao.Do đó, sự thống trị của SQL đang giảm dần và NoSQL đang trở nên phổ biến.

Diagram

Description automatically generated

Hình 2: Decline in dominance of SQL

III. AXIOMATICS OF NoSQL

A. ACID miễn phí

ACID là viết tắt của Atomicity, Consistency, Isolation và Durability. Khái niệm ACID về cơ bản xuất phát từ Môi trường SQL . Nhưng trong NoSQL chúng ta sẽ không sử dụng khái niệm ACID vì tính năng Nhất quán của SQL [6]. Trong bài báo này chúng tôi sẽ thấy cách khái niệm ACID tạo ra các vấn đề đối với NoSQL.Như trong môi trường phân tán, dữ liệu được lan truyền đến các máy khác nhau, mỗi máy lưu trữ dữ liệu của mình và bảo trì tính nhất quán là cần thiết. Ví dụ: nếu có thay đổi trong một bộ dữ liệu của bảng thì cần thay đổi trong từng bộ dữ liệu.máy chứa dữ liệu cụ thể đó. Nếu thông tin liên quan đến một bản cập nhật lan truyền ngay lập tức, thì tính nhất quán là được; nếu không, thì sự không nhất quán được thực hiện.

B. BASE

BASE là viết tắt của Tính nhất quán về cơ bản, khả dụng, trạng thái mềm và tính nhất quán cuối cùng. BASE ngược lại với ACID [1]. cơ sở dữ liệu NoSQL được phân chia giữa con đường từ ACID đến BASE. Sau khi thống nhất giao dịch, trạng thái mà chúng ta sẽ nhận được là trạng thái mềm không phải trạng thái rắn. Trọng tâm chính đằng sau BASE là tính khả dụng lâu dài [6].

Ví dụ: nghĩ về cơ sở dữ liệu trong ngân hàng, nếu hai người đang truy cập cùng một tài khoản ở các thành phố khác nhau thì dữ liệu cập nhật là cần thiết không chỉ trong thời gian mà còn cần một số cơ sở dữ liệu thời gian thực. Những cập nhật cần phải được thực hiện thường xuyên trên tất cả các máy. Một số ví dụ khác là đặt chỗ đường sắt trực tuyến, mua bán sách trực tuyến, v.v.

C. CAP

CAP là viết tắt của Tính nhất quán, Tính khả dụng và Dung sai phân vùng. CAP về cơ bản là một định lý tuân theo ba nguyên tắc

[1][2]:

(1) Dữ liệu có sẵn trên tất cả các máy phải giống nhau về mọi mặt và phải thường xuyên cập nhật trên tất cả các máy, tức là dữ liệu có tính nhất quán.

(2) Dữ liệu phải có sẵn vĩnh viễn và có thể truy cập mọi lúc, tức là có sẵn.

(3) Trong khi máy bị lỗi hoặc bất kỳ lỗi nào trong cơ sở dữ liệu của máy sẽ hoạt động tốt mà không dừng công việc của chúng, tức là phân vùng sức chịu đựng.

## IV. ROLE OF DATA ARCHITECTURE IN NOSQL

A. Các thành phần: Có bốn thành phần trong khối xây dựng của nó [6].

[1.] Ngôn ngữ mô hình hóa: Nó mô tả cấu trúc của cơ sở dữ liệu và cũng xác định lược đồ mà nó dựa vào. dữ liệu là được lưu trữ dưới dạng hàng và cột sử dụng định dạng XML. Và mỗi dữ liệu (giá trị) tương ứng với nó được gán một khóa đó là duy nhất trong tự nhiên. Để truy cập dữ liệu nhanh hơn, mô hình được xây dựng trong một môi trường phù hợp.

[2.] Cấu trúc cơ sở dữ liệu: Mỗi và mọi cơ sở dữ liệu trong khi xây dựng đều sử dụng cấu trúc dữ liệu riêng và lưu trữ dữ liệu bằng cách sử dụng thiết bị lưu trữ vĩnh viễn.

[3.] Ngôn ngữ truy vấn cơ sở dữ liệu: Tất cả các thao tác được thực hiện trên cơ sở dữ liệu là tạo, cập nhật, đọc và xóa (CURD).

[4.] Giao dịch: Trong bất kỳ giao dịch nào trong dữ liệu, có thể có bất kỳ loại lỗi hoặc lỗi nào; sau đó, máy sẽ không ngừng hoạt động.

B. Các kiểu lưu trữ dữ liệu NoSQL

Trên cơ sở định lý CAP cơ sở dữ liệu NoSQL được chia thành số cơ sở dữ liệu. Có bốn loại mới khác nhau của lưu trữ dữ liệu trong NoSQL [4].

1. Cơ sở dữ liệu giá trị khóa: Bản thân tên cơ sở dữ liệu giá trị khóa nói rằng nó là sự kết hợp của hai thứ là khóa và một giá trị. Nó là một trong những hệ thống cơ sở dữ liệu cấu hình thấp (truyền thống). Cơ sở dữ liệu Giá trị khóa (KV) là mẹ của tất cả cơ sở dữ liệu của NoSQL.

.  Khóa là mã định danh duy nhất cho một mục nhập dữ liệu cụ thể. Không nên lặp lại khóa nếu một khóa đã sử dụng không trùng lặp trongthiên nhiên. Giá trị là một loại dữ liệu được trỏ bởi một khóa.

Cơ sở dữ liệu giá trị khóa dường như ở dạng bảng băm hoặc bảng tra cứu. Trong loại cơ sở dữ liệu này, chỉ có một cách để truy vấn là với sự trợ giúp của khóa (duy nhất) và tất cả các khóa có thể đặt tên trong bất kỳ đối tượng dữ liệu nào và được sắp xếp theo thứ tự bảng chữ cái [3].

Để có tính sẵn sàng cao hơn của các cửa hàng dữ liệu, các đối tượng dữ liệu được sao chép.

Điều này được minh họa trong hình 4 bên dưới. Ví dụ: hãy lấy một ví dụ về cơ sở dữ liệu ngân hàng như trong hình 3.

|  |  |
| --- | --- |
|  | **BANK DATABASE** |
| Key | Value |
| 1 | ID:1  Joining Date: 15-July-1985 Designation: Cashier |
| 2 | ID:2  Joining Date: 19-March-1982  Designation: Manager |
| 3 | ID:3  Joining Date: 4-April-1988  Designation: Front Desk Officer |

Hình 3: Key Value (KV) databases

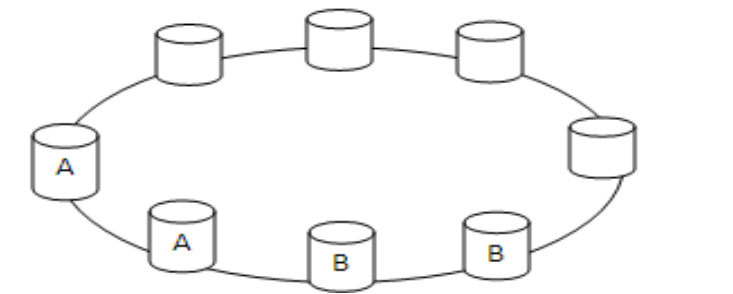
Hình 3 đã cho có hai cột đại diện cho khóa và một giá trị. Ở đây chìa khóa là duy nhất và đại diện cho giá trị của họ hoặc Các thuộc tính tương ứng với nó và dữ liệu được biểu diễn dưới dạng vòng và việc phân vùng dữ liệu được thực hiện trên cơ sở bảng chữ cái của họ (theo thứ tự sắp xếp) và dữ liệu cũng được sao chép dưới dạng vòng. Điều này như được trình bày trong phần tiếp theo**.**

1.1 **Đặc điểm của cơ sở dữ liệu khóa giá trị**

* Số khóa có thể có một bộ thuộc tính động trong cơ sở dữ liệu giá trị khóa trong quá trình lưu trữ dữ liệu.
* Dữ liệu được lưu trữ trong cơ sở dữ liệu được lưu trữ theo thứ tự bảng chữ cái.
* Tất cả các hoạt động có thể được thực hiện trên dữ liệu, tức là CRUD (Tạo, Đọc và Cập nhật và Xóa).
* Tất cả các mối quan hệ với dữ liệu được lưu trữ trong mã ứng dụng (không lan truyền rõ ràng).

**1.2 Cơ sở dữ liệu khóa Giá trị (KV) sử dụng**

* Đây là một trong những mô hình dữ liệu đơn giản nhất (chúng ta sẽ thảo luận sau) vì nó chỉ sử dụng khóa và giá trị. (như mô tả trong
* hình 3)
* Nó xử lý tải dữ liệu lớn.
* Nó chia tỷ lệ thành khối lượng lớn dữ liệu.
* Việc sao chép dữ liệu được thực hiện bằng cách sử dụng cơ sở dữ liệu ở dạng vòng. Dữ liệu sao chép được lưu trữ ở dạng vòng như cũng như theo thứ tự bảng chữ cái. Điều này được thể hiện trong hình 4 bên dưới:

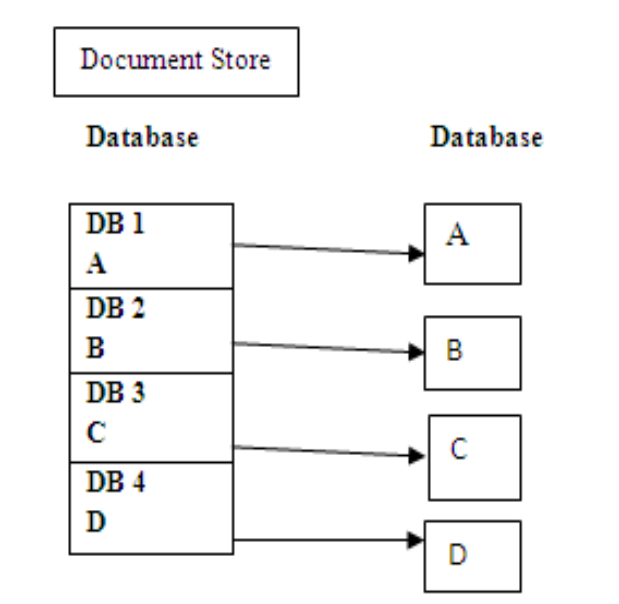


Hình 4.Phân vùng vòng và sao chép dữ liệu

2.**Kho lưu trữ tài liệu Cơ sở dữ liệu**

Cơ sở dữ liệu Cửa hàng Tài liệu là những cơ sở dữ liệu NoSQL sử dụng bản ghi làm tài liệu. Loại cơ sở dữ liệu lưu trữ này là tài liệu phi cấu trúc (văn bản) hoặc bán cấu trúc (XML) thường có tính chất phân cấp. Ở đây mỗi tài liệu bao gồm một tập hợp khóa và giá trị gần giống như trong cơ sở dữ liệu Giá trị khóa. Mỗi cơ sở dữ liệu nằm trong tài liệu lưu trữ các điểm đến các trường của nó bằng cách sử dụng các con trỏ vì nó sử dụng kỹ thuật băm. Kho lưu trữ tài liệu Cơ sở dữ liệu là lược đồ miễn phí và không cố định trong tự nhiên.

Cấu trúc của Cơ sở dữ liệu lưu trữ tài liệu được minh họa trong hình 5 dưới đây. Hình vẽ mô tả rằng nó bao gồm số lượng cơ sở dữ liệu trong kho lưu trữ tài liệu, chẳng hạn như cơ sở dữ liệu 1,2,3,4 và có id A, B, C, D nằm trong đó chỉ vào cơ sở dữ liệu của nó có một số liên quan đến nó. Cơ sở dữ liệu trỏ đến giá trị của nó bằng cách sử dụng một số khóa duy nhất nằm trong cơ sở dữ liệu. Điều này bao gồm một mảng cơ sở dữ liệu (có dạng nhóm). Điều này sẽ rõ ràng hơn sau khi lấy một ví dụ thảo luận dưới đây.



Hình 5: Cấu trúc cơ sở dữ liệu lưu trữ tài liệu

**2.1** **Đặc điểm của kho tài liệu cơ sở dữ liệu**

* Tài liệu được xử lý trong cơ sở dữ liệu bằng khóa (duy nhất) đại diện cho tài liệu đó.
* Có một số loại để tổ chức dữ liệu là bộ sưu tập, thẻ, siêu dữ liệu không nhìn thấy được và thư mục thứ bậc.
* Trong trường hợp này, chúng ta có thể sử dụng tra cứu khóa-giá trị để truy xuất tài liệu.

**2.2 Ví dụ về cơ sở dữ liệu Kho lưu trữ tài liệu**

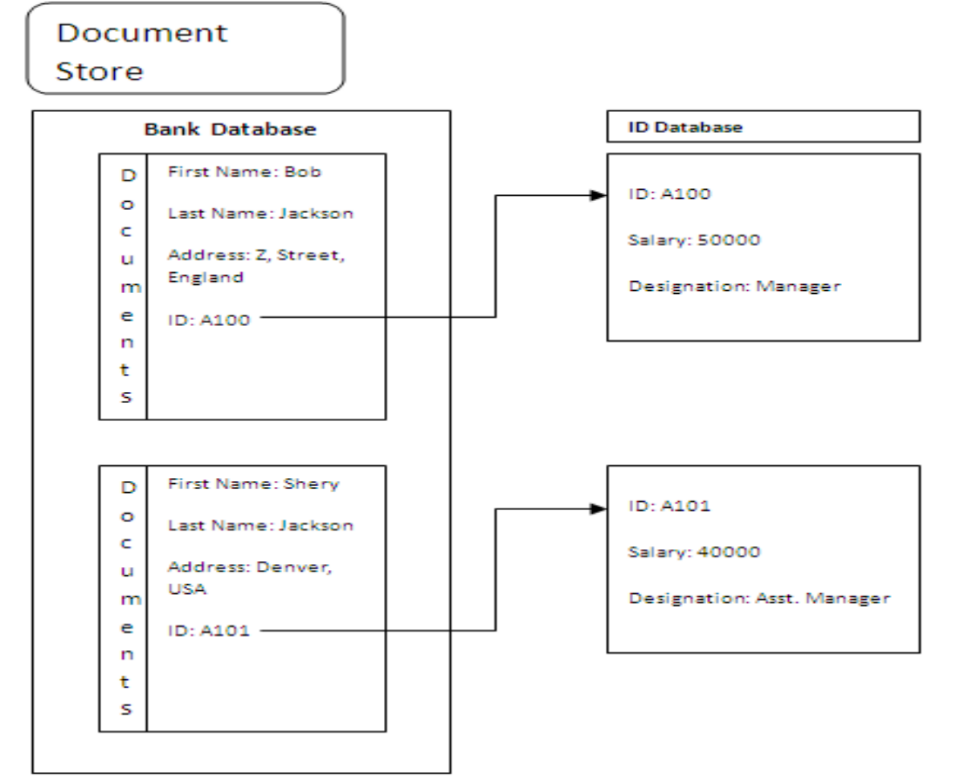
Hình 6 minh họa một ví dụ về Cơ sở dữ liệu Lưu trữ Tài liệu. Trong hình này, chúng tôi đang lấy một ví dụ về cơ sở dữ liệu ngân hàng, trong đó sử dụng ID làm khóa duy nhất. Và nó cũng bao gồm cơ sở dữ liệu ID chứa tất cả thông tin liên quan đến người đó, ai đang có ID cụ thể đó. ID trong cơ sở dữ liệu ngân hàng trỏ đến cơ sở dữ liệu ID bằng cách sử dụng con trỏ.

Cơ sở dữ liệu ngân hàng bao gồm hai tài liệu:

(1) Người có tên: Bob; Họ: Jackson; Địa chỉ: Phố Z, Anh; Mã số: A100.

(2) Một người có tên: sherry; Họ: Hilton; Địa chỉ: Denver, Hoa Kỳ; Mã số: A101.

Bây giờ chúng ta có thể coi ID của cả hai tài liệu là khóa được sử dụng để trao đổi thư từ. Khóa phải là duy nhất trong tự nhiên và là có giá trị. Chúng tôi không thể tạo khóa trùng lặp. Chúng tôi đã chọn ID làm khóa vì tên, địa chỉ không thể là duy nhất và ID của một con người luôn là duy nhất. Vì vậy, đó là lý do tại sao chúng tôi chọn ID làm khóa. Bây giờ chúng ta đang có một cơ sở dữ liệu khác là ID cơ sở dữ liệu. Cơ sở dữ liệu ID bao gồm tiền lương và chỉ định của người đó. Cơ sở dữ liệu ID bao gồm cơ sở dữ liệu ID A100, A101. Các tài liệu hiện có trong cơ sở dữ liệu ngân hàng bao gồm ID và trỏ đến dữ liệu ID cụ thể của nó. Đây là cách dữ liệu được tổ chức trong một cơ sở dữ liệu lưu trữ tài liệu. Dữ liệu được tổ chức dưới dạng bộ sưu tập, thẻ, siêu dữ liệu, v.v. Trong đó, chúng ta có thể truy xuất dữ liệu bằng cách sử dụng cặp khóa và giá trị. Điều này được minh họa rõ hơn trong ví dụ nơi chúng tôi có khóa A100 và khóa của nó các thuộc tính tương ứng bao gồm tiền lương của nó với chỉ định.



Hình 6: Cơ sở dữ liệu lưu trữ tài liệu

**3. Cơ sở dữ liệu dạng cột**

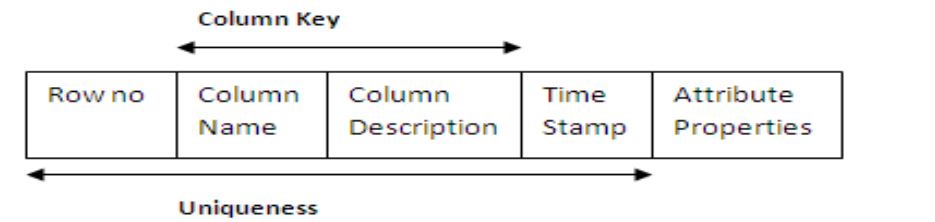
Cơ sở dữ liệu cột còn được gọi là cơ sở dữ liệu họ cột vì chúng là cơ sở dữ liệu hướng cột.

Có hai loại cơ sở dữ liệu hướng cột có chi tiết như sau:

**(1) Lưu trữ dữ liệu Wide-Column**:

Nó là một loại cơ sở dữ liệu NoSQL. Cửa hàng dữ liệu Cột rộng là những cơ sở dữ liệu được sử dụng để xử lý web,

truyền dữ liệu và tài liệu. Cấu trúc kho lưu trữ dữ liệu Column rộng như mô tả trong hình 7 bên dưới:



Hình 7: Cấu trúc kho lưu trữ dữ liệu cột Wide

Ý nghĩa của từng trường được cung cấp trong cấu trúc của cơ sở dữ liệu lưu trữ dữ liệu cột rộng được mô tả trong Hình 7 bên dưới.

|  |  |
| --- | --- |
| THUỘC TÍNH | Ý nghĩa |
| Cột không | Đó là một chìa khóa duy nhất trong tự nhiên. Nó có thể là một chuỗi hoặc một  con số. |
| Cột tên | Dữ liệu được lưu trữ trên cơ sở họ cột |
| Cột mô tả | Nó mô tả mục dữ liệu được lưu trữ |
| Dấu thời gian | Nó cho biết toàn bộ thời gian của trường hợp cụ thể |
| Kiểu dữ liệu | Giá trị hoặc thuộc tính liên quan đến khóa tương ứng đó. |

hình8: Ý nghĩa của các trường trong cấu trúc cơ sở dữ liệu cột rộng

(1) Cơ sở dữ liệu hướng cột:

Để hiểu cơ sở dữ liệu hướng cột, hãy lấy một ví dụ về cơ sở dữ liệu ngân hàng được đưa ra trong hình 9 có các trường thuộc tính là EmpID, Mức lương và chỉ định và các giá trị tương ứng với nó được mô tả trong cơ sở dữ liệu.

|  |  |  |
| --- | --- | --- |
| **EmpID** | **Salary** | **Designation** |
| 100 | 10,000 | Clerk |
| 200 | 20,000 | Assistant Manager |
| 300 | 30,000 | Manager |
| 400 | 40,000 | Zonal Head |

Hình 9: Ví dụ về cơ sở dữ liệu Ngân hàng

Biểu diễn cơ sở dữ liệu hướng hàng và cơ sở dữ liệu hướng cột:

• Cơ sở dữ liệu hướng hàng là những cơ sở dữ liệu trong đó tất cả các hàng được đặt cùng nhau từng cái một.

• Cơ sở dữ liệu hướng cột là những cơ sở dữ liệu trong đó tất cả các giá trị chứa cột được đặt cùng nhau.

Với sự trợ giúp của cơ sở dữ liệu được cung cấp ở trên, chúng tôi sẽ biểu diễn cơ sở dữ liệu theo hàng và cột, như thể hiện trong hình 10 bên dưới:

Graphical user interface, text, application, email

Description automatically generated Hình 10: Biểu diễn cơ sở dữ liệu hướng hàng và cột

3.1 Đặc điểm của cơ sở dữ liệu cột

(1) Cơ sở dữ liệu cột nhanh hơn cơ sở dữ liệu dựa trên hàng trong khi truy vấn.

(2) Trong cơ sở dữ liệu cột, việc gán đơn vị lưu trữ phải thực hiện cho từng cột.

(3) Trong DBMS dạng cột, chỉ các cột bắt buộc được đọc, vì vậy việc đọc sẽ nhanh hơn trong trường hợp này.

4. Cơ sở dữ liệu đồ thị

Cơ sở dữ liệu đồ thị dựa trên lý thuyết đồ thị. Nhìn chung, chúng ta thấy rằng đồ thị thường bao gồm các nút, thuộc tính và cạnh.

Cơ sở dữ liệu NoSQL Graph bao gồm:

(1)Các nút đại diện cho các thực thể

(2) Thuộc tính đại diện cho thuộc tính

(3)Các cạnh biểu thị các mối quan hệ [6].

Cấu trúc của cơ sở dữ liệu đồ thị như hình dưới đây

Diagram

Description automatically generated

Figure 11: Structure of graph database

4.1 Ví dụ về cơ sở dữ liệu Đồ thị

Ví dụ về Cấu trúc của cơ sở dữ liệu Đồ thị được hiển thị trong Hình 12. Nó giải thích cho chúng ta cách biểu diễn các nút, mối quan hệ và đặc tính. Ở đây có hai nút (được biểu thị bằng các vòng tròn) A và B đại diện cho NGÂN HÀNG và TÀI KHOẢN. Cả hai đang có mối quan hệ được biểu thị bằng một dòng mà ngân hàng có số lượng tài khoản. Tiếp theo là thuộc tính/thuộc tính/giá trị đại diện cho các loại tài khoản trong ví dụ được đưa ra trong hình sau. Các thuộc tính là Tiết kiệm, Hiện tại và R.D.

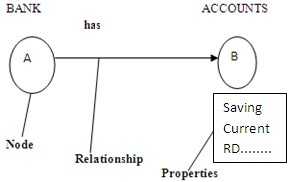
4.2 Characteristics of Graph databases

- Việc duyệt đồ thị được thực hiện với tốc độ không đổi không phụ thuộc vào tổng kích thước của đồ thị. Không có hoạt động thiết lập có liên quan làm giảm hiệu suất như đã thấy với các hoạt động nối trong RDBMS.

- Cơ sở dữ liệu đồ thị đang có hiệu suất cao trong bối cảnh truyền tải sâu của chúng.

- Chúng được sử dụng để tính toán đường đi ngắn nhất.

- Đây là những khả năng mở rộng. Nhưng sự phức tạp của nó tăng lên



Saving

Current

RD........

**Hình 12: Ví dụ về cấu trúc của cơ sở dữ liệu Graph**

## V. CHARACTERISTICS OF NoSQL

 -NoSQL không sử dụng mô hình dữ liệu quan hệ nên không sử dụng ngôn ngữ SQL.

- NoSQL lưu trữ khối lượng lớn dữ liệu.

 -Trong môi trường phân tán (dữ liệu trải rộng đến các máy khác nhau), chúng ta sử dụng NoSQL mà không có bất kỳ sự mâu thuẫn nào.

 -Nếu có bất kỳ lỗi hoặc hỏng hóc nào tồn tại trong bất kỳ máy nào, thì sẽ không có bất kỳ sự gián đoạn nào trong việc này.

 -NoSQL là cơ sở dữ liệu nguồn mở, tức là mã nguồn của nó có sẵn cho mọi người và được sử dụng miễn phí mà không cần bất kỳ chi phí chung.

 -NoSQL cho phép dữ liệu lưu trữ trong bất kỳ bản ghi nào mà nó không có bất kỳ lược đồ cố định nào.

 -NoSQL không sử dụng khái niệm thuộc tính ACID.

 -NoSQL có khả năng mở rộng theo chiều ngang dẫn đến hiệu suất cao theo cách tuyến tính.

 -Nó có cấu trúc linh hoạt hơn.

# VI.KẾT LUẬN VÀ CÔNG VIỆC TƯƠNG LAI

Mục đích chính của bài viết này là đưa ra một cái nhìn tổng quan về cơ sở dữ liệu NoSQL, về cách nó đã từ chối sự thống trị của SQL,với bối cảnh và đặc điểm của nó. Nó cũng mô tả các nguyên tắc cơ bản tạo thành cơ sở của cơ sở dữ liệu NoSQL như Định lý ACID, BASE và CAP. Thuộc tính ACID không được sử dụng trong cơ sở dữ liệu NoSQL vì tính nhất quán của dữ liệu để chúng ta biết SQL làm chậm tính nhất quán của dữ liệu như thế nào. Sau đó, trên cơ sở của định lý CAP, chúng tôi đã mô tả các loại khác nhau của Cơ sở dữ liệu NoSQL là cơ sở dữ liệu Khóa-Giá trị, Cơ sở dữ liệu Lưu trữ Tài liệu, Cơ sở dữ liệu dựa trên Cột và Cơ sở dữ liệu Đồ thị với sự giúp đỡ của một ví dụ. Ngoài tất cả những điều này, chúng tôi cũng đã mô tả đặc điểm, độ phức tạp và màn biểu diễn. Nghiên cứu sâu hơn đang diễn ra trong các công nghệ mới đang phát triển cho hoặc sau NoSQL đó là đa giác kiên trì, v.v.

**REFERENCES**

**[1.]** SilvanWeber, “ NoSQLDatabases ” http://www.christof-strauch.de /nosqldbs.pdf

[2.] Martin Fowler and Pramod Sadalage Rendered,

“NoSQLdbs- “, February8,2012,11:26, http://martinfowler.com/articles/nosql-intro.pdf

[3.] An Oracle White Paper, “Oracle NoSQL Database”,September2011,http://www.oracle.com/technetwork/database/nosqldb/learnmore/nosql-database-498041.pdf

[4.] Luis Ferreira Universidade do Minho, “Bridging the gap between SQL and NoSQL”, httpsikhote.files.wordpress.com201105artigo-mi-star1.pdf

[5.] Andrew J. Brust, Blue Badge Insights, Inc., “NoSQL and the Windows Azure platform”, April 25, 2011

[6.] DAMA - Philadelphia / Delaware Valley, the “Role of Data Architecture in NOSQL”, Wednesday January 11th,2012,http://www.damaphila.org/HaugheyNOSQL.pdf

[[