**SQL AND NoSQL DATABASES**

**Tóm tắt:** NoSQL (Không chỉ 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 tán, phi quan hệ, mã nguồn mở và có thể mở rộng theo chiều ngang (theo cách tuyến tính). NoSQL không tuân theo thuộc tính của ACID như chúng ta làm theo trong SQL. Trong bài nghiên cứu này, chúng ta đang khảo sát về NoSQL, nền tảng của nó, các nguyên tắc cơ bản như ACID, BASE và định lý CAP. Cũng trên cơ sở định lý CAP, nghiên cứu được thực hiện về các loại kho dữ liệu NoSQL khác nhau với các ví dụ, đặc điểm và ưu nhược điểm của NoSQL.

**Từ khóa: ACID, BASE, SQL, NoSQL, CAP, CURD.**

1. **GIỚI THIỆU:**

NoSQL là viết tắt của Not Only SQL. Nó được phát âm là “noseequel”. Nó là một trong những kiểu lưu trữ dữ liệu khác với cơ sở dữ liệu (đã được sử dụng trước đó) được sử dụng để lưu trữ lượng dữ liệu khổng lồ như dữ liệu trong facebook (không ngừng tăng lên từng ngày). NoSQL là một hệ quản trị cơ sở dữ liệu phi quan hệ (đôi khi được gọi là bắt nguồn từ cơ sở dữ liệu quan hệ), cơ sở dữ liệu truy xuất thông tin nhanh và có thể di động. NoSQL về cơ bản được bắt nguồn từ hệ thống cơ sở dữ liệu RDB. Cơ sở dữ liệu này 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, được phân tán về bản chất 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 tổ chức 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 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 sao chép dữ liệu. NoSQL có thể được biểu diễn một cách tượng trưng như trong hình 1:



Hình 1: Tượng trưng cho NoSQL

Hình 1 cho biết truy vấn tới cơ sở dữ liệu mà không có bất kỳ tương tác hoặc giao diện nào của ngôn ngữ SQL. Đường nghiêng trong hình cho thấy việc sử dụng cơ sở dữ liệu mà không 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 định dạng khác 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 lượng dữ liệu khổng lồ đã dẫn đến sự xuất hiện của nhiều công nghệ mới và một trong những sự nổi bật đó là NoSQL khá hữu ích trong việc lưu trữ dữ liệu. NoSQL (phi quan hệ) tương đối nhanh hơn so với cơ sở dữ liệu quan hệ .. Trước đây, trong SQL,chúng ta đang sử dụng ngôn ngữ Truy vấn để tìm nạp cũng như lưu trữ dữ liệu; đối với NoSQL, chúng ta lưu trữ các thực thể dữ liệu lớn bằng cách sử dụng các tài liệu ở định dạng XML (eXtensible Mark up Language). Về cơ bản, ngôn ngữ XML được sử dụng để lưu trữ dữ liệu có cấu trúc ở dạng con người có thể đọc được . Điều này sẽ được thảo luận kỹ hơn trong các phần sau.

1. **BACKGROUND:**

Vào năm 1960, cơ sở dữ liệu Đa giá trị được phát triển để thực hiện một số hoạt động 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 như vậy (phân cấp). Năm 1977, M[umps] được chấp nhận là ngôn ngữ tiêu chuẩn ANSI. Năm 1979, Ken Thompson đã phát triển DBM (Databases Manager) sử dụng các kỹ thuật băm để cung cấp cho người dùng khả năng truy xuất dữ liệu nhanh hơn. Vào năm 1980, một số lượng lớn các phiên bản kế thừa DBM đã được phát triển như sau:

* TDBM là một cơ sở dữ liệu kiểu DBM hỗ trợ các giao dịch nguyên tử (tuân theo các thuộc tính ACID A: Atomicity, C: Consistency, I: Isolation, D: Durability). 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.

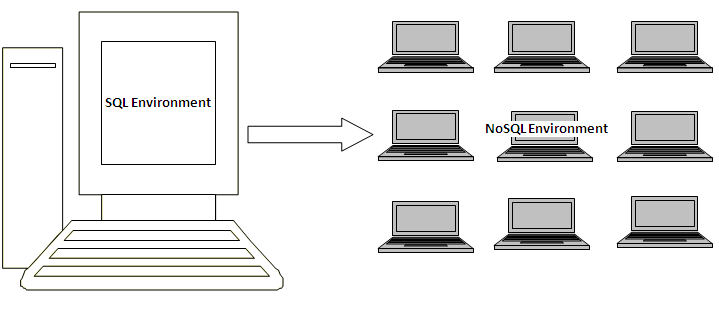
(2) TDBM cũng hỗ trợ các giao dịch nguyên tử lồng nhau sử dụng các hoạt động đa luồng. SDBM là cơ sở dữ liệu loại DBM được sử dụng vì một số lý do cấp phép.

* 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 được 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 trở thành một đối thủ cạnh tranh nặng ký với RDBMS (Relational Database Management Systems) [4]. Trong năm 2009 và 2010, đã có các hội nghị về NoSQL được tổ chức như NoSQL live, NoSQL eu, NoSQL East, v.v. Tên NoSQL này được Carlo Strozzi sử dụng lần đầu tiên vào năm 1998 làm tên tệp mà ông 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 để khắc phục điều này là mua nhiều máy (và biến chúng thành một cụm) để lưu trữ dữ liệu. Nó sẽ là một giải pháp rẻ hơn cũng như có thể mở rộng theo chiều ngang (linear way). 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.



**III. TUYẾN TÍNH CỦA NoSQL :**

1. **ACID :**

ACID là viết tắt của từ Atomicity(*tính nguyên tử*), Consistency(*nhất quán*), Isolation(*cô lập*) và Durability(*độ bền*). Khái niệm ACID về cơ bản xuất phát từ môi trường SQL. Nhưng trong NoSQL, chúng tôi sẽ không sử dụng khái niệm ACID vì tính năng nhất quán của SQL. Trong bài báo này, chúng tôi sẽ xem khái niệm ACID tạo ra các vấn đề như thế nào 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 duy trì sự nhất quán là cần thiết. Ví dụ: nếu có sự thay đổi trong một phần của bảng thì các thay đổi là cần thiết trong mỗi 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 mâu thuẫn được thực hiện.

1. **BASE:**

BASE là viết tắt của từ Basically(*cơ bản*), Available(*có sẵn*), Soft state(*trạng thái mềm*) và Eventual consistency(*tính nhất quán*). BASE ngược với ACID. Cơ sở dữ liệu NoSQL được phân chia giữa con đường từ ACID đến BASE. Sau khi nhất quán giao dịch, trạng thái mà chúng ta sẽ nhận được là trạng thái mềm không phải là trạng thái rắn. Trọng tâm chính đằng sau BASE là tính khả dụng vĩnh viễn.

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

1. **CAP:**

CAP là viết tắt của Consistency(*tính nhất quán*), Availability(*tính khả dụng*) và Partition tolerance(*phân vùng sức chịu đựng*). CAP về cơ bản là một định lý tuân theo ba nguyên tắc:

1. Dữ liệu có sẵn trên tất cả các máy phải giống nhau về mọi mặt và việc cập nhật thường xuyên được thực hiện trên tất cả các máy, tức là dữ liệu nhất quán.
2. Dữ liệu phải có sẵn vĩnh viễn và phải được truy cập mọi lúc, tức là có sẵn.
3. Trong khi máy bị lỗi hoặc bất kỳ lỗi nào trong cơ sở dữ liệu máy sẽ hoạt động tốt mà không dừng công việc của chúng, tức là phân vùng sức chịu đựng.
4. **VAI TRÒ CỦA KIẾN TRÚC DỮ LIỆU TRONG NoSQL:**

**A. Thành phần:** Có bốn thành phần trong khối xây dựng của nó.

[1.] **Modelling Language(Ngôn ngữ mô hình hóa)**: Nó mô tả cấu trúc của cơ sở dữ liệu và cũng xác định lược đồ dựa trên đó. Dữ liệu làđược lưu trữ dưới dạng hàng và cột sử dụng các 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 thích hợp.

[2.] **Database Structure(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.] **Database Query language(Ngôn ngữ truy vấn cơ sở dữ liệu):** Tất cả các hoạt động được thực hiện trên cơ sở dữ liệu đó là tạo, cập nhật, đọc và xóa (CURD).

[4.] **Transactions(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 lượng cơ sở dữ liệu. Có bốn loại kho dữ liệu mới khác nhau trong NoSQL.

**1. Giá trị cốt lõi của cơ sở dữ liệu:** Tên cơ sở dữ liệu giá trị khóa nghĩa là sự kết hợp của hai thứ là khóa và 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 các 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 về bản chất.
* Giá trị là một loại dữ liệu được trỏ bởi một khóa.

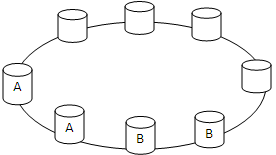
Cơ sở dữ liệu giá trị khóa dường như ở dạng bảng băm hoặc bảng tra cứu. Trong loại cơ sở dữ liệu này, chỉ có một cách để truy vấn đó là sử dụng khóa (duy nhất) và tất cả các khóa có thể đặt tên trong bất kỳ đối tượng dữ liệu nào và được sắp xếp theo thứ tự bảng chữ cái. Để có tính sẵn sàng cao hơn của 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.

|  |  |
| --- | --- |
| **Cơ sở dữ liệu Ngân hàng** | |
| Khóa | Giá trị |
| 1 | ID:1  Ngày tham gia: 15-07-1985  Chức danh: Thu ngân |
| 2 | ID:2  Ngày tham gia: 19-03-1982  Chức danh: Quản lý |
| 3 | ID:3  Ngày tham gia: 4-04-1988  Chức danh: Nhân viên lễ tân |

Trong hình đã cho, có hai cột đại diện cho khóa và giá trị. Ở đây khóa là duy nhất và đại diện cho các giá trị hoặc thuộc tính của chúng 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 chúng (theo thứ tự được sắp xếp) và dữ liệu cũng được sao chép ở dạng vòng . Điều này được thể hiện trong phần tiếp theo.

* 1. **Đặc điểm của cơ sở dữ liệu giá trị khóa:**
* 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 (Create(Tạo), Read(Đọc) và Update(Cập nhật) và Delete(Xóa)).
* Tất cả các mối quan hệ với dữ liệu được lưu trữ trong mã ứng dụng (không lan truyền rõ ràng).
  1. **Cơ sở dữ liệu giá trị khóa sử dụng (KV):**
* 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 lượng dữ liệu khổng lồ.
* 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 cũng như theo thứ tự bảng chữ cái. Điều này được thể hiện trong hình 4 dưới đây:



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

Kho lưu trữ cơ sở dữ liệu là những cơ sở dữ liệu NoSQL sử dụng bản ghi làm tài liệu. Loại cơ sở dữ liệu này lưu trữ các 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 bộ 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 công nghệ phân đoạn. Kho lưu trữ tài liệu Cơ sở dữ liệu không có lược đồ và không cố định về bản chất.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 một số 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 đang trỏ đến cơ sở dữ liệu của nó đang có một số quan hệ với 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 của nó. Điều này bao gồm một mảng cơ sở dữ liệu (ở dạng nhóm). Điều này sẽ rõ ràng hơn sau khi lấy một ví dụ được thảo luận bên dưới.

Diagram

Description automatically generated

Figure 5: Structure of document stores database

* 1. **Đặc Điểm Kho Tài Liệu:**

**Cơ sở dữ liệu :**

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

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 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 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) Một 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 liên kết . Khóa phải là duy nhất về bản chất và không phải là null. Chúng tôi không thể tạo khóa trùng lặp. Chúng tôi chọn ID làm khóa vì tên, địa chỉ không thể là duy nhất và ID của một 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 cơ sở dữ liệu ID. Cơ sở dữ liệu ID bao gồm tiền lương và chỉ định của người đó. Cơ sở dữ liệu ID bao gồm các cơ sở dữ liệu của 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 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 tôi có thể truy xuất dữ liệu bằng cặp khóa và giá trị.

Diagram

Description automatically generated

Figure 6: Document Store Databases

**3. Cơ sở dữ liệu dạng cột:**

Cơ sở dữ liệu cột cũng được gọi là cơ sở dữ liệu cột, bởi chúng là cơ sở dữ liệu về các cột. Có hai kiểu cơ sở dữ liệu về các cột, chi tiết của nó như sau:

* 1. **Lưu trữ dữ liệu Wide-Column:**

Nó là một loại cơ sở dữ liệu NoSQL. Kho lưu trữ 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:

Graphical user interface, text, application

Description automatically generated

Figure 7: Structure of Wide column data stores

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 dưới đây.

Table

Description automatically generated

Figure 8: Meaning of fields in structure of wide column databases

* 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 |

Figure 9: Example of Bank database

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

* Cơ sở dữ liệu hướng hàng là cơ sở dữ liệu trong đó tất cả các hàng được đặt cùng nhau từng cái một.
* Cơ sở dữ liệu hướng cột là cơ sở dữ liệu trong đó tất cả các giá trị chứa cột được đặt cùng nhau.
* Ta sẽ thể hiện cơ sở dữ liệu về các hàng và các cột như được trình bày dưới đây:

***Graphical user interface, text

Description automatically generated***

Figure 10: Representation of row and column oriented databases

**3.3 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 được 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)Các thuộc tính đại diện cho các thuộc tính

(3) Các cạnh đại diện cho các mối quan hệ .

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

A close-up of a sword

Description automatically generated with low confidence

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 thuộc 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 đều có mối quan hệ được biểu thị bằng một dòng trong đó ngân hàng có số lượng tài khoản. Tiếp theo là các 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 RD.

* 1. **Đặc điểm của cơ sở dữ liệu đồ thị:**
* Việc duyệt đồ thị được thực hiện với tốc độ không đổi không phụ thuộc vào tổng kích thước của đồ thị. Không có thao tác thiết lập nào liên quan làm giảm hiệu suất như đã thấy với thao tác 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.

Diagram

Description automatically generated

Figure 12: Example of structure of Graph databases

**V. ĐẶC ĐIỂM CỦA NoSQL:**

* NoSQL không sử dụng mô hình dữ liệu quan hệ do đó không sử dụng ngôn ngữ SQL.
* NoSQL lưu trữ khối lượng lớn dữ liệu.
* Trong môi trường phân tán (truyền dữ liệu đến các máy khác nhau), chúng tôi sử dụng NoSQL mà không có bất kỳ sự nhất quán nào.
* Nếu bất kỳ lỗi hoặc hỏng hóc nào tồn tại trong bất kỳ máy nào, thì trong máy này sẽ không có bất kỳ sự ngừng hoạt động nào.
* NoSQL là cơ sở dữ liệu mã 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 đang có cấu trúc linh hoạt hơn.

**VI. KẾT LUẬN VÀ CÔNG VIỆC TƯƠNG LAI:**

Mục đích chính của bài báo này là cung cấp một cái nhìn tổng quan về cơ sở dữ liệu NoSQL, về cách nó đã làm suy giảm sự thống trị của SQL, với nền tảng và đặc điểm của nó. Nó cũng mô tả các nguyên tắc cơ bản của nó tạo thành cơ sở của Cơ sở dữ liệu NoSQL như định lý về ACID, BASE và CAP. Thuộc tính ACID không được sử dụng trong cơ sở dữ liệu NoSQL vì tính nhất quán của dữ liệu, vì vậy chúng ta biết cách SQL làm giảm tính nhất quán của dữ liệu. Sau đó, trên cơ sở định lý CAP, chúng tôi đã mô tả các dạng 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 trợ giúp của một ví dụ. Ngoài ra, chúng tôi cũng đã mô tả các đặc điểm, độ phức tạp và sự hoạt động của chúng. Nghiên cứu sâu hơn đang diễn ra về các công nghệ mới đang phát sinh cho hoặc sau NoSQL, đó là các ngôn ngữ cơ sở dữ liệu bền bỉ khác, v.v.