



Trường Đại học Ngoại ngữ - Tin học Thành phố Hồ Chí Minh
Ho Chi Minh City University of Foreign Languages – Information Technology

CƠ SỞ DỮ LIỆU

Biên soạn: **Bộ môn Hệ Thống Thông Tin - HUFLIT**

GV hướng dẫn: **Trần Anh Duy**

Email: duyta@huflit.edu.vn



Trường Đại học Ngoại ngữ - Tin học Thành phố Hồ Chí Minh
Ho Chi Minh City University of Foreign Languages – Information Technology

CHƯƠNG 8

MỘT SỐ LOẠI CSDL KHÁC

Biên soạn: **Bộ môn Hệ Thống Thông Tin - HUFLIT**

GV hướng dẫn: **Trần Anh Duy**

Email: duyta@huflit.edu.vn

NỘI DUNG BUỔI HỌC

8.0. Giới thiệu

8.1. CSDL bán cấu trúc XML

8.2. CSDL NoSQL

8.3. CSDL NewSQL

8.4. Các ứng dụng

8.5. So sánh các loại CSDL.

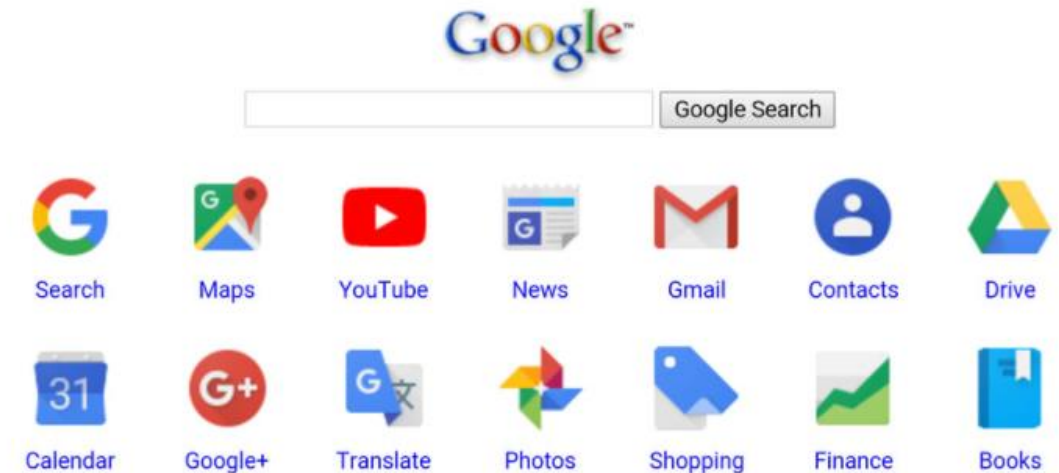
8.6. Một số hệ quản trị CSDL hiện đại

8.7. Chuyển đổi lược đồ cơ sở dữ liệu quan hệ sang NoSQL.

Bài tập

8.0. GIỚI THIỆU

- Ngày nay, nhiều công ty phải đối mặt với vấn đề **lưu trữ rất nhiều** dữ liệu và **nhiều loại** dữ liệu khác nhau.
 - Công ty như: Google, Amazon, Facebook, Twitter, ...
 - Loại dữ liệu: Chữ, số, hình ảnh, âm thanh, video, hoạt ảnh 2D, 3D, ...
 - Loại ứng dụng: Email, dữ liệu bán hàng, bản đồ, dòng trạng thái, ...
- Khái niệm **NOSQL** (Not Only SQL) ra đời.



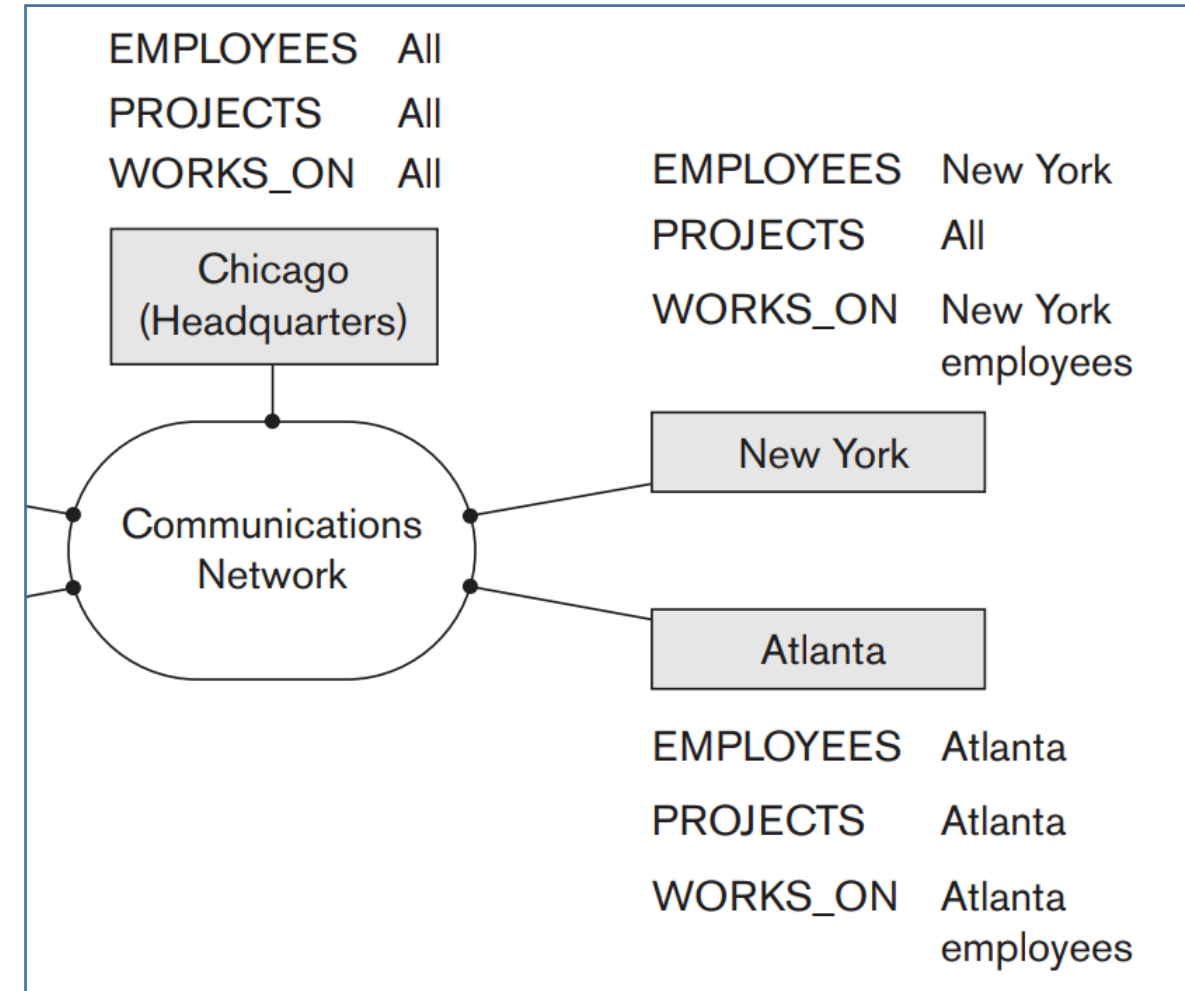
8.0. GIỚI THIỆU

- **NOSQL** (Not Only SQL) cho phép khắc phục những hạn chế của những ứng dụng CSDL truyền thống.
- Đặc điểm của **hệ thống NOSQL**:
 - Cơ sở dữ liệu phân tán.
 - Lưu trữ phân tán
 - Dữ liệu ***bán cấu trúc***
 - Tính sẵn sàng(availability)
 - Khả năng chia sẻ tập tin
 - Khả năng mở rộng
 - Khả năng truy cập dữ liệu cao.
- NOSQL hạn chế hỗ trợ *tính nhất quán (consistency)*, hạn chế hỗ trợ dữ liệu có cấu trúc và gần như không hỗ trợ câu truy vấn.

8.0. GIỚI THIỆU

Cơ sở dữ liệu phân tán (distributed database - DDB) bao gồm nhiều cơ sở dữ liệu liên quan về mặt logic và được bố trí ở những vị trí địa lý khác nhau; kết nối với nhau thông qua hệ thống mạng máy tính.

- Bao gồm nhiều máy tính, gọi là sites hoặc node, có kết nối với nhau.
- Có liên quan về mặt logic giữa các CSDL ở các node.
- Mỗi Node có thể khác nhau về mặt phần cứng hoặc phần mềm.

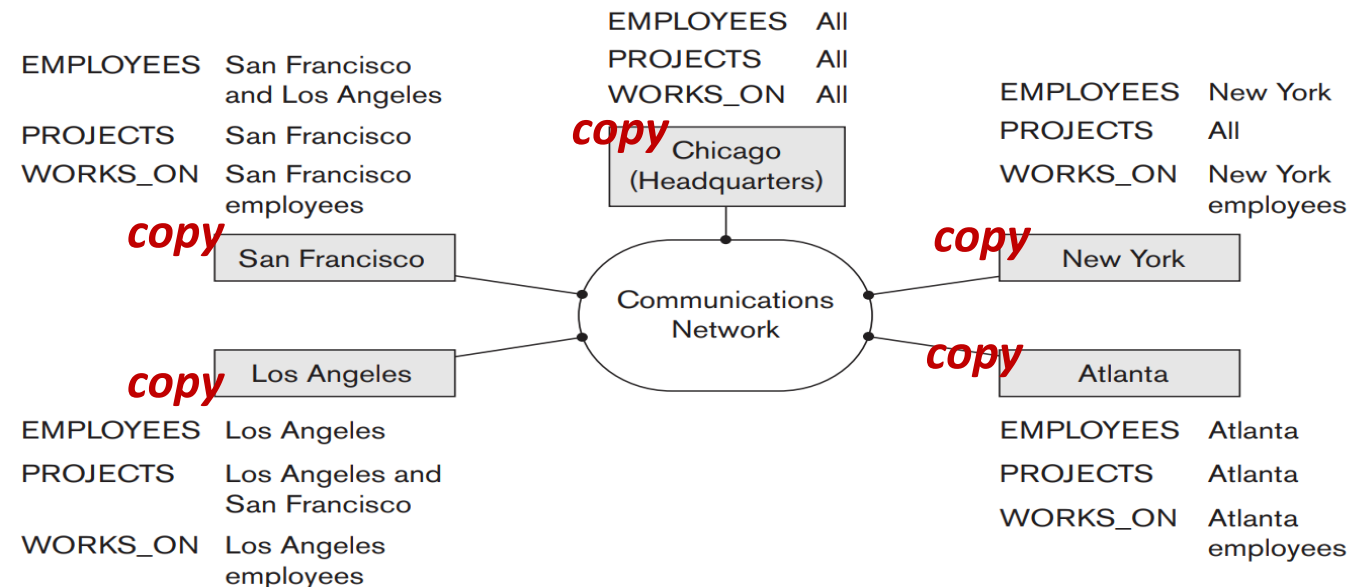


Ví dụ về CSDL phân tán

8.0. GIỚI THIỆU

Lưu trữ phân tán : CSDL sẽ được *nhân rộng(replication)* hoặc *sao chép* hoặc lưu trữ trên nhiều node khác nhau.

- Do đó, khi một node gặp sự cố thì người dùng vẫn có thể truy cập được dữ liệu (đảm bảo *tính sẵn sàng-availability*).



8.0. GIỚI THIỆU

Khả năng mở rộng: Đảm bảo khả năng lưu trữ khi dữ liệu tăng lên.

- **Mở rộng theo chiều ngang** (horizontal scalability): Khi dữ liệu tăng lên thì hệ thống phân tán sẽ được mở rộng bằng cách bổ sung thêm các node mới để lưu trữ dữ liệu.
- **Mở rộng theo chiều dọc** (vertical scalability) : Khi dữ liệu tăng lên, người ta sẽ tăng cường thêm bộ nhớ hoặc khả năng xử lý cho các node đang có

Khả năng chia sẻ tập tin: Trong nhiều ứng dụng NOSQL, các tập tin có thể có đến hàng triệu bản sao chép và các bản sao chép này dùng để đáp ứng khả năng truy cập liên tục của hàng ngàn người dùng trong một lúc.

Khả năng truy cập cao: Để tìm kiếm một đối tượng các hệ thống NOSQL sử dụng kỹ thuật băm (hashing) hoặc phân chia đối tượng theo khóa (range partitioning on object keys).

8.1. CSDL BẢN CẤU TRÚC XML

Dữ liệu có cấu trúc: là dữ liệu tuân theo một **định dạng** cụ thể nghiêm ngặt.

Ví dụ: Cho dữ liệu trong bảng phòng ban như sau:

Bảng phòng ban			
Tên phòng ban	Mã Phòng	Số điện thoại	Ngày thành lập
Nghiên cứu	5	0321456987	1998-05-22
Quản lý nhân sự	4	0321444555	1995-01-01
Giám đốc	1	0321666554	1981-06-19

Tên phòng
khác nhau

Mã phòng là
khóa ngoại

Trong đó:

- Tên phòng : Kiểu chuỗi tối đa 100 ký tự.
- Số điện thoại: Kiểu chuỗi 10 ký số
- Mã phòng: Kiểu số nguyên
- Ngày thành lập: Kiểu ngày

8.1. CSDL BÁN CẤU TRÚC XML

Có rất nhiều loại dữ liệu không thể xác định cấu trúc cụ thể.

- Thiết bị thu thập thông tin môi trường
- Dữ liệu phát sinh từ quá trình tương tác người dùng như click chuột, zoom, scroll ,...
- Dữ liệu đa phương tiện : hình ảnh, video, âm thanh, ánh sáng, ...
- Dữ liệu khai thác từ các chương trình khai thác dữ liệu.

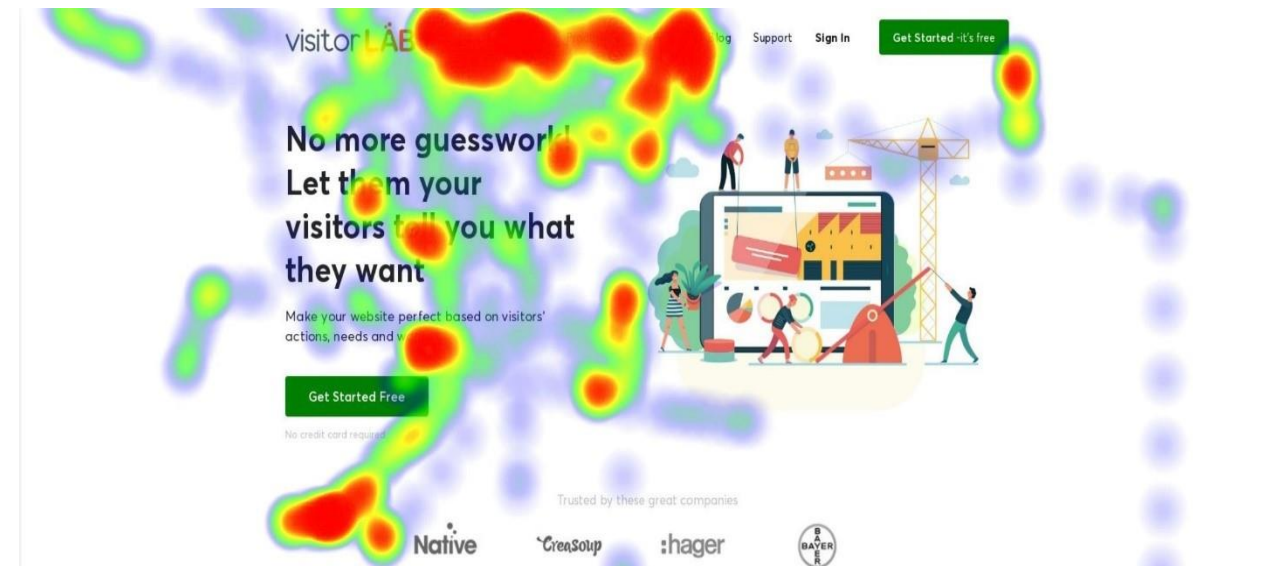


Ví dụ: Các thiết bị thu thập thông tin môi trường

8.1. CSDL BÁN CẤU TRÚC XML

Có rất nhiều loại dữ liệu không thể xác định cấu trúc cụ thể.

- Thiết bị thu thập thông tin môi trường
- Dữ liệu phát sinh từ quá trình tương tác người dùng như click chuột, zoom, scroll ,...
- Dữ liệu đa phương tiện : hình ảnh, video, âm thanh, ánh sáng, ...
- Dữ liệu khai thác từ các chương trình khai thác dữ liệu.

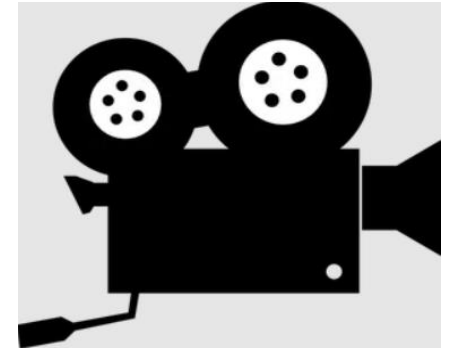


Ví dụ: thao tác click chuột vào trang Web

8.1. CSDL BÁN CẤU TRÚC XML

Có rất nhiều loại dữ liệu không thể xác định cấu trúc cụ thể.

- Thiết bị thu thập thông tin môi trường
- Dữ liệu phát sinh từ quá trình tương tác người dùng như click chuột, zoom, scroll ,...
- Dữ liệu đa phương tiện : hình ảnh, video, âm thanh, ánh sáng, ...
- Dữ liệu khai thác từ các chương trình khai thác dữ liệu.



VTV Khám Phá Thế Giới - Đế chế của Loài Kiến [Thuyết minh]
67K views • 1 year ago

VTV Khám Phá Thế Giới

VTV Khám Phá Thế Giới - Đế chế của Loài Kiến [Thuyết minh] ©Facebook VTV <https://www.facebook.com/vtvphatkhafathesegioi>



Bản tin CDTN Mùa Hè Xanh HUFLIT 2017 - Số 1
1.5K views • 5 years ago

HUFLIT Media Group

Sản phẩm do Đội hình Truyền Thông trực thuộc Ban Chỉ Huy Chiến Dịch Tình Nguyên Mùa H

Ví dụ: Video quay từ người dùng

8.1. CSDL BÁN CẤU TRÚC XML

- Một trong các mô hình lưu trữ dữ liệu bán cấu trúc là sử dụng *sơ đồ cây* (Tree).
- Trong sơ đồ cây **XML** bao gồm 3 thành phần sau: *thực thể*, *thuộc tính*, *mối quan hệ*.
- Mỗi **thực thể** được thể hiện bằng một tập các **thẻ (tag)** ký hiệu là **<tên thẻ>**
- Mỗi **thuộc tính** là một **thẻ con** chứa bên trong thực thể
- Thuộc tính cũng có thể là các **thông tin bên trong** một thẻ.

```
<PhongBan>
    <MaPhong> 5 </MaPhong>
    <TenPhong> Nghiên cứu </TenPhong>
    <SoDienThoai> 0321456987 </SoDienThoai>
    <NgayTL> 1998-05-22 </NgayTL>
</PhongBan>
```

```
<PhongBan
    MaPhong='5'
    TenPhong=' Nghiên cứu '
    SoDienThoai ='0321456987 '
    NgayTL=' 1998-05-22 '>
</PhongBan>
```


8.1. CSDL BÁN CẤU TRÚC XML

Ví dụ: Biểu diễn dữ liệu bảng phòng ban ở dạng XML

Bảng phòng ban			
Tên phòng ban	Mã Phòng	Số điện thoại	Ngày thành lập
Nghiên cứu	5	0321456987	1998-05-22
Quản lý nhân sự	4	0321444555	1995-01-01
Giám đốc	1	0321666554	1981-06-19

<CongTy>

<PhongBan>

<MaPhong> 5 </MaPhong>

<TenPhong> Nghiên cứu </TenPhong>

<SoDienThoai> 0321456987 </SoDienThoai>

<NgayTL> 1998-05-22 </NgayTL>

</PhongBan>

<PhongBan>

<MaPhong> 4 </MaPhong>

<TenPhong> Quản lý nhân sự </TenPhong>

<SoDienThoai> 0321444555 </SoDienThoai>

<NgayTL> 1995-01-01</NgayTL>

</PhongBan>

<PhongBan>

<MaPhong> 1 </MaPhong>

<TenPhong> Giám Đốc </TenPhong>

<SoDienThoai> 0321666554 </SoDienThoai>

<NgayTL> 1981-06-19</NgayTL>

</PhongBan>

</CongTy>

<CongTy>

<PhongBan MaPhong='5' TenPhong=' Nghiên cứu ' SoDienThoai ='0321456987 ' NgayTL=' 1998-05-22 ' > </PhongBan>

<PhongBan MaPhong='4' TenPhong=' Quản lý nhân sự ' SoDienThoai ='0321444555 ' NgayTL=' 1995-01-01 ' > </PhongBan>

<PhongBan MaPhong='1' TenPhong=' Giám Đốc ' SoDienThoai ='0321666554 ' NgayTL=' 1981-06-19 ' > </PhongBan>

</CongTy>

8.1. CSDL BÁN CẤU TRÚC JSON

- Một trong các mô hình lưu trữ dữ liệu bán cấu trúc là sử dụng *sơ đồ cây* (Tree).
- Trong sơ đồ cây **JSON** bao gồm 3 thành phần sau: *thực thể*, *thuộc tính*, *mối quan hệ*.
- Mỗi *thực thể* được thể hiện bằng một cặp *Key-Value*
- Mỗi *thuộc tính* là một cặp *Key-Value* chứa bên trong thực thể, phân cách bởi *dấu phẩy* (,)
- Nhiều thực thể có thể được mô tả bằng mảng một chiều của ngôn ngữ JavaScript

```
{  "PhongBan": {  
    "MaPhong":5,  
    "TenPhong": "Nghiên cứu",  
    "SoDienThoai": 321456987,  
    "NgayTL": "1998-05-22"  
  }  
}
```

Trong đó :

- Định dạng **"Key" : { Value }**
- **"PhongBan"**: {...}
- **"PhongBan"** là Key
- {...} là Value

Trong đó :

- Định dạng **"Key" : { Value }**
- **" TenPhong "**: "Nghiên cứu"
- **" TenPhong "** là Key
- "Nghiên cứu" là Value

8.1. CSDL BÁN CẤU TRÚC JSON

Ví dụ: Biểu diễn dữ liệu bảng phòng ban ở dạng XML

Bảng phòng ban			
Tên phòng ban	Mã Phòng	Số điện thoại	Ngày thành lập
Nghiên cứu	5	0321456987	1998-05-22
Quản lý nhân sự	4	0321444555	1995-01-01
Giám đốc	1	0321666554	1981-06-19

Dấu ngoặc vuông [...], [...] là mảng một chiều

```
{ "CongTy": {  
  "PhongBan": [  
    {  
      "MaPhong" : 5,  
      "TenPhong" : "Nghiên cứu",  
      "SoDienThoai" : 321456987,  
      "NgayTL" : "1998-05-22"  
    },  
    {  
      "MaPhong" : 4,  
      "TenPhong" : "Quản lý nhân sự",  
      "SoDienThoai" : 321444555,  
      "NgayTL" : "1995-01-01"  
    },  
    {  
      "MaPhong" : 1,  
      "TenPhong" : "Giám Đốc",  
      "SoDienThoai" : 321666554,  
      "NgayTL" : "1981-06-19"  
    }  
  ]  
}
```


BÀI TẬP

Bài 1: Hãy chuyển những bảng dữ liệu sau về dạng **XML**

Name	Ssn	Home_phone	Address	Office_phone	Age	Gpa
Benjamin Bayer	305-61-2435	(817)373-1616	2918 Bluebonnet Lane	NULL	19	3.21
Chung-cha Kim	381-62-1245	(817)375-4409	125 Kirby Road	NULL	18	2.89
Dick Davidson	422-11-2320	NULL	3452 Elgin Road	(817)749-1253	25	3.53
Rohan Panchal	489-22-1100	(817)376-9821	265 Lark Lane	(817)749-6492	28	3.93
Barbara Benson	533-69-1238	(817)839-8461	7384 Fontana Lane	NULL	19	3.25

Bài 2: Hãy chuyển những bảng dữ liệu sau về dạng **JSON**

<u>License_number</u>	Engine_serial_number	Make	Model	Year
Texas ABC-739	A69352	Ford	Mustang	02
Florida TVP-347	B43696	Oldsmobile	Cutlass	05
New York MPO-22	X83554	Oldsmobile	Delta	01
California 432-TFY	C43742	Mercedes	190-D	99
California RSK-629	Y82935	Toyota	Camry	04
Texas RSK-629	U028365	Jaguar	XJS	04

8.2. CSDL NOSQL

8.2.1. Key-value

8.2.2. Document database

8.2.3. Column family

8.2.4. Graph database

8.2.1. KEY-VALUE

- Những hệ thống sử dụng *mô hình dữ liệu đơn giản* dựa trên *tốc độ truy cập nhanh* bởi việc ánh xạ **key** đến **value** gắn liền với key (theo cơ chế tương tự bảng băm)
- *Value* có thể là: **mẫu tin (record)**, **đối tượng (object)**, **tài liệu (document)**, hoặc đối tượng dữ liệu phức tạp.
- Mô hình lưu trữ Key-Value tập trung vào *hiệu suất*, *tính sẵn sàng* và *khả năng mở rộng*.
- *Ví dụ*: DynamoDB, Redis, MemcacheDB, Berkley DB,...

8.2.2. DOCUMENT DATABASE

- *Document-based* hay còn gọi là hệ thống NOSQL mẫu tài liệu (document-oriented NOSQL systems) lưu trữ dữ liệu trong các tập tin văn bản.
- Tập tin lưu trữ dữ liệu được định dạng theo định dạng ***XML*** hoặc ***JSON***.
- *Ví dụ*: MongoDB, CouchDB,

```
{ "CongTy": {  
  "PhongBan": [  
    {  
      "MaPhong" : 5,  
      "TenPhong" : "Nghiên cứu",  
      "SoDienThoai" : 321456987,  
      "NgayTL" : "1998-05-22"  
    },  
    {  
      "MaPhong" : 4,  
      "TenPhong" : "Quản lý nhân sự",  
      "SoDienThoai" : 321444555,  
      "NgayTL" : "1995-01-01"  
    },  
    {  
      "MaPhong" : 1,  
      "TenPhong" : "Giám Đốc",  
      "SoDienThoai" : 321666554,  
      "NgayTL" : "1981-06-19"  
    }  
  ]  
}
```

8.2.3. COLUMN FAMILY

- Trong CSDL quan hệ, dữ liệu được lưu trữ trong các bảng.
 - Mọi thao tác truy vấn và xử lý dữ liệu đều thực hiện trên **1 hàng** trong cùng **1 thời điểm**.
- Trong khi đó, với CSDL dạng cột (*column-oriented database*) được lưu theo cột.

• Ví dụ:

EmployeeID	FirstName	LastName	Age	Salary
SM1	Anuj	Sharma	45	10000000
MM2	Anand		34	5000000
T3	Vikas	Gupta	39	7500000
E4	Dinesh	Verma	32	2000000

Dữ liệu lưu theo hàng

```
SM1 , Anuj , Sharma , 45 , 10000000
MM2 , Anand , , 34 , 5000000
T3 , Vikas , Gupta , 39 , 7500000
E4 , Dinesh , Verma , 32 , 2000000
```

Dữ liệu lưu theo cột

```
SM1 , MM2 , T3 , E4
Anuj , Anand , Vikas , Dinesh
Sharma , , Gupta , Verma ,
45 , 34 , 39 , 32
10000000 , 5000000 , 7500000 , 2000000
```

8.2.3. COLUMN FAMILY

Ưu điểm:

- Dễ dàng **thêm cột mới** (thuộc tính hoặc đối tượng mới).
- Dễ dàng **cập nhật dữ liệu** của một thuộc tính (một cột) trên tất cả các dòng dữ liệu.
- Dễ thực hiện các thao tác **tổng hợp** : SUM, COUNT, MAX, MIN, ... mà không cần duyệt qua tất cả các dòng dữ liệu.
- Ví dụ: Cassandra, HBase, Google BigTable, ...
- Ví dụ: Tính tổng lương của các nhân viên:

Dữ liệu lưu theo hàng

SM1, Anuj, Sharma, 45, 10000000
MM2, Anand, , 34, 5000000
T3, Vikas, Gupta, 39, 7500000
E4, Dinesh, Verma, 32, 2000000



10 m SM1, Anuj, Sharma, 45, 10000000



15 m MM2, Anand, , 34, 5000000



22.5 m T3, Vikas, Gupta, 39, 7500000



24.5 m E4, Dinesh, Verma, 32, 2000000

8.2.3. COLUMN FAMILY

Ưu điểm:

- Dễ dàng **thêm cột mới** (thuộc tính hoặc đối tượng mới).
- Dễ dàng **cập nhật dữ liệu** của một thuộc tính (một cột) trên tất cả các dòng dữ liệu.
- Dễ thực hiện các thao tác **tổng hợp** : SUM, COUNT, MAX, MIN, ... mà không cần duyệt qua tất cả các dòng dữ liệu.
- Ví dụ: Cassandra, HBase, Google BigTable, ...
- Ví dụ: Tính tổng lương của các nhân viên:

Dữ liệu lưu theo cột

SM1, MM2, T3, E4

Anuj, Anand, Vikas, Dinesh

Sharma, , Gupta, Verma,

45, 34, 39, 32

10000000, 5000000, 7500000, 2000000



10000000, 5000000, 7500000, 2000000

$10 + 5 + 7.5 + 2 = 24.5$ m

8.2.4. GRAPH DATABASE

- CSDL đồ thị (Graph databases) lưu trữ dữ liệu ở dạng đồ thị bao gồm các **nút** và **cạnh**.
- Trọng tâm của CSDL là thể hiện những **mối quan hệ** phức tạp **giữa các thực thể**.
- Các mối quan hệ có thể là : Sự liên hệ giữa các người dùng khác nhau trên Facebook, các tuyến đường trong Google Map, các kết nối trong một mạng máy tính, ...
- *Ví dụ*: Neo4j, FlockDB (Twitter)

Ưu điểm:

- Có thể tận dụng các thuật toán trên đồ thị như tìm đường đi ngắn nhất, tìm chu trình, ...
- Dễ dàng thể hiện sự liên hệ phức tạp giữa các thực thể.

8.3. CSDL NEWSQL

- NewSQL được biết đến như *phần mở rộng* của hệ thống quản trị *CSDL quan hệ*.
 - Cho phép đáp ứng *khả năng mở rộng* giống NOSQL
 - Vẫn duy trì *tính chất ACID*
 - Sử dụng *SQL* để tương tác dữ liệu.
- Ví dụ: Apache Trafodion, Google Spanner, ...



8.4. CÁC ỨNG DỤNG

NOSQL phù hợp với những môi trường ứng dụng sau:

- Yêu cầu **khối lượng dữ liệu** khổng lồ và tăng trưởng nhanh. Ví dụ như Facebook, Twitter, Gmail, Youtube, ...
- Yêu cầu **tốc độ truy xuất** nhanh. Ví dụ : ứng dụng thương mại điện tử, mạng xã hội, nội dung phát sóng trực tuyến, ...
- Yêu cầu **tổng hợp và phân tích** dữ liệu. Ví dụ: thống kê lượt truy cập người dùng, phân tích dữ liệu quảng cáo, phân tích dữ liệu tìm kiếm,
- Những ứng dụng với **tính sẵn sàng** cao. Ví dụ: thư viện điện tử, cổng thông tin quốc gia, hệ thống thông tin siêu thị, ...

8.5. SO SÁNH CÁC LOẠI CSDL

Đặc tính	Cơ sở dữ liệu NoSQL	Cơ sở dữ liệu quan hệ SQL
Hiệu suất	Cao	Thấp
Độ tin cậy	Thấp	Tốt
Khả dụng	Tốt	Tốt
Tính nhất quán	Thấp	Tốt
Tính sẵn sàng	Cao	Thấp
Lưu trữ dữ liệu	Tối ưu hóa cho dữ liệu lớn	Kích thước trung bình đến lớn
Khả năng mở rộng	Cao	Cao (nhưng chi phí cao hơn)
Hỗ trợ truy vấn SQL	Không hỗ trợ	Hỗ trợ theo chuẩn
Ràng buộc toàn vẹn	Phụ thuộc vào phần mềm kết nối	Đa phần các hệ thống phổ biến đều hỗ trợ

https://data.gov.vn/web/guest/news/-/asset_publisher/FRkblAs8yr3H/content/nosqlchweb

8.6. MỘT SỐ HỆ QUẢN TRỊ CSDL HIỆN ĐẠI

<https://www.mongodb.com/>

The screenshot shows the MongoDB website homepage. At the top, there is a navigation bar with the MongoDB logo, links for Products, Solutions, Resources, Company, and Pricing, a search icon, a Sign In link, and a green Try Free button. Below the navigation bar, there is a dark green hero section. On the left, it says "Build faster. Build smarter." followed by a paragraph: "Get your ideas to market faster with a developer data platform built on the leading modern database. Support transactional, search, analytics, and mobile use cases while using a common query interface and the data model developers love." Below this text are two buttons: "Start Free" and "Questions? Talk to us" with a right arrow. On the right side of the hero section, there is a stylized illustration of a document with a blue header, a white body with horizontal lines, and a green footer. The document is partially covered by a white box with a green border containing the text "_id: " followed by several horizontal lines. Below the document, there is a green box with a white circle and a white triangle, and a small white cube.

MongoDB. Products Solutions Resources Company Pricing

Search Sign In Try Free

NEW Get the latest and greatest with MongoDB 6.0 – Learn more >

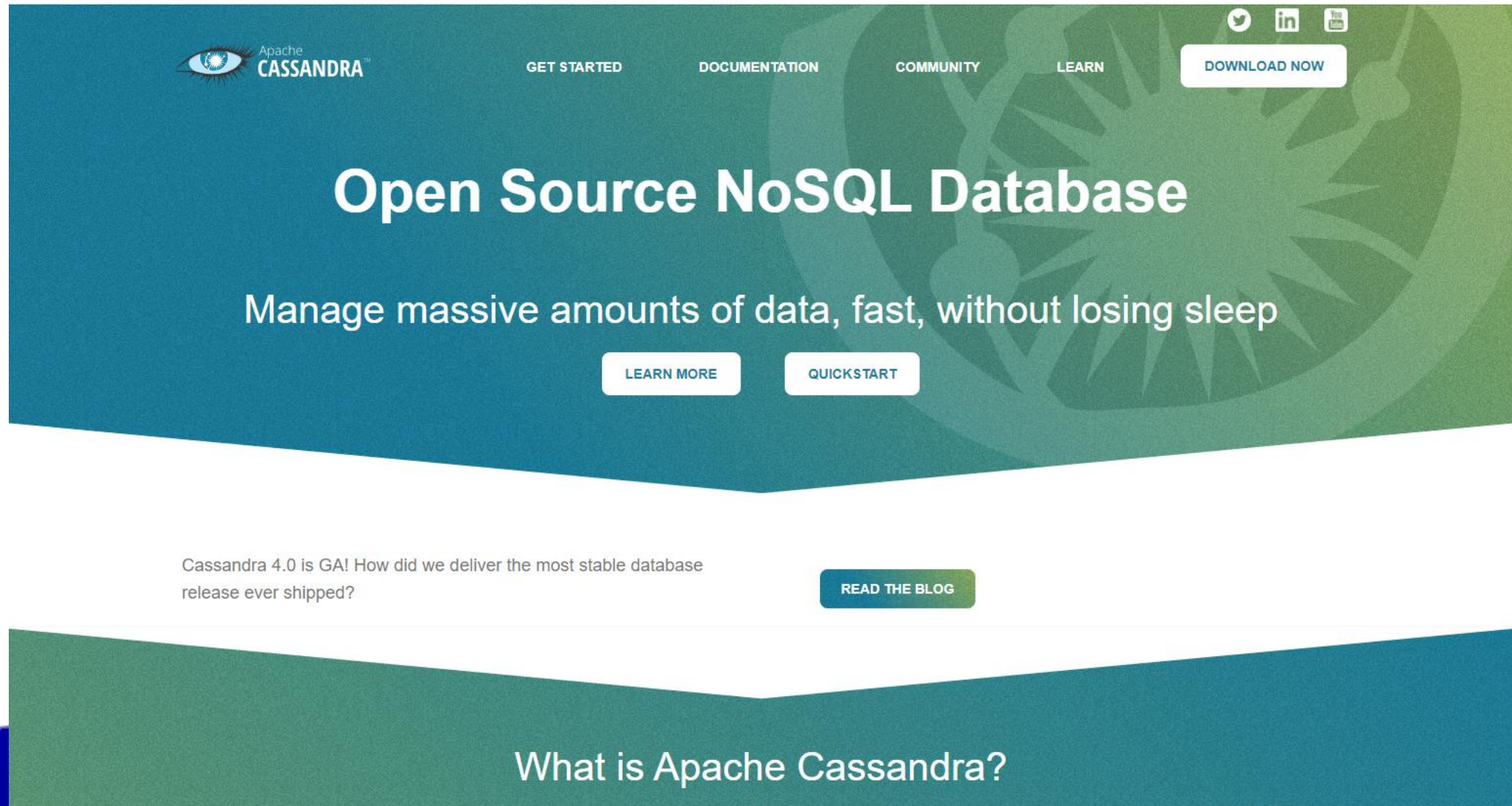
Build faster. Build smarter.

Get your ideas to market faster with a developer data platform built on the leading modern database. Support transactional, search, analytics, and mobile use cases while using a common query interface and the data model developers love.

Start Free Questions? Talk to us →

8.6. MỘT SỐ HỆ QUẢN TRỊ CSDL HIỆN ĐẠI

<https://cassandra.apache.org>



8.6. MỘT SỐ HỆ QUẢN TRỊ CSDL HIỆN ĐẠI

<https://neo4j.com>

NODES 2022: Neo4j Online Developer Education Summit | November 16-17 | Register now!

Partners ▾ Company ▾ Contact Us Support

neo4j Products Solutions Learn Developers Data Scientists [Get Started](#)

NEO4J GRAPH DATA PLATFORM

Blazing-Fast Graph, Petabyte Scale

With proven **trillion+** entity performance, developers, data scientists, and enterprises rely on Neo4j as the top choice for high-performance, scalable analytics, intelligent app development, and advanced AI/ML pipelines.

[Start Free for Developers](#) [Start Free for Data Scientists](#)

ICIJ

"[Neo4j] is a **revolutionary discovery tool** that's transformed our investigative journalism process."

ZURICH

"If I were to tell our investigators today that we were doing away with Neo4j, there would be a **huge outcry**."

Boston Scientific

Advancing science for life™

"Neo4j allows us to see relationships...that simply **would not be possible** with other tools."

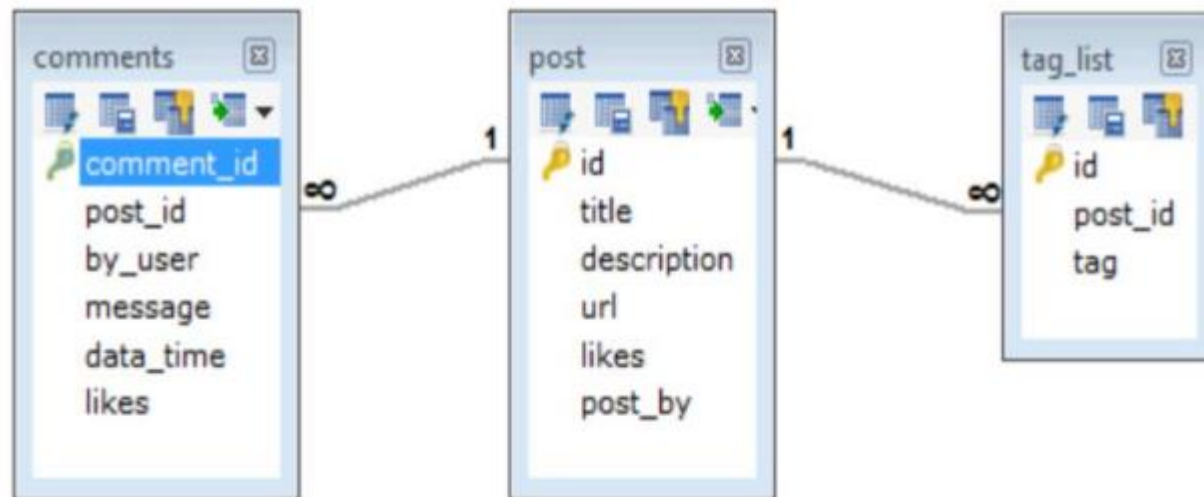
meredith

Meredith turned 14B anonymous users into 163M unique profiles, creating a better user experience and leading to **612% more web**

NODES 22
Neo4j Online Developer

November 16-17, 2022 [Save my Spot](#)

8.7. CHUYỂN ĐỔI LƯỢC ĐỒ CSDL QUAN HỆ SANG NOSQL.



Hình 1: Minh họa 3 bảng có khóa ngoại

post						tag_list		
id	title	description	url	like	post_by	id	post_id	tag
1	tit_1	des_1	url_1	100	aut_1	1	1	vacc
2	tit_2	des_2	url_2	200	aut_2	2	1	ast
3	tit_3	des_3	url_3	300	aut_3	3	1	Sputn
4	tit_4	des_4	url_4	400	aut_4			

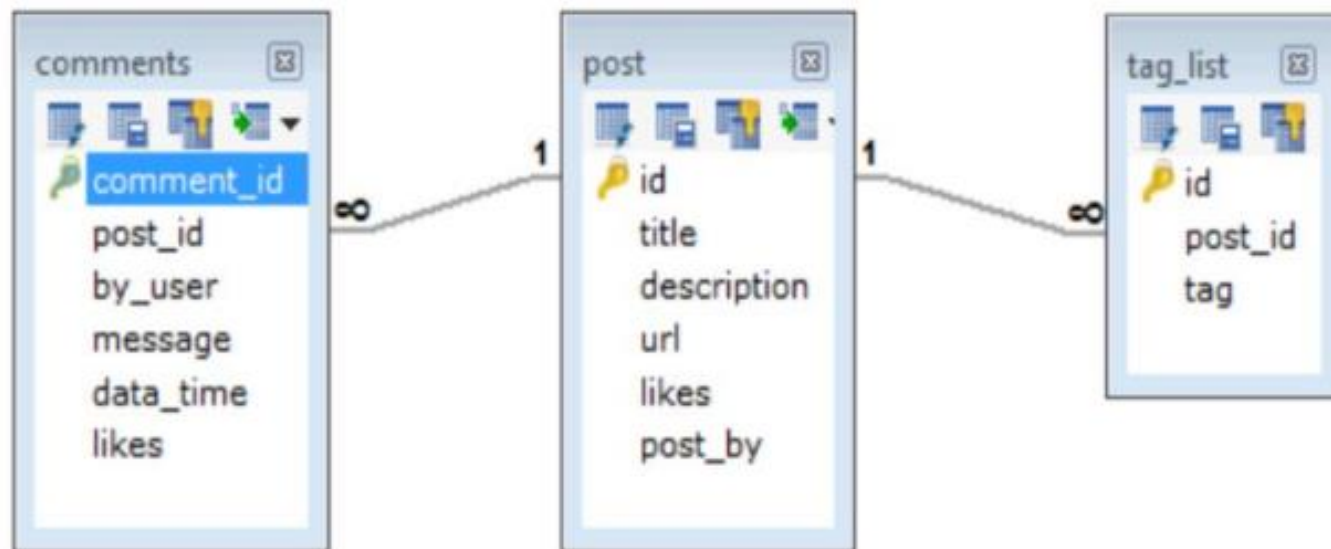
comments					
comment_id	post_id	by_user	message	data_time	likes
1	1	use_1	mess_1	10/11/12021	5

Hình 2: Hình minh họa CSDL

8.7. CHUYỂN ĐỔI LƯỢC ĐỒ CSDL QUAN HỆ SANG NOSQL.

Bước 1: Mọi đối tượng trong JSON đều bắt đầu bằng {}

```
{ }
```



8.7. CHUYỂN ĐỔI LƯỢC ĐỒ CSDL QUAN HỆ SANG NOSQL.

Bước 2: Chọn đối tượng trung tâm các quan hệ để xây dựng đối tượng gốc.

- Đối tượng được định dạng bằng cặp **key-value**
- **key** là **tên đối tượng** (tên bảng)
- **value** là **{ }**



```
{  
  "post": { }  
}
```

8.7. CHUYỂN ĐỔI LƯỢC ĐỒ CSDL QUAN HỆ SANG NOSQL.

Bước 2: Chọn đối tượng trung tâm các quan hệ để xây dựng đối tượng gốc.

- Do trong bảng có nhiều dòng dữ liệu hay có nhiều đối tượng post khác nhau nên thay value bằng ký hiệu **ngoặc vuông** đại diện cho **mảng một chiều []**

```
{  
  "post": [ ]  
}
```



post					
id	title	description	url	like	post_by
1	tit_1	des_1	url_1	100	aut_1
2	tit_2	des_2	url_2	200	aut_2
3	tit_3	des_3	url_3	300	aut_3
4	tit_4	des_4	url_4	400	aut_4

Nhiều
dòng

8.7. CHUYỂN ĐỔI LƯỢC ĐỒ CSDL QUAN HỆ SANG NOSQL.

Bước 3: Thêm đối tượng vào mảng một chiều các post.

- Một đối tượng được đặt trong cặp ngoặc nhọn { ... }
- Các thuộc tính và giá trị thuộc tính được tổ chức thành các cặp key:value
- **Key** là **tên thuộc tính** hay tên cột.
- **Value** là **giá trị** của thuộc tính.
- Các cặp key:value phân cách nhau bởi **dấu phẩy**

post					
id	title	description	url	like	post_by
1	tit_1	des_1	url_1	100	aut_1

```
{  
  "id" : 1 ,  
  "title" : "tit_1" ,  
  "description" : "des_1",  
  "url" : "url_1",  
  "like" : 100,  
  "post_by" : "aut_1"  
}
```

8.7. CHUYỂN ĐỔI LƯỢC ĐỒ CSDL QUAN HỆ SANG NOSQL.

Bước 3: Thêm đối tượng vào mảng một chiều các post.

post					
id	title	description	url	like	post_by
1	tit_1	des_1	url_1	100	aut_1
2	tit_2	des_2	url_2	200	aut_2
3	tit_3	des_3	url_3	300	aut_3
4	tit_4	des_4	url_4	400	aut_4

```
{
  "post": [
    {"id" : 1 , "title" : "tit_1" , "description" : "des_1", "url" : "url_1", "like" : 100,"post_by" : "aut_1"},
    {"id" : 2 , "title" : "tit_2" , "description" : "des_2", "url" : "url_2", "like" : 200,"post_by" : "aut_2"},
    {"id" : 3 , "title" : "tit_3" , "description" : "des_3", "url" : "url_3", "like" : 300,"post_by" : "aut_3"},
    {"id" : 4 , "title" : "tit_4" , "description" : "des_4", "url" : "url_4", "like" : 400,"post_by" : "aut_4"}
  ]
}
```

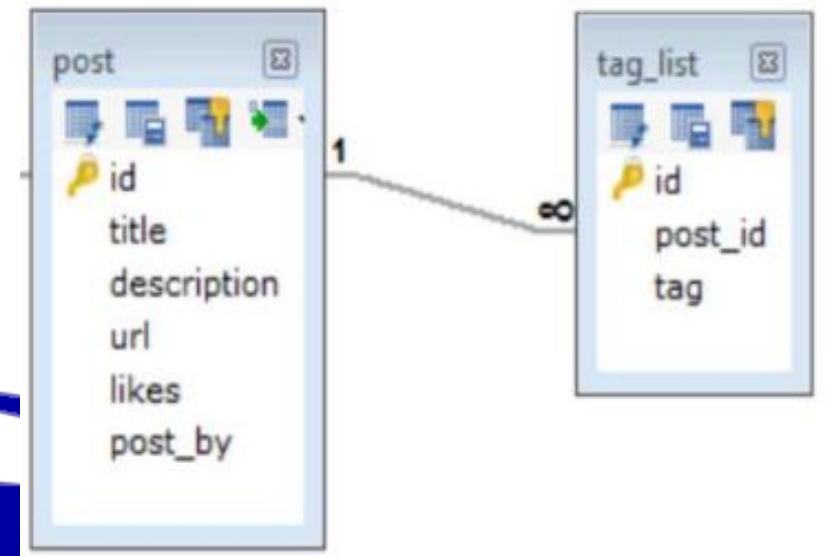
8.7. CHUYỂN ĐỔI LƯỢC ĐỒ CSDL QUAN HỆ SANG NOSQL.

Bước 4: Xây dựng quan hệ giữa các đối tượng

- **Mỗi quan hệ** giữa post và tag_list là **1-n**
- **Một** post sẽ có **nhiều** tag_list.
- Ta thêm một cặp **key:value** vào đối tượng post
- **Key** là tag_list (**tên bảng** có liên hệ)
- **Value** là **mảng một chiều** để chứa nhiều tag_list

```
{
  "id" : 1 ,
  "title" : "tit_1" ,
  "description" : "des_1",
  "url" : "url_1",
  "like" : 100,
  "post_by" : "aut_1",
  "tag_list": [ ]
}
```

tag_list		
id	post_id	tag
1	1	vacc
2	1	ast
3	1	Sputn

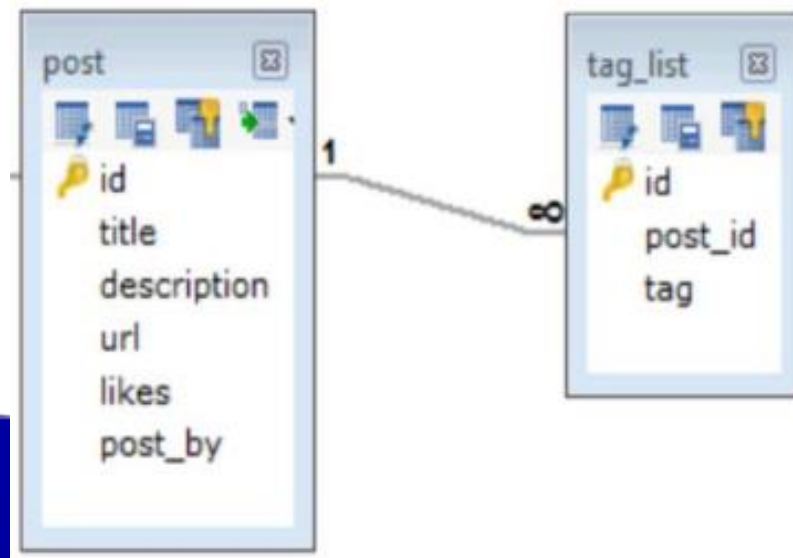


8.7. CHUYỂN ĐỔI LƯỢC ĐỒ CSDL QUAN HỆ SANG NOSQL.

Bước 4: Xây dựng quan hệ giữa các đối tượng

- Thêm dữ liệu tag_list vào.
- Mỗi tag_list sẽ là một đối tượng đặt trong cặp ngoặc nhọn {}
- Bên trong cặp ngoặc nhọn {} chứa các thuộc tính của tag_list
- Lưu ý: **Không** cần bổ sung thuộc tính **khóa ngoại** là post_id

tag_list		
id	post_id	tag
1	1	vacc
2	1	ast
3	1	Sputn



```
{  
  "id" : 1 ,  
  "title" : "tit_1" ,  
  "description" : "des_1",  
  "url" : "url_1",  
  "like" : 100,  
  "post_by" : "aut_1",  
  "tag_list": [  
    {"id":1, "tag":"vacc"},  
    {"id":2, "tag":"ast"},  
    {"id":3, "tag":"Sputn"}  
  ]  
}
```

8.7. CHUYỂN ĐỔI LƯỢC ĐỒ CSDL QUAN HỆ SANG NOSQL.

Bước 4: Xây dựng quan hệ giữa các đối tượng

- Trong trường hợp mỗi tag_list chỉ có **2 thuộc tính**
- Và **một** trong 2 thuộc tính là **khóa chính** định danh id
- Có thể **bỏ thuộc tính định danh** và chuyển thuộc tính còn lại thành các thành phần của mảng
- Khi đó, **vị trí** của các phần tử trong mảng tag_list được dùng **làm định danh** cho các phần tử thay cho id

```
{  
  "id" : 1 ,  
  "title" : "tit_1" ,  
  "description" : "des_1",  
  "url" : "url_1",  
  "like" : 100,  
  "post_by" : "aut_1",  
  "tag_list": [ "vacc", "ast", "Sputn" ]  
}
```

tag_list		
id	post_id	tag
1	1	vacc
2	1	ast
3	1	Sputn

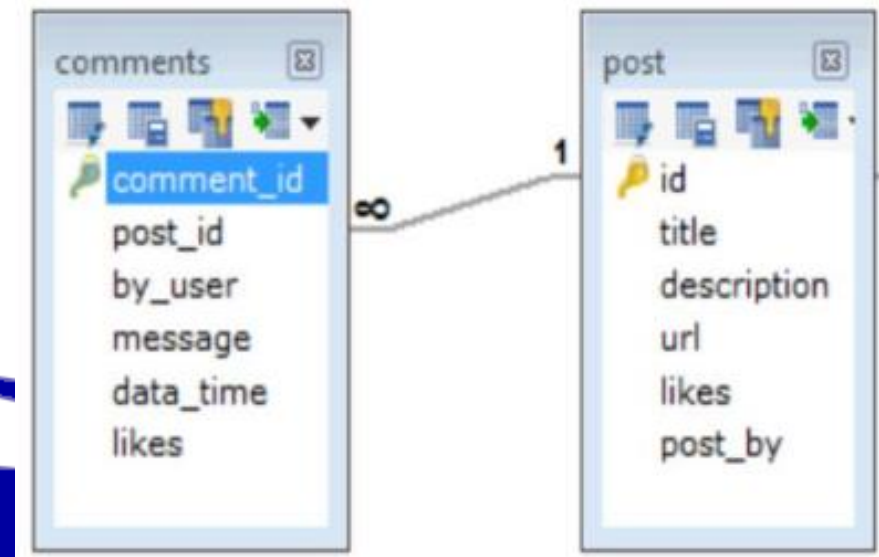
8.7. CHUYỂN ĐỔI LƯỢC ĐỒ CSDL QUAN HỆ SANG NOSQL.

Bước 4: Xây dựng quan hệ giữa các đối tượng

- **Mỗi quan hệ** giữa post và comments là **1-n**
- **Một** post sẽ có **nhiều** comments.
- Ta thêm một cặp **key:value** vào đối tượng post
- **Key** là comments (**tên bảng** có liên hệ)
- **Value** là **mảng một chiều** để chứa nhiều comments

```
{
  "id" : 1 ,
  "title" : "tit_1" ,
  "description" : "des_1",
  "url" : "url_1",
  "like" : 100,
  "post_by" : "aut_1",
  "tag_list" : [ "vacc", "ast", "Sputn" ],
  "comments" : []
}
```

comments					
comment_id	post_id	by_user	message	data_time	likes
1	1	use_1	mess_1	10/11/12021	5



8.7. CHUYỂN ĐỔI LƯỢC ĐỒ CSDL QUAN HỆ SANG NOSQL.

Bước 4: Xây dựng quan hệ giữa các đối tượng

- Thêm dữ liệu comment vào.
- Mỗi comment sẽ là một đối tượng đặt trong cặp ngoặc nhọn { }
- Bên trong cặp ngoặc nhọn { } chứa các thuộc tính của comment
- Lưu ý: **Không** cần bổ sung thuộc tính **khóa ngoại** là post_id

comments					
comment_id	post_id	by_user	message	data_time	likes
1	1	use_1	mess_1	10/11/12021	5

```
{
  "id" : 1 ,
  "title" : "tit_1" ,
  "description" : "des_1",
  "url" : "url_1",
  "like" : 100,
  "post_by" : "aut_1",
  "tag_list" : [ "vacc", "ast", "Sputn" ],
  "comments" : [
    { "comment_id" : 1,
      "by_user" : "user_1",
      "message" : "mess_1",
      "date_time" : "10/11/2021",
      "like" : 5
    }
  ]
}
```

8.7. CHUYỂN ĐỔI LƯỢC ĐỒ CSDL QUAN HỆ SANG NOSQL.

Kết quả

```
{
  "post": [
    {"id" : 1 , "title" : "tit_1" , "description" : "des_1", "url" : "url_1", "like" : 100, "post_by" : "aut_1",
      "tag_list" : [ "vacc", "ast", "Sputn" ],
      "comments" : [
        { "comment_id" : 1,
          "by_user" : "user_1",
          "message" : "mess_1",
          "date_time" : "10/11/2021",
          "like" : 5
        }
      ]
    },
    {"id" : 2 , "title" : "tit_2" , "description" : "des_2", "url" : "url_2", "like" : 200, "post_by" : "aut_2"},
    {"id" : 3 , "title" : "tit_3" , "description" : "des_3", "url" : "url_3", "like" : 300, "post_by" : "aut_3"},
    {"id" : 4 , "title" : "tit_4" , "description" : "des_4", "url" : "url_4", "like" : 400, "post_by" : "aut_4"}
  ]
}
```

8.7. CHUYỂN ĐỔI LƯỢC ĐỒ CSDL QUAN HỆ SANG NOSQL.

```
<?xml version="1.0" encoding="UTF-8" ?>
<root>
  <post>
    <id>1</id>
    <title>tit_1</title>
    <description>des_1</description>
    <url>url_1</url>
    <like>100</like>
    <post_by>aut_1</post_by>
    <tag_list>vacc</tag_list>
    <tag_list>ast</tag_list>
    <tag_list>Sputn</tag_list>
    <comments>
      <comment_id>1</comment_id>
      <by_user>user_1</by_user>
      <message>mess_1</message>
      <date_time>10/11/2021</date_time>
      <like>5</like>
    </comments>
  </post>
```

```
<post>
  <id>2</id>
  <title>tit_2</title>
  <description>des_2</description>
  <url>url_2</url>
  <like>200</like>
  <post_by>aut_2</post_by>
</post>
...
<post>
  <id>4</id>
  <title>tit_4</title>
  <description>des_4</description>
  <url>url_4</url>
  <like>400</like>
  <post_by>aut_4</post_by>
</post>
</root>
```

BÀI TẬP

Bài 3: Chuyển đổi lược đồ quan hệ sau sang JSON

Bảng Sản phẩm

Mã Sản Phẩm	Tên Sản phẩm	Giá bán
A1	Pepsi	5000
E2	Cocacola	6000
S3	Mirinda	7000

Bảng Hóa Đơn

Mã Hóa Đơn	Ngày bán	Mã Sản Phẩm	Số Lượng
HD01	1/11/2022	A1	5
HD01	1/11/2022	E2	8
HD01	1/11/2022	S3	7
HD02	1/11/2022	A1	6
HD02	1/11/2022	E2	9

Khóa ngoại: *Sanpham.Mã Sản phẩm* và *HoaDon.Mã Sản phẩm*

BÀI TẬP

Bài 4: Chuyển đổi lược đồ quan hệ sau sang JSON

Bảng Khách Hàng

Mã khách hàng	Tên khách hàng	Chiết khấu
KH01	Văn Quan	15%
KH15	Hồng Ngọc	16%
KH24	Sơn Trà	17%

Bảng Hóa Đơn

Mã hóa đơn	Ngày Lập	Mã khách hàng	Tổng tiền
HD01	1/11/2022	KH01	4500
HD02	1/11/2022	KH15	6000
HD03	1/11/2022	KH24	7000
HD04	2/11/2022	KH01	3000
HD05	2/11/2022	KH15	2000

Khóa ngoại: *KhachHang.MaKH* và *HoaDon.MaKH*

BÀI TẬP

Bài 5: Chuyển đổi lược đồ quan hệ sau sang JSON

Bảng nhân viên

Mã nhân viên	Tên nhân viên	Hệ số lương
NV01	Thế Ngọc	3.1
NV32	Minh Nghi	4.2
NV54	Cẩm Loan	5.3

Bảng Hóa Đơn

Mã hóa đơn	Ngày Lập	Mã nhân viên	Tổng tiền
SK01	3/11/2022	NV01	4500
SK02	3/11/2022	NV32	6000
SK03	3/11/2022	NV54	7000
SK04	2/11/2022	NV32	3000
SK05	2/11/2022	NV54	2000

Khóa ngoại: *NhanVien.MaNV* và *HoaDon.MaNV*

BÀI TẬP

Bài 6: Chuyển đổi lược đồ quan hệ sau sang JSON

Bảng nhân viên

Mã nhân viên	Tên nhân viên	Hệ số lương
NV01	Thế Ngọc	3.1

Bảng Sản phẩm

Mã Sản Phẩm	Tên Sản phẩm	Giá bán
A1	Pepsi	5000

Bảng Khách Hàng

Mã khách hàng	Tên khách hàng	Chiết khấu
KH01	Văn Quan	15%

Bảng Hóa Đơn

Mã hóa đơn	Ngày Lập	Mã nhân viên	Mã Sản Phẩm	Mã khách hàng	Số Lượng
SK01	3/11/2022	NV01	A1	KH01	5

Khóa ngoại:

- NhanVien.**MaNV** và HoaDon.**MaNV**
- KhachHang.**MaKH** và HoaDon.**MaKH**
- SanPham.**MaSP** và HoaDon.**MaSP**