Lecture 14

# Introduction to NOSQL

# The Other Aspect of Relational DBMS

RDBMSs put a lot of emphasis on keeping data consistent.

- Atomicity, Consistency, Isolation, Durability (ACID)
- Entire database is consistent at all times

Focus on consistency may affect flexibility and scalability

# No SQL

Not Only SQL

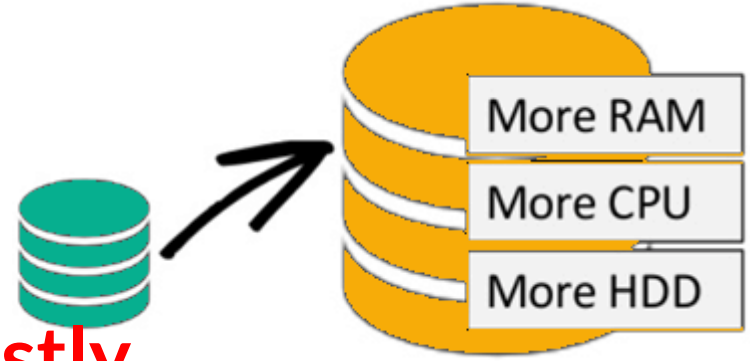
NoSQL Database is non-relational

Scales Out better than relational DB

System response time becomes slow when using RDBMS for massive volumes of data

Solution :

Scale Up : Upgrade Hardware




**Very Costly**

Scale Out : Distribute Database load on multiple servers

# Doesn't require object-relational mapping and data normalization

# Types of NOSQL Databases

Key - Value Pair Based

Document Based

Column Based

Graph Based

# Key - Value Pair Based

Data is stored in key/value pairs

Key-value pair storage databases store data as a hash table where each key is unique

| Key |
| --- |
| wilfried |
| seppe |
| bart |
| jeanne |


| Hash | Key |
| --- | --- |
| 01 | (wilfried,34) |
| 03 | (seppe,30) |
| 07 | (bart,46) |
| 08 | (jeanne,19) |

# Column Based

Column-oriented databases work on columns

Every column is treated separately.

Values of single column databases are stored contiguously.

They deliver high performance on aggregation queries like SUM, COUNT, AVG, MIN etc. as the data is readily available in a column.

Based on BigTable paper by Google

Other examples HBase, used in FB messenger

# Document Based

Stores and retrieves data as a key value pair

The value part is stored as a document.

The document is stored in JSON or XML formats.

| Col1 | Col2 | Col3 | Col4 |
|---|---|---|---|
| Data | Data | Data | Data |
| Data | Data | Data | Data |
| Data | Data | Data | Data |

**Document 1**

```
{
  "prop1": data,
  "prop2": data,
  "prop3": data,
  "prop4": data
}
```

**Document 2**

```
{
  "prop1": data,
  "prop2": data,
  "prop3": data,
  "prop4": data
}
```

**Document 3**

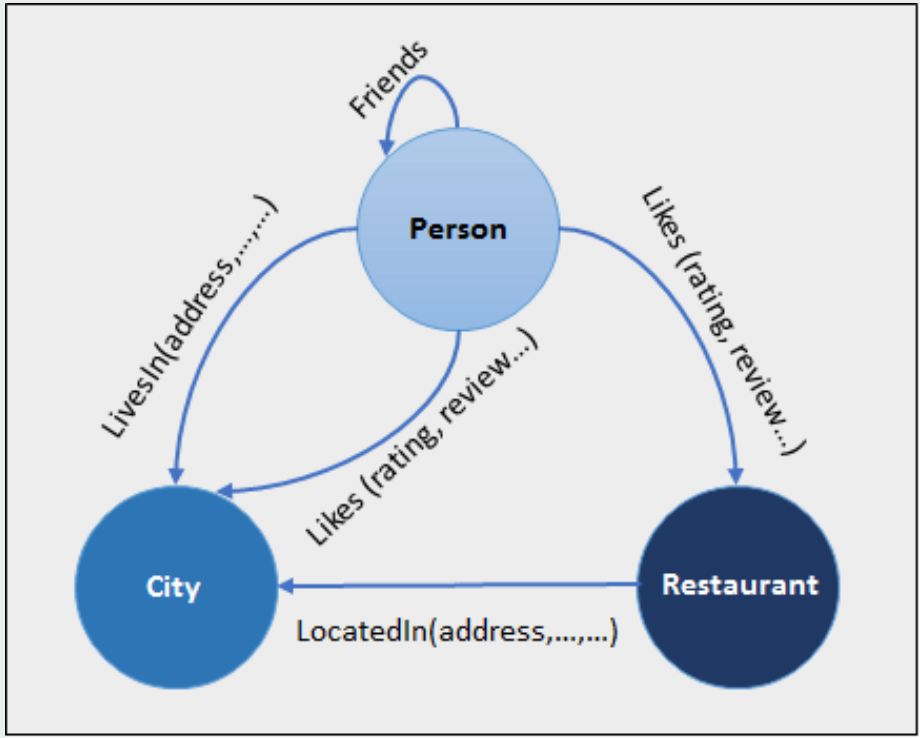
```
{
  "prop1": data,
  "prop2": data,
  "prop3": data,
  "prop4": data
}
```

# Graph Based

Stores entities as well the relations amongst those entities.

The entity is stored as a node with the relationship as edges

Graph based databases are mostly used for social networks, logistics, spatial data.



# More on NOSQL

NoSQL stores compromise consistency in favor of availability, speed and partition tolerance

NoSQL databases offer a concept of eventual consistency in which database changes are propagated to all nodes

# More on NOSQL

Queries for data might not return updated data immediately

OR

Might result in reading data that is not accurate which is a problem known as **stale reads.**

# Summary

NOSQL is a non-relational DMS

Either schema-free or has relaxed schemas

Avoids joins

Easy to scale

NOSQL can handle structured, semi-structured, and unstructured data with equal effect

# Questions?