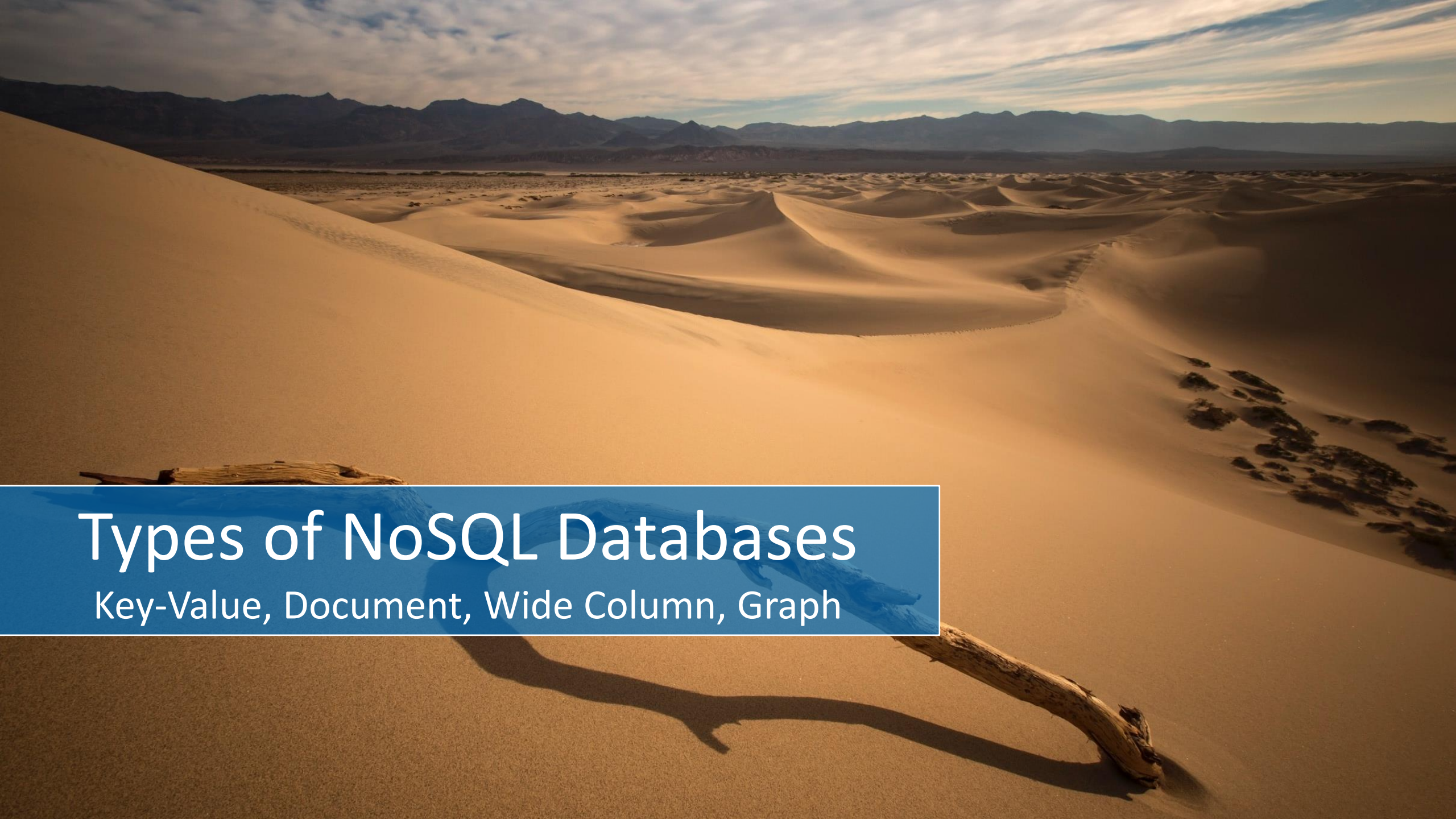
Cloud Computing for Beginners

Database Technologies

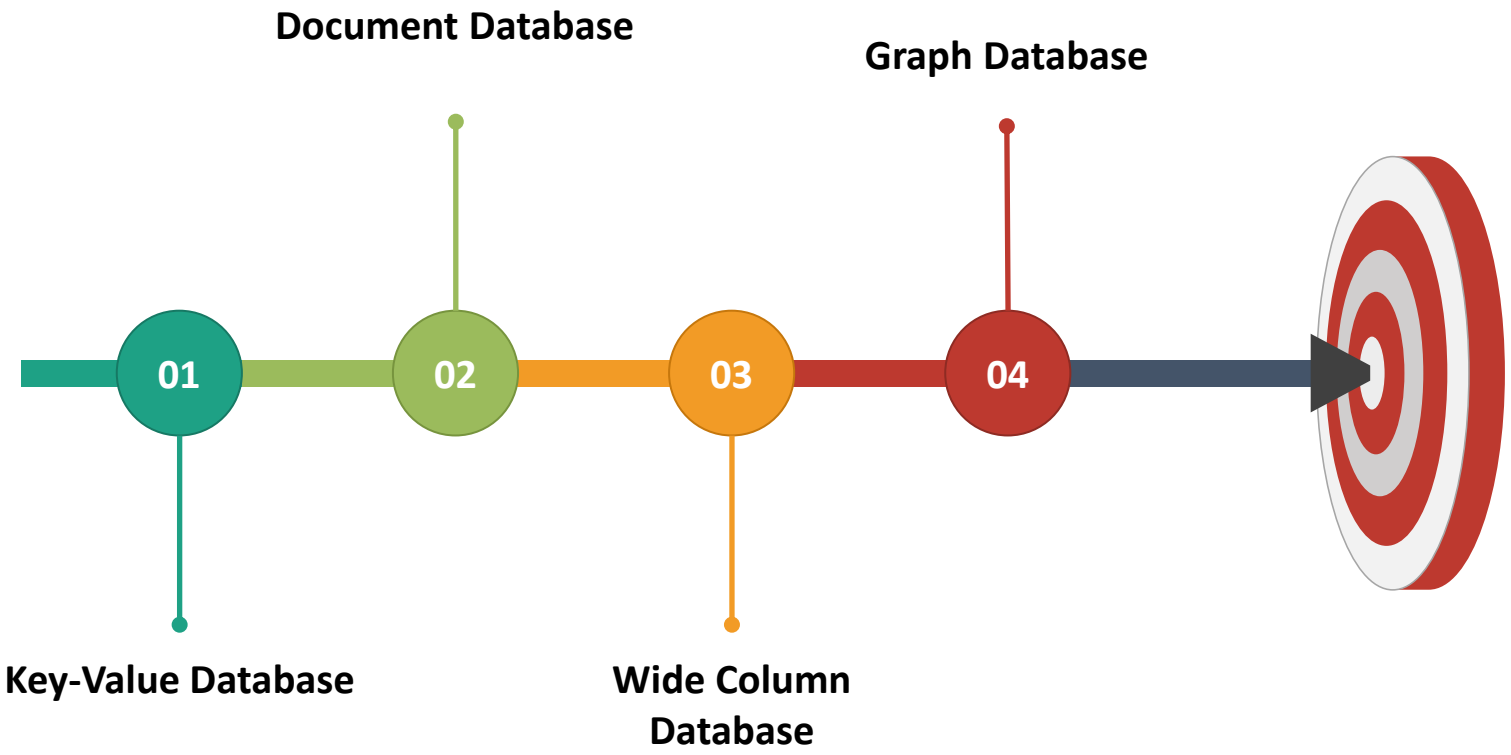
By Idan Gabrieli



Types of NoSQL Databases

Key-Value, Document, Wide Column, Graph

Types of NoSQL Databases



#1 - Key-Value Database (NoSQL)

- The **simplest type** of NoSQL databases
  - Simple data model
- Store data in a collection of **key-value pairs**
  - Key 1 → Value 1
  - Key 2 → Value 2
  - Key 3 → Value 3
- There are **no tables**, **no attributes** associated with tables
- Every value in a key-value pair can have completely different types of fields
  - Key 1 → **John Beri**
  - Key 2 → **12053-111-222**
  - Key 3 → **C:\List of Images\Image01.jpg**

#1 - Key-Value Database (NoSQL)

- **Operations**
  - **Put**(key, value)
  - **Get**(key)
    - Simple model → fast query → “give me the key and get the value”
    - No option to perform complex queries (e.g. filter by value)
  - **Delete** (key)
- **Where it is used?**
  - **Quick access** time is needed for large quantities of **simple data**
  - Simple lookups using the value of a key or by a range of keys
  - Not optimized for query by value/s
    - WHERE name=“Idan”
- **Use Cases**
  - Frequently implemented as a **caching solution** (in-memory + key value)

#1 - Key-Value Database (NoSQL)

- **Website shopping cart**
  - Each user login and starts to add products to a shopping cart
  - The data is relatively **simple**
  - High volumes - millions of customers playing with their shopping carts
  - Key-value database
    - Key = Customer ID
    - Values = Items added to the shopping cart
- **Key-Value Databases solutions**

redis

Amazon DynamoDB

Azure Cosmos DB

MEMCACHED

#2 - Document Database (NoSQL)

What is Document Database?

- **Document Database**
  - Group **key-values pairs** into objects called **documents**
  - Each document is a file that is encoded using some file format
    - **JSON**, XML, or YAML
  - Each document is given a unique ID, like a key to pull out a complete document
- **Document in a JSON Format**

```
{
        “FirstName”: “Idan”,
        “Phone”:”1111-2222”,
        “Address”:”5 Great Place”
}
```

#2 - Document Database (NoSQL)

What is Document Database?

- **Document Database**
  - Applications can retrieve documents using the **document key**
  - There are no “empty fields” in a document
  - A document database has **no fixed schema**
  - Each document in a collection of **unique fields**
  - Add/remove data from specific documents **without affecting** the other documents

#2 - Document Database (NoSQL)

Use Case Example

- **Managing a Product Catalog**

```
{
    "Product ID": "111",
    "Product Name": "Car Toy",
    "Product Desc": "Remote Car Toys for Kids",
    "Age Range": "5-10",
    "Distance": "25m",
    "Supplier-ID": "200"
    "Supplier-ID": "Cars for Kids LTD"
}
```

```
{
    "Product ID": "222",
    "Product Name": "Cordless Drill",
    "Product Desc": "Drilling tool for experts",
    "Volts": "20",
    "Speed-RPM": "750",
    "Supplier-ID": "300"
    "Supplier-ID": "Black & Decker"
}
```

#2 - Document Database (NoSQL)

Well-known Document databases

Amazon DynamoDB

Couchbase

Azure Cosmos DB

mongoDB®

Firebase

CouchDB

#3 - Wide Column Database (NoSQL)

What is Wide Column Database?

- **Wide Column Database**
  - A type of NoSQL database
  - Also called
    - Column-oriented database
    - Columnar database
    - Column-family database
  - **A “row”** is a list of values related to the **same column**

#3 - Wide Column Database (NoSQL)

Before the Age of Wide Column Databases

- Traditional relational databases are **row-oriented**
  - Each row will have a **row key**
  - Each **field** within the row will be stored together in a table
  - Adding **more rows** in a table
  - A single column represents an **attribute** or a field inside the row

Number	First Name	Last Name	Phone
1	Idan	Gabrieli	111-222
2	Bob	Marley	333-444
3	Maria	Carey	555-666

#3 - Wide Column Database (NoSQL)

Before the Age of Wide Column Databases

- Traditional relational databases are **row-oriented**
  - **Storage Level**
    - One-dimensional, big line
      - 1, Idan, Gabrieli, 111-222, 2, Bob, Marley, 333-444, 3, Maria, Carey, 555-666
    - When query only two columns – **First** and **Last Name**
      - It will scan **all rows and all columns**
        - 1, Idan, Gabrieli, 111-222, 2, Bob, Marley, 333-444, 3, Maria, Carey, 555-666
      - Regardless of **which columns** are required
    - For example - assuming
      - 1 x Row = 10 Kbytes, 1M rows → 10 Gigabytes
      - 2 columns = 2 Kbytes, 1M rows → 2 Gigabytes

#3 - Wide Column Database (NoSQL)

What is Wide Column Database?

- **Wide Column Database**
  - Column-oriented database\columnar database\column-family database
  - **Flips this storage system**
  - **A row** is a list of values related to the **same column**

Row-oriented

Number	First Name	Last Name	Phone
1	Idan	Gabrieli	111-222
2	Bob	Marley	333-444
3	Maria	Carey	555-666

Column-oriented

Number	1	2	3
First Name	Idan	Bob	Maria
Last Name	Gabrieli	Marley	Carey
Phone	111-222	333-444	555-666

#3 - Wide Column Database (NoSQL)

What is Wide Column Database?

- **Wide Column Database**
  - **Storage Level**
    - One-dimensional, big line
      - 1, 2, 3, Idan, Bob, Maria, Gabrieli, Marley, Carey, 111-222, 333-444, 555-666
    - Query only two columns – First and Last Name
      - 1, 2, 3, Idan, Bob, Maria, Gabrieli, Marley, Carey, 111-222, 333-444, 555-666
    - Columns are organized into **groups called families**
  - **Reduce the amount of data** to load from the disk
    - Improve performance and query time
    - An important factor for big data analysis
    - Well-suited for OLAP-oriented workloads
  - **Better Data Compression**
  - **Scalability**
    - Designed to scale "out" using distributed cluster

#3 - Wide Column Database (NoSQL)

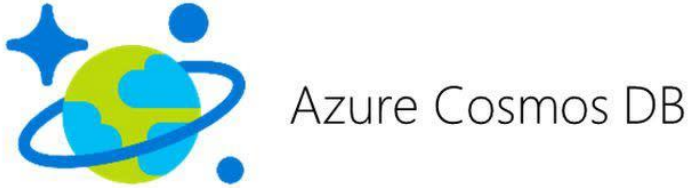
What is Wide Column Database?

- **Disadvantages**
  - **Writing data** into a column-oriented database is a **slow process**
  - Transactions must be separated into columns and compressed
  - **Row-oriented** database is a much better solution for OLTP (Transactions)

#3 - Wide Column Database (NoSQL)

Well-known Wide Column databases

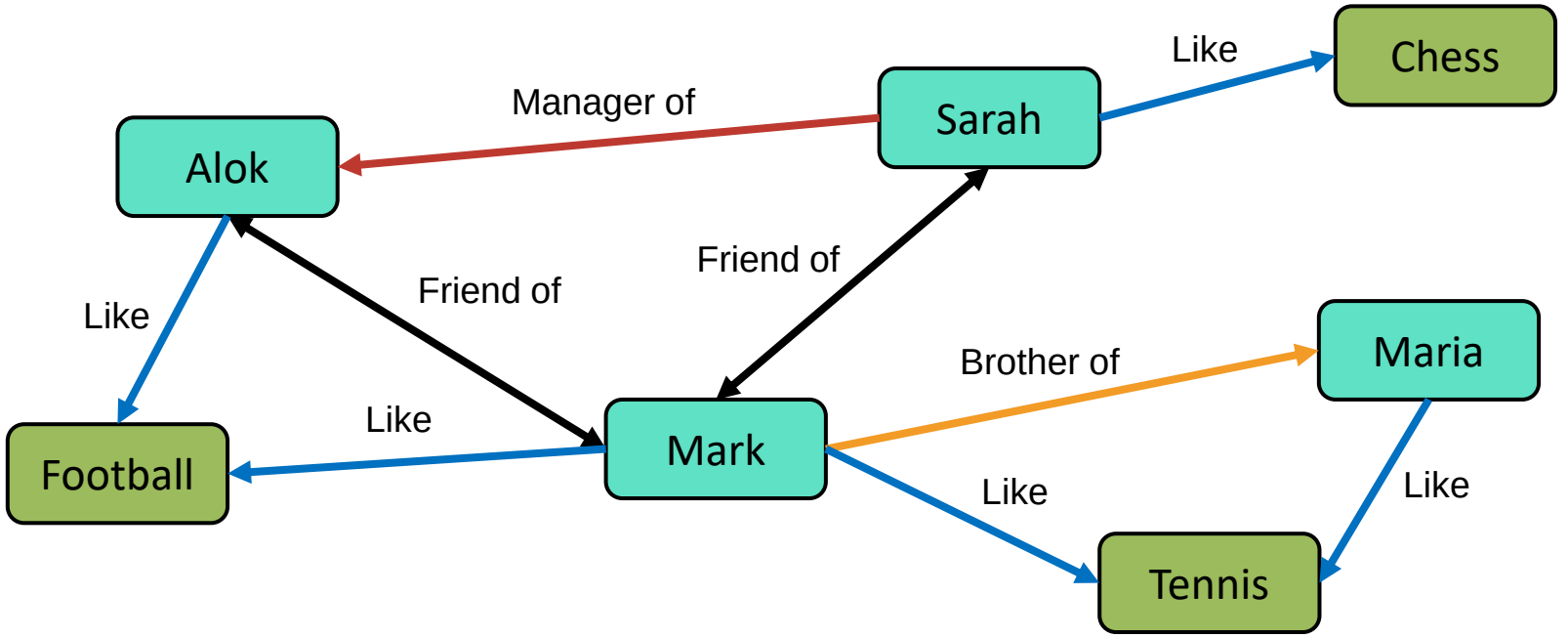
Google Cloud Bigtable



#4 - Graph Database (NoSQL)

What is Graph Database?

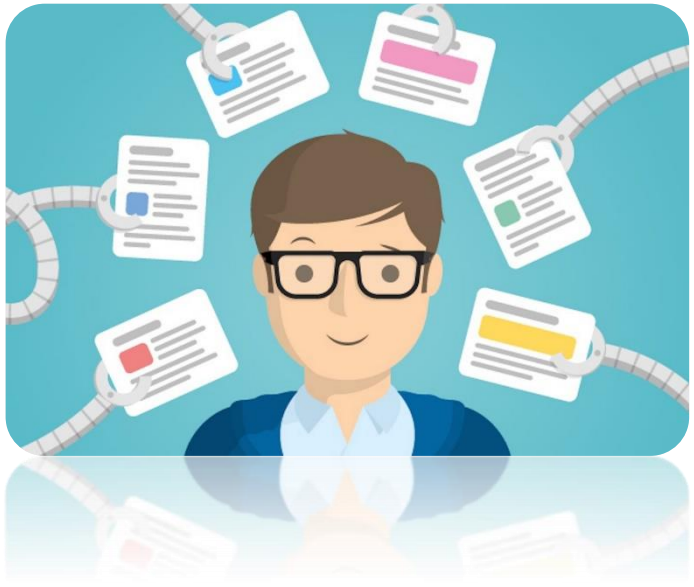
- **Represents data as a Graph**
  - Connections between data items are as important as the data items themselves
  - Two types of information
    - **Nodes** - represents an entity (user, product, habit…)
    - **Edges** - represents how two nodes are associated



#4 - Graph Database (NoSQL)

What are the Typical Use Cases?

- **Recommendation Engines**
  - E-commerce Websites
  - Stores the complex relationships between information categories
  - E.g. : Customer details, customer interests, purchase history, friends, etc.
  - Used to make **product recommendations**
- **Fraud Detection**
  - Relations between small pieces of information to detect patterns



#4 - Graph Database (NoSQL)

Well-known Graph databases