Module 6: Database Selection

Text chapter 15

University of Illinois
at Springfield

What is Database to Choose to Use

In general, when choosing a databases, consider:

- The types of queries used to read, insert, update, and delete data

- The volume of reads and writes

- Tolerance for inconsistent data in replicas

- The nature of relations between entities and how that affects query patterns

- Availability and disaster recovery requirements

- The need for flexibility in data models

- Latency requirements

What is Database to Choose to Use

- **Key-value databases are well suited to applications that have frequent small reads and writes along with simple data models.**

- **Key-value databases generally have simple query facilities that enable you to look up a value by its key.**
  - **Some key-value databases support search features that provide for somewhat more flexibility.**

University of Illinois
at Springfield

What is Database to Choose to Use

- **Key-value databases are well-suited for things like:**
  - **Tracking transient attributes in a web application, such as a shopping cart**
  - **Storing configuration and user data information for mobile applications**
  - **Storing large objects, such as images and audio files**

University of Illinois
at Springfield

What is Database to Choose to Use

- **Document databases are designed for flexibility.**
  - **If an application requires the capability to store varying attributes along with large amounts of data, then document databases are a good option.**

- **Document databases provide for embedded documents, which are useful for denormalizing.**

- **Document databases improve on the query capabilities of key-value databases with indexing and the capability to filter documents based on attributes in the document.**

- **Document databases are probably the most popular of the NoSQL databases because of their flexibility, performance, and ease of use.**

What is Database to Choose to Use

Documented databases are well suited to a number of use cases, including:

- Back-end support for websites with high volumes of reads and writes

- Managing data types with variable attributes, such as products

- Tracking variable types of metadata

- Applications that use JSON data structures

- Applications benefitting from denormalization by embedding structures within structures

University of Illinois
at Springfield

What is Database to Choose to Use

Column family databases are designed for large volumes of data, read and write performance, and high availability.

These database management systems run on clusters of multiple servers.

If your data is small enough to run with a single server, then a column family database is probably more than you need; consider a document or key-value database instead.

University of Illinois
at Springfield

What is Database to Choose to Use

Column family database are good options for:

- Applications that are geographically distributed over multiple data centers

- Applications that can tolerate some short-term inconsistency in replicas

- Applications with dynamic fields

- Applications with the potential for truly large volumes of data, such as hundreds of terabytes

University of Illinois
at Springfield

What is Database to Choose to Use

Problem domains that lend themselves to representations as networks of connected entities are well suited for graph databases.

One way to assess the usefulness of a graph database is to determine whether instances of entities have relations to other instances of entities.

What is Database to Choose to Use

Domains well suited for using graph databases include

- Network and IT infrastructure management
- Identity and access management
- Business process management
- Recommending products and services
- Social networking

What is Database to Choose to Use

- **NoSQL and relational databases are complementary.**
  - Relational databases offer many features that protect the integrity of data and reduce the risk of data anomalies.
  - Relational databases incur operational overhead providing these features.
  - In some use cases, performance is more important than ensuring immediate consistency or supporting ACID transactions. In these cases, NoSQL databases might be the better solution.
  - Choosing a database is a process of choosing the right tool for the job.

University of Illinois
at Springfield