NoSQL Databases

From: *Principles of Database Management*
W. Lemahieu, S. Vanden Broucket, B. Baeses
Cambridge University Press, 2018

Chapter 11

Introduction

- The NoSQL movement
- Key–value stores
- Tuple and document stores
- Column-oriented databases
- Graph-based databases
- Other NoSQL categories

The NoSQL Movement

- RDBMSs put a lot of emphasis on keeping data consistent.
  – Entire database is consistent at all times (ACID)
- Focus on consistency may hamper flexibility and scalability
- As the data volumes or number of parallel transactions increase, capacity can be increased by
  – Vertical scaling: extending storage capacity and/or CPU power of the database server
  – Horizontal scaling: multiple DBMS servers being arranged in a cluster

The NoSQL Movement

- RDBMSs are not good at extensive horizontal scaling
  – Coordination overhead because of focus on consistency
  – Rigid database schemas
- Other types of DBMSs needed for situations with massive volumes, flexible data structures, and where scalability and availability are more important ➔ NoSQL databases

The NoSQL Movement

- NoSQL databases
  - Describes databases that store and manipulate data in formats other than tabular relations, i.e., non-relational databases (NoREL)
- NoSQL databases aim at near-linear horizontal scalability by distributing data over a cluster of database nodes for the sake of performance as well as availability
- Eventual consistency: the data (and its replicas) will become consistent at some point in time after each transaction (but *continuous consistency* not guaranteed)

The NoSQL Movement

	Relational Databases	NoSQL Databases
Data paradigm	Relational tables	Key–value (tuple) based Document based Column based Graph based XML, object based Others: time series, probabilistic, etc.
Distribution	Single-node and distributed	Mainly distributed
Scalability	Vertical scaling, harder to scale horizontally	Easy to scale horizontally, easy data replication
Openness	Closed and open source	Mainly open source
Schema role	Schema-driven	Mainly schema-free or flexible schema
Query language	SQL as query language	No or simple querying facilities, or special-purpose languages
Transaction mechanism	ACID: Atomicity, Consistency, Isolation, Durability	BASE: Basically Available, Soft state, Eventual consistency
Feature set	Many features (triggers, views, stored procedures, etc.)	Simple API
Data volume	Capable of handling normal-sized datasets	Capable of handling huge amounts of data and/or very high frequencies of read/write requests

Key–Value Stores

- Key–value-based database stores data as (key, value) pairs
  - Keys are unique
  - Hash map, or hash table or dictionary

Key–Value Stores

```
import java.util.HashMap;
import java.util.Map;
public class KeyValueStoreExample {
	public static void main(String... args) {
		// Keep track of age based on name
		Map<String, Integer> age_by_name = new HashMap<>();

		// Store some entries
		age_by_name.put("wilfried", 34);
		age_by_name.put("seppe", 30);
		age_by_name.put("bart", 46);
		age_by_name.put("jeanne", 19);

		// Get an entry
		int age_of_wilfried = age_by_name.get("wilfried");
		System.out.println("Wilfried's age: " + age_of_wilfried);

		// Keys are unique
		age_by_name.put("seppe", 50); // Overrides previous entry
	}
}
```

Key–Value Stores

- Keys (e.g., "bart", "seppe") are hashed by means of a so-called **hash function**
  - A hash function takes an arbitrary value of arbitrary size and maps it to a key with a fixed size, which is called the hash value
  - Each hash can be mapped to a space in computer memory

Key
wilfried
seppe
bart
jeanne

Hash function →

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

Key–Value Stores

- NoSQL databases are built with horizontal scalability support in mind
- Distribute hash table over different locations
- Assume we need to spread our hashes over three servers
  - Hash every key ("wilfried", "seppe") to a server identifier
  - index(hash) = mod(hash, nrServers) + 1

Key–Value Stores

Key
wilfried
seppe
bart
jeanne

Hash function →

Hash	Index (= mod 3 + 1)
01	2
03	1
07	2
08	3

Server #1	Entries
03	(seppe,30)

Server #2	Entries
01	(wilfried,34)
07	(bart,46)

Server #3	Entries
08	(jeanne,19)

Sharding!

Key–Value Stores

- Example: Memcached
  - Implements a distributed memory-driven hash table (i.e., a key–value store), which is put in front of a traditional database to speed up queries by caching recently accessed objects in RAM
  - Caching solution

Key–Value Stores

```
import java.util.ArrayList;
import java.util.List;
import net.spy.memcached.AddrUtil;
import net.spy.memcached.MemcachedClient;

public class MemCachedExample {
 public static void main(String[] args) throws Exception {
 List<String> serverList = new ArrayList<String>() {
 {
 this.add("memcachedserver1.servers:11211");
 this.add("memcachedserver2.servers:11211");
 this.add("memcachedserver3.servers:11211");
 }
 };
```


Key–Value Stores

```
MemcachedClient memcachedClient = new MemcachedClient(
AddrUtil.getAddresses(serverList));

// ADD adds an entry and does nothing if the key already exists
// Think of it as an INSERT
// The second parameter (0) indicates the expiration - 0 means no expiry
memcachedClient.add("marc", 0, 34);
memcachedClient.add("seppe", 0, 32);
memcachedClient.add("bart", 0, 66);
memcachedClient.add("jeanne", 0, 19);


// SET sets an entry regardless of whether it exists
// Think of it as an UPDATE-OR-INSERT
memcachedClient.add("marc", 0, 1111); // <- ADD will have no effect
memcachedClient.set("jeanne", 0, 12); // <- But SET will
```


Key–Value Stores

```
// REPLACE replaces an entry and does nothing if the key does not exist
// Think of it as an UPDATE
memcachedClient.replace("not_existing_name", 0, 12); // <- Will have no effect
memcachedClient.replace("jeanne", 0, 10);

// DELETE deletes an entry, similar to an SQL DELETE statement
memcachedClient.delete("seppe");

// GET retrieves an entry
Integer age_of_marc = (Integer) memcachedClient.get("marc");
Integer age_of_short_lived = (Integer) memcachedClient.get("short_lived_name");
Integer age_of_not_existing = (Integer) memcachedClient.get("not_existing_name");
Integer age_of_seppe = (Integer) memcachedClient.get("seppe");
System.out.println("Age of Marc: " + age_of_marc);
System.out.println("Age of Seppe (deleted): " + age_of_seppe);
System.out.println("Age of not existing name: " + age_of_not_existing);
System.out.println("Age of short lived name (expired): " + age_of_short_lived);

memcachedClient.shutdown();

 }
}
```


Key–Value Stores

- Request coordination
- Consistent hashing
- Replication and redundancy
- Eventual consistency
- Stabilization
- Integrity constraints and querying

Eventual Consistency

- Membership protocol does not guarantee that every node is aware of every other node *at all times*
  - it will reach a consistent state over time
- State of the network might not be perfectly consistent at any moment in time, though will become eventually consistent at a future point in time
- Many NoSQL databases guarantee so-called **eventual consistency**

17

Eventual Consistency

- Most NoSQL databases follow the **BASE** principle
  - Basically Available, Soft state, Eventual consistency
- **CAP theorem** states that a distributed computer system cannot guarantee the following three properties at the same time:
  - Consistency (all nodes see the same data at the same time)
  - Availability (guarantees that every request receives a response indicating a success or failure result)
  - Partition tolerance (the system continues to work even if nodes go down or are added)

18

Eventual Consistency

- Most NoSQL databases sacrifice the consistency part of CAP in their setup, instead striving for eventual consistency
- The full BASE acronym stands for:
  - Basically Available: NoSQL databases adhere to the availability guarantee of the CAP theorem
  - Soft state: the system can change over time, even without receiving input
  - Eventual consistency: the system will become consistent over time

19

Stabilization

- The operation which repartitions hashes over nodes in case nodes are added or removed is called **stabilization**
- If a consistent hashing scheme is being applied, the number of fluctuations in the hash–node mappings will be minimized.

20

Integrity Constraints and Querying

- Key–value stores represent a very diverse gamut of systems
- Full-blown DBMSs versus caches
- Only limited query facilities are offered
  - e.g. put and set
- No means to enforce structural constraints
  - DBMS remains agnostic to the internal structure
- No relationships, referential integrity constraints, or database schema can be defined

Tuple and Document Stores

- A **tuple store** is similar to a key–value store, with the difference that it does not store pairwise combinations of a key and a value, but instead stores a unique key together with a vector of data
- Example:
  - `marc -> ("Marc", "McLast Name", 25, "Germany")`
- No requirement to have the same length or semantic ordering (schema-less!)

Tuple and Document Stores

- Various NoSQL implementations do, however, permit organizing entries in semantical groups (aka *collections* or tables)
- Examples:
  - `Person:marc -> ("Marc", "McLast Name", 25, "Germany")`
  - `Person:harry -> ("Harry", "Smith", 29, "Belgium")`

Tuple and Document Stores

- **Document stores** store a *collection of attributes* that are labeled and unordered, representing items that are semi-structured
- Example:

```
{
 Title = "Harry Potter"
 ISBN = "111-1111111111"
 Authors = [ "J.K. Rowling" ]
 Price = 32
 Dimensions = "8.5 x 11.0 x 0.5"
 PageCount = 234
 Genre = "Fantasy"
}
```

Tuple and Document Stores

- Most modern NoSQL databases choose to represent documents using JSON

```
{
    "title": "Harry Potter",
    "authors": ["J.K. Rowling", "R.J. Kowling"],
    "price": 32.00,
    "genres": ["fantasy"],
    "dimensions": {
            "width": 8.5,
            "height": 11.0,
            "depth": 0.5
    },
    "pages": 234,
    "in_publication": true,
    "subtitle": null
}
```


Tuple and Document Stores

- Items with keys
- Filters and queries
- Complex queries and aggregation with MapReduce
- SQL after all …

Items with Keys

- Most NoSQL document stores will allow you to store items in tables (collections) in a schema-less manner, but will enforce that a primary key be specified
  - e.g. Amazon's DynamoDB, MongoDB (_id)
- A primary key will be used as a partitioning key to create a hash and determine where the data will be stored

Filters and Queries

```
import org.bson.Document;
import com.mongodb.MongoClient;
import com.mongodb.client.FindIterable;
import com.mongodb.client.MongoDatabase;
import java.util.ArrayList;
import static com.mongodb.client.model.Filters.*;
import static java.util.Arrays.asList;

public class MongoDBExample {
 public static void main(String... args) {
 MongoClient mongoClient = new MongoClient();
 MongoDatabase db = mongoClient.getDatabase("test");

 // Delete all books first
 db.getCollection("books").deleteMany(new Document());
// Add some books
db.getCollection("books").insertMany(new ArrayList<Document>() {{
add(getBookDocument("My First Book", "Wilfried", "Lemahieu", 12, new String[]{"drama"}));
add(getBookDocument("My Second Book", "Seppe", “vanden Broucke", 437, new String[]{"fantasy", "thriller"}));
add(getBookDocument("My Third Book", "Seppe", “vanden Broucke", 200, new String[]{"educational"}));
add(getBookDocument("Java Programming", "Bart", "Baesens", 100, new String[]{"educational"}));
}});
```

Filters and Queries

```
// Perform query
FindIterable<Document> result = db.getCollection("books").find(
and( eq("author.last_name", "vanden Broucke"),
eq("genres", "thriller"),
gt("nrPages", 100)));

for (Document r : result) {
System.out.println(r.toString());
// Increase the number of pages:
db.getCollection("books").updateOne(
new Document("_id", r.get("_id")),
new Document("$set",
new Document("nrPages", r.getInteger("nrPages") + 100)));
}
mongoClient.close();}

        public static Document getBookDocument(String title,
         String authorFirst, String authorLast,
         int nrPages, String[] genres) {
         return new Document("author", new Document()
         .append("first_name", authorFirst)
         .append("last_name", authorLast))
         .append("title", title)
         .append("nrPages", nrPages)
         .append("genres", asList(genres));}}
```


Filters and Queries

Document{{_id=567ef62bc0c3081f4c04b16c, author=Document{{first_name=Seppe, last_name=vanden Broucke}}, title=My Second Book, nrPages=437, genres=[fantasy, thriller]}}

Filters and Queries

```
// Perform aggregation query
AggregateIterable<Document> result = db.getCollection("books")
          .aggregate(asList(
                    new Document("$group",
                              new Document("_id", "$author.last_name")
                                        .append("page_sum", new Document("$sum",
"$nrPages")))));

for (Document r : result) {
 System.out.println(r.toString());
}
```

Document{{_id=Lemahieu, page_sum=12}}
Document{{_id=Vanden Broucke, page_sum=637}}
Document{{_id=Baesens, page_sum=100}}

Filters and Queries

- Queries can still be slow because every filter (such as “author.last_name = Baesens”) entails a complete collection or table scan
- Most document stores can define a variety of indexes
  - unique and non-unique indexes
  - compound indexes
  - geospatial indexes
  - text-based indexes

SQL After All

- GROUP BY-style SQL queries are convertible to an equivalent map–reduce pipeline
- Many document store implementations express queries using an SQL interface
- Couchbase also allows defining foreign keys and performing join operations

```
SELECT books.title, books.genres,
authors.name
FROM books
JOIN authors ON KEYS books.authorId
```


SQL After All

- Many RDBMS vendors start implementing NoSQL by the following:
  - Focusing on horizontal scalability and distributed querying
  - Dropping schema requirements
  - Support for nested data types or allowing storing JSON directly in tables
  - Support for map–reduce operations
  - Support for special data types, such as geospatial data

Column-Oriented Databases

- A **column-oriented DBMS** is a database management system that stores data tables as sections of columns of data
- Useful if:
  - Aggregates are regularly computed over large numbers of similar data items
  - Data are sparse, i.e., columns with many null values
- Can also be an RDBMS, key–value, or document store

Column-Oriented Databases

- Example

Id	Genre	Title	Price	Audiobook price
1	fantasy	My first book	20	30
2	education	Beginners guide	10	null
3	education	SQL strikes back	40	null
4	fantasy	The rise of SQL	10	null

- Row-based databases are not efficient at performing operations that apply to the entire dataset
  - Need indexes which add overhead

Column-Oriented Databases

- In a column-oriented database, all values of a column are placed together on disk

Genre: fantasy:1,4 education:2,3
Title: My first book:1 Beginners guide:2 SQL strikes back:3 The rise of SQL:4
Price: 20:1 10:2,4 40:3
Audiobook price: 30:1

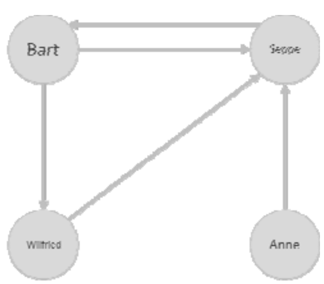
- A column matches the structure of a normal index in a row-based system
- Operations such as find all records with price equal to 10 can now be executed directly
- Null values do not take up storage space anymore

Column-Oriented Databases

- Disadvantages
  - Retrieving all attributes pertaining to a single entity becomes less efficient
  - Join operations will be slowed down
- Examples
  - Google BigTable, Cassandra, HBase, and Parquet

Graph-Based Databases

- **Graph databases** apply graph theory to the storage of information of records
- Graphs consist of **nodes** & **edges** (“follows” relation)



Graph-Based Databases

- One-to-one, one-to-many, and many-to-many structures can easily be modeled in a graph
- Consider the N–M relationship between books and authors
- RDBMS needs three tables: Book, Author and Books_Authors
- SQL query to return all book titles for books written by a particular author would look like this:

```
SELECT title
FROM books, authors, books_authors
WHERE author.id = books_authors.author_id
 AND books.id = books_authors.book_id
 AND author.name = "Bart Baesens"
```

Graph-Based Databases

- In a graph database (using **Cypher query language** from Neo4j)


```
MATCH (b:Book)<-[:WRITTEN_BY]-(a:Author)
WHERE a.name = "Bart Baesens"
RETURN b.title
```

41

Graph-Based Databases

- A graph database is a hyper-relational database, in which JOIN tables are replaced by more interesting and semantically meaningful relationships that can be navigated and/or queried using graph traversal based on graph pattern matching.

42

Graph-Based Databases

- Cypher Overview (Neo4j)
- Exploring a social graph

43

Graph Databases

- Location-based services
- Recommender systems
- Social media (e.g., Twitter and FlockDB)
- Knowledge-based systems

44

Other NoSQL Categories

- XML databases
- OO databases
- Database systems to deal with time series and streaming events
- Database systems to store and query geospatial data
- Database systems such as BayesDB which let users query the probable implication of their data

Evaluating NoSQL DBMSs

- Most NoSQL implementations have yet to prove their true worth in the field
- Some queries or aggregations are particularly difficult; map–reduce interfaces are harder to learn and use
- Some early adopters of NoSQL were confronted with some sour lessons
  - e.g., Twitter and HealthCare.gov

Evaluating NoSQL DBMSs

- NoSQL vendors start focusing again on robustness and durability, whereas RDBMS vendors start implementing features to build schema-free, scalable data stores
- NewSQL: blend the scalable performance and flexibility of NoSQL systems with the robustness guarantees of a traditional RDBMS

Evaluating NoSQL DBMSs

	RDBMSs	NoSQL Databases	NewSQL
Relational	Yes	No	Yes
SQL	Yes	No	Yes
Column stores	No	Yes	Yes
Scalability	Limited	Yes	Yes
Eventually consistent	Yes	Yes	Yes
BASE	No	Yes	No
Big volumes of data	No	Yes	Yes
Schema-less	No	Yes	No

Conclusion

- The NoSQL movement
- Key–value stores
- Tuple and document stores
- Column-oriented databases
- Graph-based databases
- Other NoSQL categories