

Topic: NoSQL Databases (Excluding MongoDB)

Executive Summary

We examined the dynamic world of NoSQL databases in this lecture, with an emphasis on how they manage unstructured data and scale horizontally. The main conclusions are as follows:

1. **Overview of NoSQL Databases:**

- o NoSQL databases, such as Cassandra and DynamoDB, provide more flexibility than conventional relational databases.
- o They work best in situations where data needs to scale seamlessly and comes in a variety of forms.

2. **NoSQL Trends:**

- o Graph databases: Neo4j and related programmes are excellent at simulating intricate relationships. Contemplate recommendation engines and social networks.
- o Time-Series Databases: For real-time analytics and the Internet of Things, InfluxDB effectively handles time-stamped data.
- o Multi-Model Databases: These databases offer versatility by combining features from several NoSQL models.

Organisations can address the issues brought on by ever expanding data quantities by carefully implementing NoSQL databases. NoSQL gives us the tools to traverse the data-driven future, whether it be for collecting sensor data, optimising recommendation engines, or guaranteeing real-time insights. Recall that the NoSQL world is always changing, thus being up to date is crucial. Continue your research, testing, and trend-following.

Main Idea

Introduction

NoSQL databases capacity to handle a wide range of data types and their scalability, flexibility, and adaptability have completely changed the data management industry. This section will examine the fundamental concept of NoSQL databases and go over its practical uses.

Example Scenario: E-Commerce Recommendation Engine

Think of an online store that seeks to offer its customers customised product recommendations. Relational databases have always had difficulty keeping up with the dynamic nature of real-time updates, product interactions, and user preferences. This is the sweet spot for NoSQL databases.

The Algorithm: Building a Recommendation Engine

1. Data Modeling:

- o Because graph databases, like Neo4j, can depict intricate relationships, that is why we have chosen them.
- o Nodes stand for people, things, and interactions (likes, purchases, and views).
- o Edges identify different forms of interactions and link users to products.

2. Graph Traversal:

- o We search the graph for a user's associated products after they log in.
- o We take into account both direct interactions, like purchases, and indirect relationships, like preferences of similar users.

3. Recommendation Generation:

- o We find similar people based on the user's graph neighbourhood.
- o We suggest products that the present user hasn't used but these comparable users have.
- o Recommendations (e.g., through content-based approaches or collaborative filtering) are rated according to their relevance.

4. Real-Time Updates:

- o We update the graph as users engage with the products (liking a new item, for example).
- o Suggestions are dynamically modified in response to the changing graph structure.

When to use it

1. Unstructured Data:

- o Situation: When working with data that doesn't neatly fit into columns and rows.

- o A few examples are records, multimedia information, sensor readings, and posts on social media.

- o Use of NoSQL: NoSQL databases support a variety of data formats by offering flexible schema design.

2. High Write Throughput:

- o Situation: Applications that require fast data input and updates.

Examples include real-time analytics, event-driven systems, and IoT platforms.

- o Use of NoSQL: NoSQL databases efficiently handle high write demands due to its distributed architecture.

3. Scalability Demands:

- o Scenario: Scaling horizontally becomes crucial as data volumes increase exponentially.

- o As an illustration, consider content delivery networks, recommendation engines, and e-commerce platforms.

- o Application of NoSQL: NoSQL databases ensure smooth scalability by distributing data among nodes.

4. Graph-Based Relationships:

- o Scene: Representing intricate connections between different elements.

- o As an illustration, consider social media, fraud detection, and recommendation engines.

- o NoSQL usage: Neo4j and other graph databases are excellent in expressing and querying interconnected data.

5. Time-Series Data:

- o Situation: Managing timestamp-based data points.

- o Server logs, financial market tickers, and sensor data are a few examples.

- o Application of NoSQL: Time-series databases, such as InfluxDB, are optimised for effective time-stamped data storage and retrieval.

6. Polyglot Persistence:

- o Situation: When not every use case can be satisfied by a single data model.

- o An illustration would be to combine key-value access, graph queries, and document storage.

- o NoSQL usage: Multi-model databases provide flexibility without compromising efficiency.

In conclusion, NoSQL databases perform exceptionally well in situations where conventional relational databases struggle, such as managing complex relationships, supporting a variety of data kinds, or scaling horizontally. Select the best NoSQL solution for your particular domain and data types.

References

NoSQL Referential Data:

- Stack Overflow. ["NoSQL Referential Data."](#)

Name: Kartik Vedi

Student ID:300374518

CSIS 4260 Section: 001

Seminar: 5

Topic: Data Velocity in the Facebook Industry

Executive Summary

Take a voyage into the centre of Facebook's virtual world, where an amazing data story is created by the harmonious combination of user interactions, posts, and videos. The Facebook sector has seen a phenomenal increase in data velocity in recent years, which has turned the platform into a living tapestry of human experiences. Similar to digital storytellers, users add to this constantly growing collection, creating a rich mosaic of exchanges, pictures, and feelings.

Facebook's environment has changed dramatically over the last ten years, moving from a text-focused platform to one that is rich in multimedia. This change, together with Facebook's community's worldwide expansion, has resulted to an unprecedented amount of broad and diversified data being generated. Cutting-edge algorithms appear as the guiding hands as this digital tapestry unfolds, sifting through the complex data strands to extract significant insights.

This session explores the changes that have occurred over the last five to ten years, focusing on the pulse of this dynamic ecosystem. It reveals the technological complexities involved in storing data, the techniques used to handle this digital harvest, and the analytical tools that give life to the enormous dataset, revealing the relationships and human tales that are concealed therein. Come along as we interpret the language of Facebook's data velocity, where each exchange adds to a shared story

and each data point becomes a brushstroke on the vibrant picture of our common humanity.

Essentially, the central thesis centres on the mutually beneficial association between people and algorithms, forming a digital symphony in which data is not merely an input but a dynamic factor that shapes the Facebook experience. In order to transform the data into meaningful connections, narratives, and interactions inside the virtual fabric of Facebook's worldwide community, the algorithms are essential to humanising the data.

Main Idea

An invisible tapestry of data is woven every second in the vast digital universe of Facebook, where billions of individuals share their thoughts, feelings, and lives. Envision a vibrant virtual metropolis where people converse, exchange images and videos, respond to posts, and express their feelings. This vibrant conversation results in an incredible deluge of data that is essential to the platform's survival.

As we explore the core of this digital city, we see that data output has grown exponentially over the past ten years. Previously a language-focused environment, it has evolved into a multimedia-rich one where users can easily tell tales by combining text, images, and videos.

The amount and variety of data created has increased at an unprecedented rate as a result of this evolution and Facebook's community's global expansion.

Advanced algorithms operate as digital stewards, carefully sorting through the torrent of data to help traverse this immense sea of information. Imagine these algorithms as knowledgeable librarians who are able to classify messages, identify trends, and make sense of the complex web of user interactions. They put in a lot of effort to keep everything in order and make sure that every bit of information adds something significant to the overall story of the Facebook experience.

Content suggestion is one area where this algorithmic skill is demonstrated. Facebook's algorithms begin to work when a user interacts with a post or video; they do this by examining the user's past interactions, preferences, and comparable user behaviours. The algorithm takes on the role of an artistic curator in this personalised trip, offering a well curated selection of content that fits the user's interests and digital footprint. This complex dance between artificial intelligence

and user engagement demonstrates how data is not only gathered but also turned into a personalised, interesting story.

Essentially, the central thesis centres on the mutually beneficial association between people and algorithms, forming a digital symphony in which data is not merely an input but a dynamic factor that shapes the Facebook experience. In order to transform the data into meaningful connections, narratives, and interactions inside the virtual fabric of Facebook's worldwide community, the algorithms are essential to humanising the data.

When to use it

Imagine browsing through your Facebook stream and easily selecting images and videos that are relevant to your interests and are updated in real-time to suit your tastes. The clever use of Facebook's data velocity—especially when combined with real-time analytics and content delivery—is what makes this personalised journey magical.

In Facebook's busy virtual metropolis, where emotions are transmitted quickly and interactions happen in the blink of an eye, quick data processing becomes critical. Imagine a situation in which a user exhibits joy or excitement in response to a live event, reacting enthusiastically. A useful snapshot of the user's present mental state is created from the data produced by this instantaneous interaction.

Real-time analytics play a crucial role in this dynamic environment by promptly analysing these responses, comprehending changing user attitudes, and adjusting content distribution as needed. Imagine that this is a real-time conversation between friends, with responses that are influenced by what is being said. Facebook's real-time analytics function in a similar manner, interacting instantaneously with users to make sure the material they view conforms to their changing tastes and interests.

This becomes particularly important for preserving the core component of the Facebook experience—user engagement. The platform is all about creating deep connections, and real-time user behaviour analytics allow content to be sent that personally speaks to users. Facebook's data velocity enables the company to suggest material that encapsulates a user's digital experience at any given time, much like a buddy may suggest a book depending on your mood.

Furthermore, situations like live events, breaking news, or popular themes make this real-time adaption especially relevant. Imagine a

Facebook Live performance that is happening in real time. The data created by participants sharing their responses, experiences, and comments creates a dynamic canvas. In order to provide a shared and immersive experience, real-time analytics quickly assess this data and adjust the content distribution to fit the audience's collective pulse.

Facebook's user experience is essentially made more human by the deliberate application of data velocity in real-time analytics and tailored content delivery. It makes the platform fluid and dynamic, allowing user interactions to alter the story as it unfolds and encouraging participation and connection within the virtual social fabric.

References

<https://harshthakare70.medium.com/how-facebook-stores-manages-and-manipulate-thousands-of-terabytes-of-data-with-high-speed-and-7aa2a3bce61>