

The impact of NoSQL databases for data management and analysis in smart cities.

Prof. Anupa Rajesh Haridas

Assistant Professor,
School of Computer Studies,
Sri Balajai University,Pune,India

Abstract - India is a developing country. In 21st century, the majority of the population is shifting from rural to urban areas. This can lead to a lot of issues in urban areas. As per the data collected in the 2011 census, approximately 31% of India's population lives in cities and contributes 63% of the country's GDP. By 2013, nearly 40% of India's population will live in cities and contribute 75% of the country's GDP. Prime Minister Narendra Modi launched the smart cities mission in India on June 25, 2015. The project aims to improve the quality of life for people living in India and promote economic growth. [1]

In this research paper, I am attempting to find solutions using technologies. Technology is becoming an increasingly important part of cities' urban sustainability strategies, as new technologies offer solutions that benefit citizens. Cities are looking to integrate smart systems into their industrial, infrastructure, educational, and social activities. The concept of a Smart City is relatively new. The idea of a SC first appeared in 1993 when Singapore city presented itself as an “intelligent city”. [2]

The digitalization of the city arrived after some time, however, it is important to note that these two concepts are distinct from each other. Many governments are considering adopting the smart city solutions implementation in their cities – implementing big data applications that support smart city components to reach the required level of sustainability and improve the living standards.

Index Terms - Smart city, Big Data, Data lake, Big Data Analytics, NoSQL Database, IoT devices.

I. INTRODUCTION

Data is the life blood of a smart city. The main technical challenges to achieve smart cities include optimizing resource utilization, providing uninterrupted service at all times, reducing costs, and reducing resource consumption. Effective data management doesn't just mean collecting and storing data. It also means sharing and combining data so that it's accessible, analyzed and usable across agencies, organizations, and even society. To address these challenges requires new technologies for modelling and processing data. Traditional relational databases often struggle to handle the large volumes of diverse and unstructured data generated by smart city sensors and IoT devices. NoSQL databases have emerged as a promising solution due to their ability to handle scalability, flexibility, and performance requirements. This paper explores the impact of NoSQL databases on data management and analysis in smart cities, examining their advantages, challenges, and real-world applications.

II. Details

Concept of Smart City:

A Smart City is a city that is managed using intelligent technologies that improve the quality of services provided to citizens. Smart City's uses technology for access to public transport, traffic management, water and power management, law enforcement, schools and hospitals, e-governances etc. [3]As part of urban mobility enhancement, around 1,300 smart mobility projects have been completed, with 383 projects nearing completion. [4]Over 70% of projects under the Smart Cities Mission (SCM) are in alignment with the UN's Sustainable Development Goals (SDGs) on cities, clean water and sanitation, clean energy and economic growth. [5]



Fig . 1 Smart City

Implementation of Smart city:

Smart city projects generally use sensors and connected IoT devices to collect and analyze data. This data is used to optimize city operations. Smart City uses data lake technology to store data in enormous capacity and there are many types of databases. One of the main and key features of big data technologies is NoSQL database.



Fig.2: Implementation of Smart city

III.Theory

Need of NoSQL for smart city Big Data Analytics:

NoSQL databases (aka "not only SQL") are non-tabular databases and store data differently than relational tables.NoSQL databases come in a variety of types based on their data model. The main types are document, key-value, wide-column, and graph. They provide flexible schemas and scale easily with large amounts of data and high user loads.[6]

1)Data Model:

In traditional Database data is stored in Relation format or Table format i.e. Structured Format .

In No-SQL data is saved in Unstructured format such as document,key-value,wide-column and graph.

2)Scalability :

In Traditional RDBMS are designed for vertical scalability, where you scale up by adding more resources to a server is costly.

NoSQL databases are designed for horizontal scalability, where you scale out by adding more servers to a distributed cluster.

3)Flexibility:

Traditional RDBMS is not much flexible as it follows Structured Data.

NoSQL can handle a large amount of various types of unstructured data regardless of their data type.

4)No-SQL also insures ACID properties.

As data in smart cities are generated continuously and in large amount it is better way to use NoSQL.

Concept of Data Lake:

A data lake is a centralized repository that allows us to store all structured and unstructured data at any scale. It enables to store vast amounts of raw data in its native format until it is needed. Unlike traditional data warehouses, which require you to structure data before storing it, a data lake allows you to store data in its raw form and structure it later as needed for different types of analysis.

Layer Architecture of smart city:

1)Data Acquisitions Layer: A large number of sensors and connected devices used to provide a wide range of services.

2)Middleware: Information /data storage and processing of Big Data tools are available.

3)Applications: Implementation and usage of the applications inside the city by designated users. [7]

Concept of Big Data:

Big data is a large, diverse collection of information that is growing at a rapid rate. It includes the amount of information, the speed or scale at which it is generated and collected (known as the “three v’s” of big data), and the variety or scale of the data points that it covers. i.e. Variety, Velocity and Volume.[8]

The data types include structured, semi structured and unstructured. This data is processed for valuable insights and utilized in machine learning projects and predictive modeling and other cutting-edge analytics applications. Big Data enables organizations to solve problems in their business through Big Data Analytics. The amount of Big Data increases exponentially over time.

Big Data Analytics :

Big data analytics is the process of examining large amounts of heterogeneous data to uncover information — such as hidden patterns, correlations, market trends and customer preferences that can help organizations make informed business decisions, which improve business-related outcomes.

Big Data Flows in Smart City:

- 1.Data professionals collect data from (IOT devices) a variety of different sources. It is a mix of semi-structured and unstructured data.
- 2.Data is prepared and processed. After data is collected and stored in a data lake, data professionals must organize, configure and partition the data properly for analytical queries.
- 3.Data cleansing is done by using different tools to improve quality of data, removes error, unwanted /irrelevant data.
- 4.The collected, processed and cleaned data is analyzed with analytics software. [7]

IV. Case Study

Case Study of NoSQL used in Smart City Big Data Analysis:

This case study focuses on a hypothetical smart city deploying a comprehensive parking management system powered by NoSQL databases. The system comprises IoT-enabled parking sensors, mobile applications for users, and a centralized data processing platform. MongoDB, a popular NoSQL database, is chosen for its flexibility, scalability, and real-time data processing capabilities.



Fig.3: Parking Management System

- 1)IoT Parking Sensors: Deployed across parking spaces within the city, these sensors detect the presence or absence of vehicles in real-time. Each sensor transmits data regarding occupancy status, location, and timestamp to the central database.
- 2)Mobile Applications: Citizens and visitors utilize mobile applications to access real-time parking availability, reserve parking spots, and make payments. These applications interface with the NoSQL database to fetch and update parking information seamlessly.
- 3)NoSQL Database (MongoDB): MongoDB serves as the backbone of the parking management system, storing and managing the vast volume of parking-related data. Its document-oriented structure allows for flexible schema design, accommodating diverse data types such as sensor readings, user profiles, and transaction records.
- 4)Data Analytics Engine: With MongoDB's framework and analytics tools, the smart city extracts valuable information from parking data. This information helps decision-making processes, such as optimizing parking allocation, predicting demand trends.

V.Objectives of Research:

- 1) Introducing the smart city concept: Provide a need for technologies to transform a city into a smart city that can have a positive impact on urban life.
- 2)Examine the Role of Data in Smart Cities: Provide an overview of the importance of data in optimizing various aspects of urban living.
- 3)Identify Challenges in Data Management: Discuss the limitations of traditional relational databases in handling the large volumes of diverse and unstructured data generated by smart city.
- 4)Introducing NoSQL Databases: Defining NoSQL databases and highlight their characteristics, including scalability, flexibility, and support for diverse data types, which make them suitable for addressing the challenges of data management in smart cities.
- 5)Present Use Cases of NoSQL in Smart Cities: Provide real-world examples and case studies demonstrating the application of NoSQL databases in areas such as real-time parking management system.
- 6)Highlight Future Trends: Explore emerging trends and opportunities for NoSQL databases in smart cities.
- 7)Conclusions: Summarize the key findings of the paper, provide recommendations for further research and exploration in this field.

VI.Methodology Used for research:

The research method uses a secondary source i.e. Review of Literature and articles, news etc. Conduct a comprehensive review of existing literature on NoSQL databases, data management, and smart city initiatives. Identify key studies, frameworks, and methodologies relevant to the research topic.

VII.Conclusion:

NoSQL databases play a critical role in supporting real-time systems by providing scalable, flexible, and high-performance data storage and processing capabilities so it is better to use NoSQL for smart city Big Data Analysis.

VIII.References:

- 1)<https://www.magicbricks.com/blog/smart-cities-mission/127407.html>
- 2)T. M. Heng and L. Low, "The intelligent city: Singapore achieving the next lap: Practitioners forum," *Technology Analysis & Strategic Management*, vol. 5, no. 2, pp. 187–202, 1993
- 3)<https://www.techtarget.com/iotagenda/definition/smart-city>
- 4)<https://www.financialexpress.com/business/infrastructure-smart-cities-mission-modi-govt-reveals-big-update-on-projects-6650-completed-so-far-in-100-cities-3368921/>
- 5)<https://indianexpress.com/article/india/sustainable-development-at-heart-of-smart-cities-has-lessons-for-global-south-report-8959228/>
- 6), Pramod J. Sadalage and Martin Fowler" *NoSQL Distilled: A Brief Guide to the Emerging World of Polyglot Persistence Paperback*".
- 7)Smart City – Big Data Analytics - KR Architecture World (industry40.co.in)
- 8) V. Mayer-Schönberger and K. Cukier, "Big Data: A Revolution That Will Transform How We Live, Work, and Think," Eamon Dolan/Houghton Mifflin Harcourt, 2013.