

# Graph Database-An Overview of its Applications and its Types

Neha Tyagi<sup>1</sup>, Neelam Singh<sup>2</sup>

(M.Tech Scholar Dept CS & E, GEU, Dehradun)

(Assistant Professor, Dept of Computer Science & ENG, GEU, Dehradun )

## Abstract:

Relational database has been around for many decades and is one of the pioneering database technology for efficient storage and retrieval of structured data. There are number of technologies to handle and process complex structured and unstructured data in an efficient manner with Graph Database as one of the promising technology. With increasing data volume, net clients and huge growth in applications, there is a need to switch from traditional database approach to approach like graph database. This paper provides the general overview of graph database their uses and its application area in which it works very well.

*Keywords*— Graph database, NOSQL, Relational database, Graph storage.

## I. INTRODUCTION

From the last decades there has been a huge growth in volume of data .The data is increasing day by day .The relational database was used from late 1960s for data storage and retrieval of data and its applications. Retrieving of data has been done by using the SQL query language. The relational database was efficient to store the data in forms of tables. But there was some limitations of relational database that was its limited ability to store the data. Relational database was not able to store the connected data, big data problems including complex interconnected data [1]. Storing retrieving and manipulating such information becomes very difficult by using RDBMS approaches. So there arrived a data structure called graph database. It could be presented in network model database from late 1960's.CODASYL, which had characterized COBOL in 1959, which defined the Network database language in 1969.Variou changes to graph database showed up in mid 1990's. In the mid-late 2000 commercial ACID graph database for example Neo4j and Oracle Spatial and graph become available. This paper discuss about the graph database and the properties of graph database, applications of graph database and some of the leading graph database.

## II. GRAPH DATA BASE

Graph Database is currently the most powerful technology for storing the data. It is NOSQL database that uses graph theory to store, map and query relationships. The four building block of graph database are Nodes, Relationship, Properties and the labels .Nodes represents entities, relationship that connects one nodes to other node through a line and properties are information about the data and every

node can have labels attached, used to represent the roles. Fig 1 shows the basic structure of graph database representing its nodes, properties and relationship. It is emerged a powerful technology for demonstrating information when an attention is to focus on relationship between the entities in the data design model [4]. Used for handling very huge, thick and interrelated data sets. Its design allow development of predictive model and discovery of relationship. The highly interconnected nodes with the relationship allows very rapid traversal, the benefit of using this is that the traversals are very localized no need to consider other in consequential information. So, Graph database provide a better alternative to Relational database. With the help of graph database we can do many things easily with it we can represent many things in a quite simple manner. Science, chemistry, social networking, social web applications can be presented in very normal shape. There has been expanded growth in graph to represent the social network data. There is huge growth in data and that was not possible to store and query with relational database whereas graph database is very useful in understanding complex and dynamic connection in very connected data to create very worthwhile results [6]. Generally used with transactions systems, in this the relation can be pre characterized or made by the clients and the technology is to query highly interconnected data.

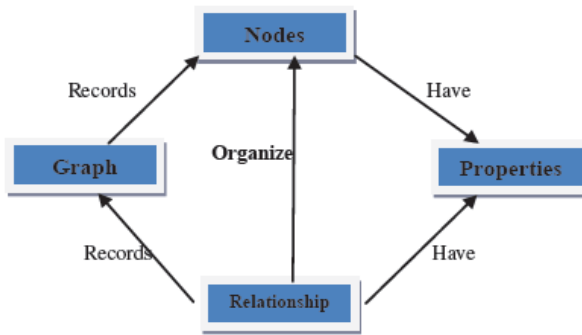


Fig 1 Representing Graph Database

A. GRAPH DATABASE PROPERTIES

- 1) *Graph Storage:* Graph database have a concept of storage which deals with persistence and queries over connected data. It uses native graph storage that means that each node has to handle all the outgoing relationship it has and each relationship it has been associated with. Native graph uses Index Free Adjacency that means that every node is directly connected to its incoming and outgoing relationship [1]. Not all the graph databases uses the native storage but some uses the graph data into the relational system, it can be object oriented or any other useful information system. The benefit of using native storage is that it is basically build for better performance and its scalability , provide very rapid traversals whereas non -native depends on the non-graph backend (for ex SQL).Native storage is greatly useful in great traversal.
- 2) *Graph processing:* Graph processing system are designed to tackle an alternate kind of issue. They provide multidimensional analytical queries concentrating on high performance and to manage transactional oriented queries. In graph processing exists a property called index free adjacency, which means that node is directly connected to its incoming and outgoing relationships. Where it acts like an index to other nodes which is much cheaper than having global indexes, if having this local index the query time has been decreased. Index free adjacency is key for high performance, traversal, queries. Few definitions require that they use index free adjacency meaning that the associated nodes physically points to each other in database. Database that behaves from the user view point behaves like a graph database. Native graph processing uses to describe graph data with index free adjacency. They are designed to solve different types of issues.

III. APPLICATIONS OF GRAPH DATABASE

Graph database shine when working in areas about data interconnectivity .The graph database is very useful. In such

the relations between the data and the data that are on the same level. For this kind of data the relational database does not work, for this kind of data the graph database structure works well [7].Relational database can also be used but in a limited manner and in an expensive way.

A. SOCIAL GRAPHS

Social networking is the online social network service. Clients tend to share their exercises (Facebook), substance (Twitter), photographs, recordings, or proficient data with companions on their sites. The graph representation of the system usually treats the clients as the nodes and the connections (friends, follow, likes). It has become very popular now a days. As we are living in a modern society where social network has taken a major role in humans life. It has not taken role in public area but also in graph research. It present a significant measure of information, as well as present vast graph data issues for the society. The graphs, nodes in that stores nodes of individual as well as link nodes of media, relationships and messaging [1]. Considering business utilization of SNS. Companies promote marketing information from the data, content, activities from clients or the advertisement that clients like o their page. To make balance relationship between the user experience and the business reason, decisively focusing on the promotion to every client.it is the primary task in the action of social network service [6]. For this users using machine learning techniques such as collaboration filtering to search the people (nodes) and the entities to provide the results. Now a days individuals begin to utilize the connection(edges) between individuals with people entity information to get more accurate results on the suggestions .This process depends on mining the graph data itself which requires high calculation and storage capacities .The graph database framework prove very great ability for academic and industry clients. Fig 2 representing example of social networking that showing nodes relationship and other relationship associated with it.

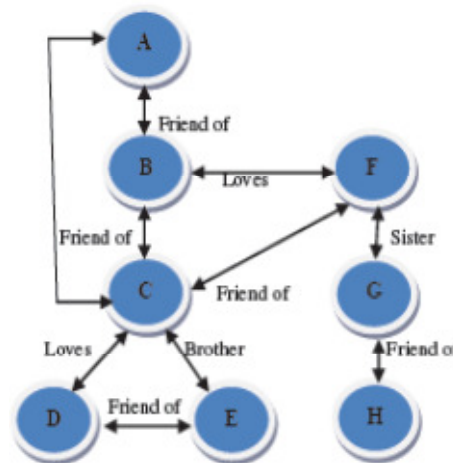


Fig 2 Example of Social Network

### B. WEB GRAPH

Web is basically a graph data and information linked together .It can be defined as linked data which supports the fast scattering of vast scale structured data through standards. URI (Unique Resource Identifiers) that build the formation of different types of data anywhere on the web .Structured Information gives standards for the data to be connected. Links of similar resources connect the data to form groups or cluster of data. This graph database presents application in c search and data collection [6].

### C. CHEMICAL BIOLOGICAL NETWORKS

The graph structure is broadly used in the field of science like chemical and biology. In science or bioinformatics region graph theory can be seen as a very powerful tool to offer people to discover new features .There is branch in chemical called chemical graph theory. Chemical theory is modeled as graphs by assigning atoms as nodes and bonds the edges between them. Biological data can also be represented in a similar manner [6].The graph data is very important for many important operations on drug discovery and analysis on many fields. Many of graph operations are focused on pattern recognition. It is done by finding frequent sub graphs of given graph. If these is done with the help of traditional database. Then operational would have taken very significantly more time because of the recursive nature of traversing a graph. It uses graph databases to provide complex web information that includes genes, enzymes and protein [1].

### D. MASTER DATA MANAGEMENT

Basic to the operation of the business. It incorporates data including clients, customers, products, providers, departments, geologies and cost centers. In mostly very large organizations. The data is handled at very large and different places having lot of overlap and redundancy in very different formats with different quality and it means access them. It is the act of recognizing, cleaning and gathering the data. Its key concern is overseeing change after sometime, organization union, business rule change, joining new wellspring of information. Graph database does not provide full master data management. They are provide to demonstrating storing, queries of hierarchies, master metadata and master data models. A graph data database organized yet schema free data modes for ad hoc, variable and extraordinary structure schema anomalies that emerge when there are numerous repetitive data source allowing at the same time for fast development for master data model in accordance with changing business model.

### E. RECOMMENDER SYSTEM

It provides user an information about the product by predicting interest on different sorts of data. With the help of recommender system we get to know the area of

interest of the client or user is interested on. The two techniques apply to the graph database Item to Item correlation prescribes items in view of relationship between new things and the item in which the user have already shown the interest [1].Apart from that User to User system gives forecasting based on correlations made by looking how the clients communicates with the system. Combine both these frameworks offer very intense approach to use or gather lots of information and also helps us to find the user behavior and the relationship of user with that product whether it may be the positive review or it may be the negative ones. Many organizations find this is the best way to promote their clients and the way to appreciate [1]. In below fig 3 the person are rating to movie .From this the film production gets to know about the rating details of the movie.

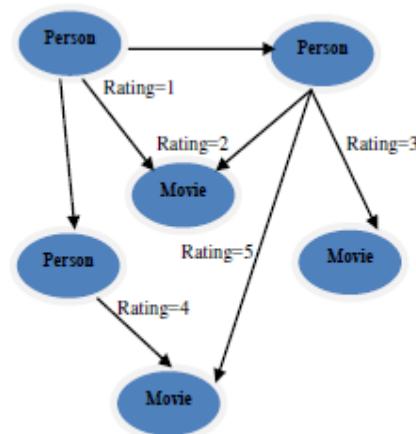


Fig 3 Example of recommender system

## IV. TYPES OF GRAPH DATA BASE

### A. NEO4J

One of the most popular and leading graph database. It is one of the open source options and it is the product of neo4j technologies. Most popular alternatives due to its dual free software/commercial model [7]. It is written in java and can be conveyed on numerous frameworks [2]. It Allows client to provide fraud detection system based on connected intelligence which provides upgraded levels of experiences contrasted with algorithms that uses statistical analysis and pattern recognition. It has been in production for more than five years .Neo4j is an embedded, disk based, fully transactional java persistence engine that

stores data structures in graph rather than the tables. It is more scalable, flexible, provides great performance [4].

**B. FLOCK DB**

It was created by twitter for relationship related analytics. There is no arrival of flock DB and there are some debate with respect to whether it can be genuinely alluded to as a graph database. Without traversal it is only a persistent graph but not a graph database. It is fascinating because it is easy to understand that a vertex in this belong to person and the edge either defines the property [4]. Flock DB uses the MYSQL as basic database storage system. NOSQL databases sometimes use SQL framework as their essential database framework. In Flock DB application server each request is independent from all others requests [4]. The greatest contrast between Flock DB and other graph database is graph traversal. In twitter model has no need for traversing social graph instead it is only worried about the direct edges (connection) on given node [3]. It can distribute resources elsewhere. However because of it is mostly used largest sites in the world and because of this it is very simpler to use then other graph DB.

**C. ALLEGRO GRAPH**

It is graph database developed around W3C for resource description framework. It is made for dealing with linked data and the semantic web. It supports Prolog, SPARQL. It is a restrictive result of Franz Inc. which advertises various semantic web items including LISP based development tool. It is modern system that uses efficient memory utilization in combination with disk- based storage, enabling billions of quads while providing greatest performance. It is a system to load, store and query data. It has java, prolog and python interfaces. It basically provides full support to ACID properties, rollback, check pointing and the commit and also support full and fast recoverability. And great feature which it provides is automatic and dynamic indexing.

**D. INFINITE GRAPH**

It is restrictive graph database from objectivity, the organization behind the object database of seminal name. It will likely make graph database with virtual unlimited scalability. Infinite graph is now being utilized by the department of defense. Its objectivity is high performance, scalable graph database that allows storage of information. It is highly specialized graph database. The objective will keep on supporting to authorized clients and recommend to java engineer who wish to use graph analysis. Particular features like path finding have been converge into the database, better improvement, quality confirmation, support and maintenance of products, cost efficient [3]. It provides powerful navigation.

**E. ORIENT DB**

It was the multi model open source NOSQL DBMS that combines the power of graphs and flexibility of documents into the scalable, high performance operational database. It is multi model database that is written in java and supporting graph document, key/value and object models.

It also supports schema-less, schema-full and schema fixed models. Orient DB has a very great security system based on users and support querying and extended for graph traversal, uses several indexing mechanism based on B tree and extendible hashing and hash index. The development of Orient DB will be based on open source community. Some of the best features provided by orient DB are quick installation, pattern matching, full transactional supports ACID and as well as commercial support is also available. Also provides security features, it is distributed so it can run anywhere. It graph structured data model that supports web technologies easy to understand and use.

Table 1 Comparison of types of graph database.

Features Functionality	Neo 4j	Flock DB	Allegro DB	Infinite Graph	Orient DB
Source	Open Source	Open Source	Property	Property with open Source Extension	Open Source
Consistency	ACID Possible	ACID	ACID	Flexible(ACID to Relaxed)	ACID
Query Method	Cypher	SQL	SPARQL Prolog	Graph Navigation API procedure language Qualification	Extended SQL
Language	Java, Python, Java Script	Java, Ruby	Java, Ruby, Python, Perl	Java, {C, C++}	Java
Platform	Windows, Unix, Mac OS x	Various	Linux, Mac OS x	Windows, Unix, Mac OS x	Windows, Linux, + any Java Platform
Graph Model	Directed Graph	Adjacency list of directed graph	XML	Labeled Directed Graph	Directed Graph

## V. CONCLUSION

Graph Database have realized another method for demonstrating and navigating interconnected information or data which is unparalleled in information storage. With the upcoming of generation review frameworks for ex NEO4J utilizing graph database issues can be addressed without falling back on restricting usage on Relational database. The graph database has application for natural, semantic, web, organization and recommender framework that provide type of data model only which they can provide. The graph database provide the structural type queries then the relational database. In full context character searches, the graph databases proves to be more significant. The indexing mechanism provides by the graph database made the queries less efficient, it has brought another pattern of modeling data and solving complex issues. They perform superior to social database, numerous big organizations using graph database to handle their large amount of data and queries. In this paper the general overview of graph database is given in it. Its application in real cases and overview of different types of graph database that organizations are using to handle their data.

## REFERENCE

- [1] JUSTIN J. MILLER, "GRAPH DATABASE APPLICATIONS AND CONCEPTS WITH NEO4J" *PROCEEDING OF SOUTHERN ASSOCIATION FOR INFORMATION SYSTEM CONFERENCE, ATLANTA, GA, USA MARCH 23, 2013.*
- [2] CHAD CLICKNAIR, MICHAEL MACIAS, "COMPARISON OF GRAPH DATABASE AND A RELATIONAL DATABASE" *ACMSE APRIL 15-17, 2017.*
- [3] HARSHA R. VYAWAHARE, DR P.P. KARDE, "ON OVERVIEW OF GRAPH DATABASE MODEL" *INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH IN COMPUTER AND COMMUNICATION ENGINEERING VOL.3, ISSUE 8, AUGUST 2015.*
- [4] PALLAVI MADAN, ANUZ SAXENA, "REVIEW: GRAPH DATABASES" *INTERNATIONAL JOURNAL OF ADVANCED RESEARCH IN COMPUTER SCIENCE AND SOFTWARE ENGINEERING, VOL 4, ISSUE 5, MAY 2014.*
- [5] ADRIAN SILVESW, DOINA CARAGEA, "GRAPH DATABASE" *ARTIFICIAL INTELLIGENCE RESEARCH LABORATORY DEPARTMENT OF COMPUTER SCIENCE.*
- [6] SHEFALI PATIL, GAURAV VASWAMI, "GRAPH DATABASES- AN OVERVIEW" *INTERNATIONAL JOURNAL OF COMPUTER SCIENCE INFORMATION TECHNOLOGY, VOL 5 (1), 2014.*
- [7] PRADEEP JADHAV, RUHI OBEROI, "COMPARATIVE ANALYSIS OF GRAPH DATABASE MODELS USING CLASSIFICATION AND CLUSTERING BY USING WEKA TOOL" *INTERNATIONAL JOURNAL OF ADVANCED RESEARCH IN COMPUTER SCIENCE AND SOFTWARE ENGINEERING, VOL 5, ISSUE 2, FEBRUARY 2015.*
- [8] NISHTHA JATANA, SAHIL PURI, MEHAK AHUJA, ISHITA KATHURIA, DISHANT GOSAIN, "A SURVEY AND COMPARISON OF RELATIONAL AND NON-RELATIONAL DATABASE" *INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT), VOL 1 ISSUE 6, AUGUST - 2012 ISSN: 2278-0181.*

[9] MARY FEMY P.F, RESHMA K.R, SUREKHA MARIAM VARGHE, "OUTCOME ANALYSIS USING NEO4J GRAPH DATABASE" *INTERNATIONAL JOURNAL ON CYBERNETICS & INFORMATICS (IJCI) VOL. 5, No. 2, APRIL 2016.*

[10] ROBERT MCCOLL, DAVID EDIGER, "A PERFORMANCE EVALUATION OF OPEN SOURCE GRAPH DATABASES" *FEBRUARY 16, 2014, ORLANDO, FLORIDA, USA.*

[11] MIKE BUERLI, "THE CURRENT STATE OF GRAPH DATABASES" *DEPARTMENT OF COMPUTER SCIENCE CAL POLY SAN LUIS OBISPO [MBUERLI@CALPOLY.EDU](mailto:MBUERLI@CALPOLY.EDU) DECEMBER 2012.*