ISSN: 1813-4890

Research on the Development and Application of Database Technology in the Background of Big Data

Lili Shi[a], Zhengqiu Weng, Dafeng Gong*

College of Artificial Intelligence, Wenzhou Polytechnic, Wenzhou, Zhejiang 325035, China

[a]sll520qiu@qq.com, *dafenggong@qq.com

Abstract

The arrival of the age of big data enables everyone to enjoy the convenience of data for life.Under the background of big data era, the application of database technology has penetrated into all aspects of people's lives and played an extremely important role in today's information fields.This paper takes the big data era as the background, starts with the development process of database technology, expounds the characteristics of database technology in the big data era, analyzes the application status of database technology and common database technology, and looks forward to its future development trend.

Keywords

Big Data; Age of Big Data; Database Technology.

1. Introduction

With the continuous development and progress of modern information technology, database technology has played an extremely important role in today's information fields, and has also become an important way and means to obtain data information resources.In the age of big data, all walks of life and all fields can not do without the support of database technology in all kinds of data management and data sharing.The huge data volume, diversified data types and other changes make the traditional database technology upgrade under the promotion of the environment, and its ability to process data is more powerful.Moreover, with the continuous integration of big data with various industries and fields, database technology has been applied more widely and received more and more attention.This paper takes the big data era as the background, starts with the development process of database technology, expounds the characteristics of database technology in the big data era, analyzes the application status of database technology and common database technology, and looks forward to its future development trend.

2. Introduction to the Database Technology

Database technology is a core technology of information system and a method of computer aided data management. It studies how to organize and store data and how to obtain and process data efficiently. It is a technology to process, analyze and understand the data in the database by studying the basic theories and implementation methods of database structure, storage, design, management and application.

Database technology mainly refers to effective storage and management of information data, optimization of data structure, design of data management mode, etc. through relevant technical measures[1].Therefore, the application of database can greatly improve our ability to use data analysis to deal with problems in the age of big data.From here, we can see that database is a science to study data management, and database technology is a means to apply this science.As an important part of modern information science and technology, database technology is the core of modern computer data processing. Its fundamental goal is to solve the problem of data sharing. It provides a scientific solution for the processing of large amounts of data in computers, which can effectively reduce data redundancy and ensure data security in the process of data processing, It can also improve the efficiency of data retrieval and processing in the process of data application[2].

3. The Development Course of Database Technology

The development of database technology has become an important part of modern information technology and the foundation and core of modern computer information systems and computer application systems.Database technology was first produced in the mid-1960s. Today, in recent decades of history, its rapid development and wide range of use are far behind other technologies.Data model is the core and foundation of database system. Database technology can be divided into three development stages according to the progress of data model: the first generation of mesh and hierarchical database system; The second generation of relational database system; Third generation database system and emerging database system.

3.1. Primary Stage - First Generation Database System

Both hierarchical and mesh models are formatted models. They have common characteristics from architecture, database language to data storage management, and are the first generation database systems.

3.1.1. Representative of the First Generation Data System

In 1969, the hierarchical model database management system IMS was developed by IBM.

The Database Task Force (DBTG) under the CODASYL (American Society for Database System Languages) systematically studied and discussed database methods, and the DBTG report was put forward in the 1960s. The method proposed by DBTG is based on the mesh structure, which is a typical representative of the mesh model database.

3.1.2. Features of the First Generation Database System

①An architecture that supports three levels of patterns (external, internal, and internal). There is conversion (or mapping) function between modes.

②The storage path is used to represent the relationship between data. This is one of the main differences between a database system and a file system. The database not only stores data, but also stores the relationship between data. The relationship between data is represented and realized by access path in hierarchical and mesh database systems.

③Independent data definition language. Hierarchical data system and mesh database system have independent data definition language, which is used to describe the three-level schema of database and mutual image.

④Data manipulation language for navigation. The data query and data manipulation language of hierarchical and mesh databases is a navigational procedural language with records one at a time.

3.2. Intermediate Stage - Second Generation Database System

The relational database system supporting the relational data model is the second generation database system.

In 1970, researchers of IMB Company proposed the relational model of database, initiated the research on database relational methods and relational data theory, and laid a theoretical foundation for relational database technology.

ISSN: 1813-4890

The 1970s was the era of theoretical research and prototype development of relational databases.

①The theoretical foundation of the relationship model is established, and the specification of the relationship model that is accepted by people is given.

②Research on relational data language, including relational algebra, relational calculus, SQL and QBE, and establish SQL as the language standard of relational database.

③A large number of RDBMS prototypes have been developed, and a series of key technologies in system implementation, such as query optimization, transaction management, concurrency control, fault recovery, have been overcome. This not only greatly enriched the database management system implementation technology and database theory, but also promoted the industrialization of databases.

3.3. Advanced Stage - The Third Generation Database System

Although the data model of the second generation database system describes the structure of displaying world data and some important interrelationships, it still cannot capture and express the rich and important semantics of data objects.

The third generation database system is characterized by more diversified data models and data management functions to meet the requirements of a wide range of complex new applications. The third generation DBMS should have three basic features (three basic principles):

①The third generation database system shall support data management, object management and knowledge management.

In addition to providing traditional data management services, the third generation database system will support richer object structures and rules, and should integrate data management, object management and knowledge management.

②The third generation database must maintain or inherit the technology of the second generation database system.

The third generation database system should inherit the existing technology of the second generation database system; Maintain the non procedural data access mode and data independence of the second generation database system, which can not only support object management and rule management, but also better support the original data management and support the query required by most users.

③The third generation database must be open to other systems.

The open line of database system is to support database language standards; Support standard network protocols on the network; The system has good portability, connectivity, scalability and interoperability.

3.4. Future Development Trend - Emerging Database System

In the age of big data, the amount of data is growing explosively, and the data storage structure is becoming more and more flexible. The emerging business needs of increasingly changing have led to the existence of databases and application systems becoming more and more abundant. These changes have constantly challenged various capabilities of the database, promoted the evolution of database technology towards model expansion and architecture decoupling, and presented a development trend of learning from each other and constantly integrating with cloud computing, artificial intelligence, block chain, privacy computing, new hardware and other technologies. It can be summarized in three directions.

①The multi-mode database realizes one database for multiple purposes, uses a unified framework to support mixed load processing, and uses AI to achieve management autonomy, improving ease of use and reducing use costs.

ISSN: 1813-4890

②Make full use of emerging hardware and combine it with cloud infrastructure to enhance functions and performance.

③The privacy computing technology is used to help improve the security capability, and the block chain database is used to assist in data storage and traceability, so as to improve the credibility and security of data.

4. Technical Features of Database in the Background of Big Data

4.1. Uniformity

In the environment of big data, data processing needs to be uniform. Data processing efficiency will be reduced if there is no uniform data. Data processing efficiency is directly related to the efficiency of big data application, and uniform data generation needs to depend on database uniformity. There are many types of traditional databases, together with the increase of data types brought by the development of the times, it is difficult to achieve unified management of different types of data. Under the background of the diversity of traditional database system building modes, incompatible phenomena of different types of databases often occur. In the face of the development of the big data era, database technology is upgraded, different types of databases are unified, and the characteristics of database technology uniformity in the big data environment are generated.

The incompatibility and inconsistency between traditional databases lead to idle and wasted data resources. After the unified construction of databases, the efficiency of data processing has been further improved. Enterprises and public services continue to innovate on the basis of the unified construction of databases, data storage, management and other functions continue to optimize, and data utilization efficiency improves to promote the development of productivity and ultimately promote social development[3].

4.2. Sharing

The development of big data environment depends on the increasing efficiency of data information processing. Only the efficiency of data processing can meet the needs of big data processing and development and utilization. There is not enough data processing efficiency, and the big data environment lacks a foundation. The functions of data storage and management need to be realized by database technology. Database technology upgrade and optimization are the precondition for the formation of data environment. Sharability is the technical feature of database technology in the process of optimization and upgrade. Sharability of database technology promotes efficient data transmission and utilization. Data sharing makes it easier, faster and faster to obtain, transfer and use data resources. Increased utilization of data information promotes the generation of big data environment.

During the specific optimization and upgrade process of database technology, database sub-nodes are more scientific and reasonable, database technology produces more functions, and sharing is easier to achieve. The features of database technology integration and sharing promote the development of big data environment, improve the utilization efficiency of data information resources, and further enhance the value of database technology[4].

4.3. Abundance Resources

The most unique feature of the stored resources is that they are rich in information and comprehensive in content. For example, CNKI is an internationally leading online publishing platform integrating journals, doctoral theses, master's theses, conference papers, newspapers, reference books, yearbooks, patents, standards, Chinese studies, and overseas literature resources. The database allows people to access rich information in a very short time, so that people can get what they need and realize the optimal allocation of resources. The database

ISSN: 1813-4890

classifies and integrates various information resources, so that people can obtain the knowledge information they need according to their own needs. Moreover, the database is rich in content and resources. It integrates the information resources of various industries. Therefore, the information is comprehensive and specific, and the content is rich, which can meet the needs of people in all walks of life.

5. Current Status of Database Technology Application in the Context of Big Data

5.1. Distributed Clustered Database System

Distributed database system is composed of several computer systems. Each computer is like a warehouse storing data information. Different computers connect to each other through the network to share information. All computer systems are organically integrated into a whole through database technology. Each computer in the system can extract and manage data information from other computers, and the performance of multiple computers can be combined through database technology to provide users with better services[5].

Distributed databases are logically a unified whole and physically stored on different physical nodes. An application can access databases distributed in different geographic locations through a network connection. Its distribution shows that the data in the database is not stored on the same site. Rather, it is not stored on the same computer's storage device. This is the difference from a centralized database. From the user's point of view, a distributed database system is logically the same as a centralized database system, and users can perform global applications on any site. Just as the data is stored on the same computer and managed by a single database management system (DBMS), users don't feel different.

In a distributed database, materialized views are involved. The so-called materialized views refer to reports that have multiple base tables. By comparing and analyzing the views, you can know that they hold various data information in the reports. When there is a certain change in the base table data, through the use of materialized view to achieve data preservation, users can use manual refresh to synchronize the data, ensuring the authenticity and accuracy of the data. In materialized view, it allows users to remotely transfer and manage data on the local platform, but this operation only reads replica information [6]. The distributed cluster database architecture diagram, as shown in Figure 1:

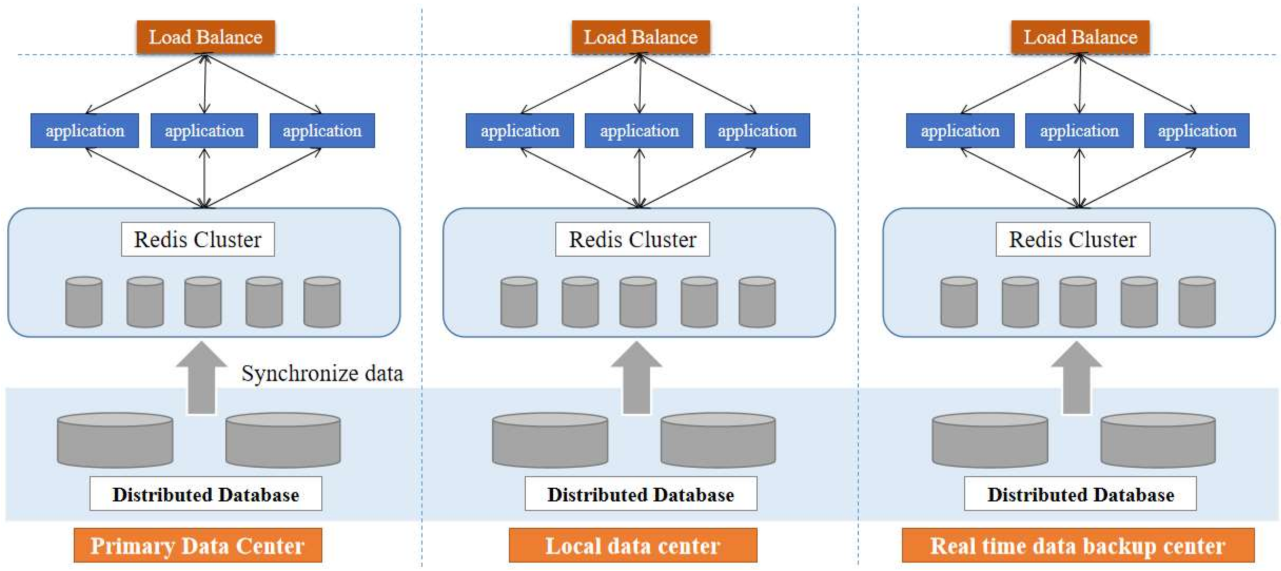

Figure 1. Distributed Cluster Database Architecture Diagram

ISSN: 1813-4890

5.2. Real-time Database System

In recent years, the application of real-time database system has successfully achieved the integration of process production data of field control system, established the data communication between enterprise management system and the underlying control system, and laid a foundation for the comprehensive integration of real-time information of the whole factory process.

Real-time database system is the supporting software for developing real-time control system, data collection system, CIMS system, etc. In the process industry, a large number of real-time database systems are used for control system monitoring, advanced control and optimization control, and provide real-time data services and a variety of data management functions for enterprise production management and scheduling, data analysis, decision support and remote online browsing.

Real-time database has become the basic data platform of enterprise information technology. It can directly collect and obtain all kinds of data in the process of enterprise operation in real-time, and convert it into effective public information for all kinds of business, to meet the requirements of real-time information integrity, consistency and security sharing among enterprise production management, enterprise process monitoring and enterprise operation management. It can build a bridge between enterprise automation system and management information system. Help the professional management departments of the enterprise use these critical real-time information to improve the operational efficiency of production and sales[7].

6. Common Database Technologies in the Context of Big Data

6.1. Cloud Database Technology

With the increasing popularity of computer networks, the amount of data information generated in the network is also growing. Traditional local databases have been unable to support such a large amount of data information storage. Therefore, cloud databases based on cloud technology can more effectively cope with this situation, thereby improving the efficiency of database application and providing better support for computer network services.

Cloud computing can be divided into three different categories, IaaS, PaaS, and SaaS, depending on the type of service. With the wide application of SaaS, cloud databases have been generated and developed, greatly improving the storage capacity of databases, eliminating duplicate resources, providing good convenience for software upgrade and hardware update. Cloud databases have many advantages, such as supporting efficient resource distribution, multi-tenant form, high availability, good scalability, and so on. Cloud database technology is a major direction in the future development of database technology[8].

(1) Features of Cloud Databases

1) Distributed storage of data

Because cloud databases are virtual databases, they can be modeled differently depending on the level of cloud database service purchased by users. Typically, cloud databases split up data and store it in different virtual machines based on the number of concurrent connections and the size of the capacity. And in the process of splitting, it has a strong intelligent algorithm to achieve the goal of improving storage efficiency.

2) High availability of data services

Traditional databases have the drawback of low availability, while cloud databases can choose storage behavior based on different types of database products and hosts. Usually, redundancy is a mechanism that means having multiple database hosts in the process of data processing. Its main goal is to be able to switch data services within a short period of time in the event of a

ISSN: 1813-4890

host or database failure, so that you can see the high availability of data services in cloud databases.

3) Isolation of data services

When there are multiple database instances running in the same physical host, there must be competition for resources between databases, which will seriously affect the quality and efficiency of database services. However, using a cloud database isolates the database instances, and since there is no impact between them, the quality and efficiency of the database instances are guaranteed.

(2) Common Cloud Databases

1) Amazon Cloud Database

As a leader in the cloud database market, Amazon has a positive role to play in the development of cloud databases. In Amazon, not only EC2 computing services and S3 storage services are provided to users, but cloud-based database service Dynamo is also provided. In addition, at this stage, Amazon has been working with other manufacturers and can be well used in different database platforms[9].

2) Google Cloud Database

Big Table represents Google's cloud databases, and its primary purpose is to process both formatted and semi-formatted data.

3) Microsoft Cloud Database

The Azure platform is Microsoft's cloud database product, which enables the use of SQL Server in the network to create, query and edit data in the database. The features of Azure platform mainly include the following points: ①Azure is a relational database;②Ability to support a large number of data types; ③ Support for stored procedures; ④ Supporting transactions in the cloud.

4) Aliyun Database

Ali Cloud Relational Database Service (RDS) is a stable, reliable and scalable online database service. Based on Alibaba Cloud Distributed File System and SSD disk high performance storage, RDS supports MySQL, SQL Server, PostgreSQL, PPAS (Postgre Plus Advanced Server, highly compatible with Oracle database) and MariDB TX engine, and provides a complete set of solutions for disaster tolerance, backup, recovery, monitoring, migration, etc. to thoroughly solve the troubles of database maintenance.

5) Huawei Cloud Database

Currently, Huawei's database tool services support both open source database services and self-study GaussDB database services. Currently, mature commercial services include DRS, DAS, DDM, which provides cloud-based data replication services for customers.

6) Baidu Cloud Database

RDS (Relational Database Service) is a stable, reliable and scalable online database service. Based on Baidu Cloud Distributed File System and High Performance Storage, RDS supports MySQL, SQL Server, PostgreSQL and PPAS (Postgre Plus Advanced Server, a highly compatible Oracle database) engine, and provides a complete set of solutions for disaster tolerance, backup, recovery, monitoring, migration, etc. to thoroughly solve the troubles of database operation and maintenance.

6.2. Graph Database Technology

With the continuous development of big data technology, how to collect data from various fields, and mining the value of big data through data association and analysis has become the focus of attention of various industries. In the exploration of technology, graph processing technology provides a new way for data to drive better development of the industry.

6.2.1. Chart Database Features

Graphic data, which structurally preserves knowledge by changing entity and point of relationship, is a model expression based on transaction-related relationship and has natural interpret-ability of data, which is highly recommended by academia and industry. In data association analysis, traditional relational databases require a large number of association operations, which are acceptable in the case of small-scale data. However, as the data size increases, the performance of the association operations will be exponentially degraded.

Compared with traditional relational databases and NoSQL databases, graph databases provide efficient related queries and complete entity information with their rich and complete relational expression. Graph databases have obvious advantages in processing large-scale and highly correlated data. Compared with relational databases, graph databases have a decentralized and distributed storage architecture, and integrate mature graph theory algorithms to enable faster relational queries[6].

At present, the main structure of graphics database in the market is shown in Figure 2. As a whole, the hierarchical design mode is composed of three layers: interface layer, computing layer and storage layer.

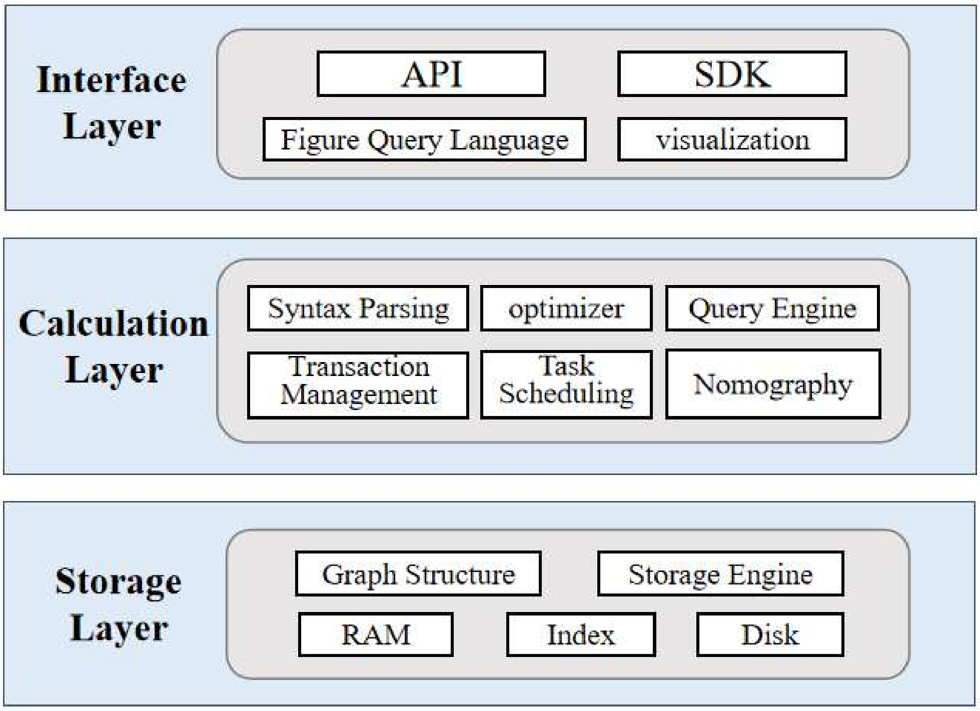

Figure 2. Diagram Database System Architecture

6.2.2. Common Map Database Products in China

(1) HugeGraph

HugeGraph is an open source distributed map database of Baidu. Supports standard Apache Tinkerpop Gremlin graph query language, attribute maps, and 100 billion-level relational data. Supports multiple backend storage (Cassandra, HBase, RocksDB, MySQL, PostgreSQL, ScyllaDB); Supports all kinds of indexes (secondary index, range index, full-text index, union index, all without relying on third-party index libraries); Provides a visual Web interface for graphic modeling, data import, and graphic analysis; Provides import tools to support importing data from a variety of data sources into diagrams, including CSV, HDFS, relational databases (MySQL, Oracle, SQL Server, PostgreSQL); Supports REST interfaces and provides 10+ generic graph algorithms; Supports integration with data systems such as Hadoop, Spark GraphX.

(2) GDB

GDB (Graph Database) is a self-developed supporting attribute map model developed by Alibaba to handle highly connected data queries and storage, and a real-time and reliable online map database service. Support for the TinkerPop Gremlin query language can help users quickly build applications based on highly connected datasets. GDB is a cloud-native, self-service database service, which means that GDB is an out-of-the-box service that can be used by public cloud applications without having to use some open source or other commercial versions to host operations and maintenance. GDB is easy to use, existing graph systems are more inclined to analysis classes, GDB has faster query and update performance, and graph database GDB has millisecond query time and millions of traversal capabilities.

(3) GeaBase

GeaBase (Graph Exploration and Analytics Database) is a simple and easy-to-use high performance - Financial - distributed - real-time - graphics database developed entirely by Ant Golden Garment. With its unique data organization and distributed parallel computing algorithm, GeaBase can quickly and efficiently query the relational information of data, thus satisfying various scenarios in the financial field for very large complex relational networks. In GeaBase, you can model online data, update data real-time through the client, import data flexibly in a variety of ways (through ODPS online import data and file upload offline import data), and conduct visual query testing.

6.3. Time Series Database Technology

A time series database is called a time series database (TSDB). Specialized databases for storing and managing time series data are optimized databases for ingesting, processing, and storing time stamp data. Compared with regular relational database SQL, the biggest difference is that time-series databases are databases recorded at regular intervals indexed by time.

Time series databases generally have the following characteristics:

(1) High throughput write capability

This is tailored to the feature that time series business continuously produces a large amount of data. To achieve high throughput writing of the system, two basic technical requirements must be met: the system has horizontal scalability and single-machine LSM architecture.

(2) Data Hierarchical Storage/TTL

This is a customized technical feature for the thermal and cold properties of time series data. Hierarchical data storage requires the ability to store data at the nearest hour level in memory, data at the nearest day level in SSD, data older in cheaper HDD, or to expire directly using TTL.

(3) High compression rate

There are two considerations for providing a high compression rate. One is cost savings, which is easy to understand. Compressing 1T data to 100G can reduce 900G of hard disk overhead, which is tempting for the business. Another aspect is that compressed data can be more easily stored in memory. For example, the last 3 hours'data is 1T, and I only have 100G of memory now. If not compressed, 900G of data will be forced to put on the hard disk, which will result in very high query overhead. Compression will put all 1T of data in memory and very good query performance.

(4) Efficient time window query capability.

The query requirements of time series business are divided into two categories: one is real-time data query, which reflects the status of the current monitoring objects; The second is mainly to query the historical data of a certain period of time, which is very large. At this time, it needs to optimize the query for a large number of data in the time window.

(5) Multidimensional query capability

ISSN: 1813-4890

Time series data usually has multiple dimension labels to depict a single data. How to query efficiently based on random dimensions is a problem that must be solved. Bitmap indexing or inverted indexing techniques are often considered.

(6) High efficient aggregation capacity

A common requirement for time series operations is aggregate statistical report queries, such as the total number of exceptions to an interface in the last day in a sentinel system or the maximum time spent executing an interface. Pre-aggregation is the most sophisticated way in which basic aggregation is done as data is written in.

(7) Bulk Deletion Ability

Sequential business needs batch deletion of expired data.

(8) Transactional capability is usually not required

Unlike traditional relational databases, traditional relational databases focus on adding, deleting, checking and transaction functions, while time series databases write large amounts of data and read and query data for a period of time.

The most popular open source time series databases in the industry are InfluxDB, OpenTSDB, Prometheus, Graphite and so on.

7. Future Trends of Database Technology

7.1. Multi-scene Fusion

From a product perspective, different scenarios have different characteristics and different requirements for database read-write performance, throughput, consistency, and so on. To support different requirements in different scenarios, database diversity is an essential choice. For example, in the scenario of the Internet of Things, the amount of data written is especially large and the real-time requirement is especially high, but the data is naturally time-ordered and static, so time-series databases have advantages over traditional transactional databases. The HTAP/NewSQL databases, multimodal databases, unified management platforms in the market today meet the needs of enterprise simplification and integration, so "integration" in the context of multiple scenarios is also a trend that should not be ignored.

7.2. Open Source Model

From the perspective of industrial development, on the one hand, the open source mode improves the "efficiency" of database product development. It opens the source code and avoids the redevelopment of basic programs by developers. On the other hand, it also helps the technological "innovation" of products. Open source communities maximize the global resources and provide space for developers to exchange and discuss, thus accelerating the birth of innovation inspiration. For manufacturers, although seemingly, enterprises can not obtain direct benefits by deploying open source, in this process, they can lay out the ecological construction of products (including personnel training, market education, practice feedback, enterprise culture, product influence, supporting peripheral products, etc.) so as to gain a favorable strategic position for themselves. At the turning point of the development of mobile Internet to industrial internet, open source mode is not the breaking edge for cloud manufacturers, traditional manufacturers and emerging manufacturers to expand their markets.

7.3. Intelligence

With the development of mobile Internet to industrial internet, data grows exponentially every day and presents multimodal characteristics. In the face of complex and massive data, more and more kinds of databases appear, and the scope of debugging needs to be wider and wider. However, most of the optimization tasks still fall on the DBA, which needs to be manually tuned so that the manual ability can not keep up with the development of the database. Artificial intelligence can make up for the shortage of human capabilities and solve many database problems that exist for many years, such as resource scheduling, index design and optimization, query optimization, load balancing design, cache failure, etc. AI uses optimization algorithms to effectively predict, analyze, and automate tasks, reducing labor costs and greatly improving database performance. Especially in the future, with the wider popularity of databases on the cloud, intelligent resource scheduling will become the next topic that each supplier needs to face[10].

8. Conclusion

In the big data environment, the requirements for data processing efficiency, security, stability and other issues are constantly improving. The optimization and upgrading of database technology is the key to solving data processing problems in the big data environment. Distributed database systems and real-time database systems play an important role in improving database security, operating efficiency, etc., and promoting database technology to better deal with the huge amount of information, complex content, security Stability and many other issues, ultimately better meet the personalized needs of users.

In the big data environment, cloud database technology, graph database technology and time series database technology are widely and deeply integrated with multimedia, artificial intelligence and other fields. Database technology will get more improvement and produce more types with the development of the times. Database technology will gradually progress and improve, driven by technology in other fields, and provide support for progress in other fields to promote the development of big data environment.

Acknowledgments

The article was supported by the Project of the First Batch of Ideological and Political Demonstration Courses in Zhejiang Province(ZheJiaoHan [2021]No.47),the Course name: Relational Database Application (No.37).

The article was also supported by the "Curriculum ideological and political" projects of Wenzhou Polytechnic in 2021 (No.WZYSZZY2104),the Major Education Reform Project of Wenzhou Polytechnic (No.WZYZD202202), Wenzhou key scientific and technological projects (No.ZG2020031)First batch of Ideological and political Demonstration Courses in Zhejiang Province.

References

[1] You Yanhou.The Application of Database Technology in the Era of Big Data. [J].China Computer& Communication,Volume 23, 2018. PP 15.

[2] Cai Ming.Research on the Application of Database Technology in the Age of Big Data[J].Scientific and Technological Innovation,Volume 5, 2019. PP 94.

[3] Li Xueren.Database Technology and Its Aapplication in Big Data Environment[J].Electronic Technology & Software Engineering. Volume 19, 2021. PP 150.

[4] Dai Yun.Research on the Status Quo and Development Direction of Database TechnologyBased on Big Data Background[J].Digital Technology &Application, Volume 37, 2019.PP 86-87.

[5] Li Haifeng.Characteristics and application countermeasures of database technology in the era of big data[J].Wireless Internet Technology,Volume 4, 2019.PP 147-148.

[6] LIU Yu-Ning, FAN Bing-Bing. Survey on Graph Database Development[J]. Computer Systems & Applications,Volume 8,2022.PP 1-16.

[7] Foead D, Ghifari A, Kusuma MB, et al. A systematicliterature review of A* pathfinding. Procedia ComputerScience, 2021.PP 179: 507–514.

[8] D'Agostino D, Liò P, Aldinucci M, et al. Advantages of usinggraph databases to explore chromatin conformation captureexperiments. BMC Bioinformatics, Volume 2,2021.PP 43.

[9] Rebecca Taft, Irfan Sharif, Andrei Matei, et al. The resilient GEO - dis-tributed SQL database[C]/ /. In Proceedings of the 2020 ACM SIG-MOD International Conference on Management of Data (SIGMOD'20) .Association for Computing Machinery,2020.

[10] Lin Ma, Dana Van Aken, Ahmed Hefny, et al. Query - based workloadforecasting for self - driving database management systems[C]/ / In Proceedings of the 2018 International Conference on Management of Data (SIGMOD'18) . Association for Computing Machinery, 2018.