Technical Analysis on NoSQL Database under Background of Big Data

Chongli Zhong

Xi'an Eurasia University Xi'an City, Shaanxi Province 710065 China

Keywords: Big Data; Nosql; Database; Technical Analysis

Abstract: With the rapid development of Internet technology, network intelligence has become a trend. With the development of network intelligence, the demand for processing capacity of server's increases correspondingly, and the amount of data that the server needs to process also presents an explosive growth. Especially after entering the era of big data, the emergence of diversified data such as e-commerce transactions and network interactions makes the data scale of servers grow significantly. High Performance makes more pressure on servers in reading and writing. The hard disk managing database storage fails to meet the requirements on current input and output, which lags behind the development of the times of this regard. Due to the priority of data records and low efficiency of SQL, it is difficult to meet the demand of massive data.

The amount of data increased dramatically after entering the new period, so the requirements on processing data by processors also increased. As the relational database is limited in storage record, the SQL query is lablled with low efficiency and is hard to meet the requirements for massive data storage, the rapid increase of server nodes are difficult to meet the demand of the service, and it is difficult to expand relational database, thus the requirements on high expansibility and high availability are difficult to be satisfied. Therefore, the research and development of NoSQL is conducted and NoSQL is applied into practical scenarios. With the vigorous development of science and technology and the rapid development of society, the intelligent development of modern technology has been gradually paid attention to by the general public, and the national information level has become an important symbol for measurement of the development national level of modernization .The development of NoSQL database has become the focus of current research.

1. Development of NoSQL Database

With the development of social science and technology, NoSQL database has been rapidly applied in many fields after its birth, and many professional fields are highly attached to NoSQL database with high dependence, which has been deeply penetrated into people's production activities and life, playing a key role in the rapid development of national economy. With the values of NoSQL database is gradually explored, people show more and more emphasis on NoSQL database, but there are many barriers to the development of the related work in practical application of technology of NoSQL database. Prospects for application of NoSQL database are very broad; as a result, the research on it for the benefit of society is conducive to the development of economy, science and technology.

Not Only SQL refers to an abnormally relational distributed principle that does not have a data storage system that conforms to the ACID principle and is designed in accordance with typical technical theories and principles. In simple terms, the CAP theory is a theory that can meet the requirements of multiple fault tolerance. A distributed system fails to meet the requirements of multiple systems (at least two systems).The design principle of BASE emphasizes reading and writing efficiency, data capacity, and scalability requirements of the system. Compared with the traditional database, the database is simpler and the design focuses on distributed data management. The NoSQL is characterized with many advantages in many aspects, such as strong scalability, high performance, high flexibility and high performance in actual application. In fact, NoSQL databases discard the connectivity of traditional data and information, which enables the databases independent of each other, so it is very easy to establish the distribution between the databases. The reading and writing of NoSQL database are shown with good performance, because of the simple

distributed and simple structure of the database and the storage data format at any time can be achieved. Under the premise of not affecting the performance, NoSQL database can be easily structured with a more usable architecture, and the establishment of database under this background can be better in carrying out big data processing [1].

2. System Architecture of NoSQL Database

There are many solutions to NoSQL database, including master-slave, P2P and other two structures, and the two structures are quite different, the operation and maintenance protocol is also quite different [2].

2.1 In structure design for Master-slave, the storage system is responsible for managing the system and for monitoring the status of the Slave node. To allocate the data to be stored in each Slave node, which can be regarded as the port for data storage. Only one Master node is designed and the operation state of this node will affect the overall performance of the entire data storage system. Therefore, multiple sub-nodes are needed to ensure the normal operation of the nodes and ensure the normal operation of the system under intensive operation, thus the fault tolerance of the system can be assured. Slave node is responsible for maintaining index table of data storage in a certain region. With the increase of data storage volume in the system, multiple nodes at horizontal level should be expanded for the system. In the architecture system, Master system should be in a monitoring state, and the communication among nodes of SLave should avoid to reduce the communication cost [3].Every Slave node should report its operation situation and load cleaning to MAster node regularly, and M node plays a role in unified regulation and allocation of storage tasks.As for search engines, the typical and excellent big data storage structure systems include Bitable, Apache and Habse of Google.

2.2 P2P structure belongs to the design with ring structure, the Node in the system by Consistent and Hashing algorithm is formed the calculation logic with a ring structure, every Node can store data to ensure that there is no M (Master) nodes. Besides, it will be easier to add Node, and it will be feasible as the data exchanges are conducted between the new nodes and adjacent nodes, which will not result in a very significant effect on the system [4].The internal information of the nodes needs to be displayed at the whole. There are no data systems such as Cassandra and Dynamo in this system.

3. Data Storage

3.1 Key-Value

There are different types of databases, so there exists greater differences for data storage models. The Key - value is the more frequently used storage type, the data are stored in form of e Key value and data model serves as a map, the Key refers to the keyword to find the address of data, the value is used for storage of the contents of data. The data model and hash function of this relation are used to realize the key to value mapping, so as to improve the reading and writing ability and storage ability of the data for rapid data processing and the query is made through major key, but the data can be available through Key.

3.2 Document Storage

Document storage can be used in the same way as defining a table structure. Document storage can be used in a variety of formats and it is capable of storing system logs and unstructured data. Different from the key-value storage format, data can be obtained through complex query criteria [5].This system is not endowed with the ability in processing things and is short of the capability of Join. Other processing methods can be basically achieved by this system. Therefore, NoSQL database can be easily used by this system, and the systems using this mode include MongoBD and CouchBD.

3.3 Column-based Storage

Column storage is adopted to store relevant data in form of units, which is preferred to enter into the data storage system in column units, the query of data can be achieved at random by means of retrieval. The use of columns enables the scalability of the database to be extended. Even if the data is increased, the processing speed will not be reduced. Therefore, a large amount of data is stored in the system. There are many differences regarding column storage database and the unit storage system, so the application difficulty is also relatively large. This type of database includes Cassandra et al. [6].

3.4 Image Storage Figure

There is inadequate research on the image-based database, only a few databases can storage images, the databases with the capability to store images and the image-based database is built based on image theory, similarly, the nodes and attributes are adopted. Nodes serve as entities, which hold information about attributes and nodes, and edges are used to connect nodes to represent the relationship among them. Some data sets of graph database are very fast and the images can be directly mapped to the application degree of objects. Neo4j [7] is a more typical storage model system [7].

4. Key Technology of NoSQL Database

4.1 Workplace Technology

NoSQL database provides workplace technology with client API functions for creating, updating, reading and deleting data. The basic feature of the function is the key value. Get key value is read and the corresponding value is taken out. Delete means to delete a specified value. The PUt key means to create or update the specified data. There are also many functions in NoSQL database that are convenient for customers to use and are capable of meeting the requirements for users to process more consistent data. For example, data conversion of the same types is not allowed in this system, and summarization of data with filtering and operation can be achieved for arbitrary expressions, which are then combined with NoSQL, Map Reduce and other data. The latter is a large-scale computing framework, which can be the input and output framework of this model by the NoSQL database [8].

4.2 Database

NoSQL database is the data that needs to be processed, and the database increase with the development of the society as a whole, the scalability of the system needs to be improved and it is very important to add and remove nodes. At present, the commonly used technology is Consistent Hashing technology, which can as far as possible decrease or change the key-value mapping relationship to meet the requirements on monotonicity when removing or adding nodes [9].As it is necessary to add a node, the data should be in counter clockwise and the value of the former node should be placed to the next node. If a node is removed, the data on the node should be migrated to the next node in clockwise direction. Adding or removing service nodes from the system as a whole will only affect neighboring nodes without causing influence on other nodes.

4.3 Duplication

In order to avoid problems of data nodes, resulting in data loss, and ensure high availability of the data system, data can be backed up after replication and transferred to multiple nodes[10].The implementation mechanism of data replication is capable of distributing the data to the secondary nodes according to certain strategies by using the communication between nodes and secondary nodes. NoSQL database uses distributed service framework to regulate and coordinate internal nodes. The service node is responsible for writing requests, and the writing function is completed by the subsidiary nodes.

Conclusion

In conclusion, NoSQL database has strong extensibility, practicability and usability, which make up for the defects of traditional database. Among the traditional database and there exists the problems of high concurrency performance of data, defects in the processing technology. Therefore, this technology adopts a non-relational way to solve the problem in big data storage and management to replace the existing problems in the traditional database in a short time, the emergence of NoSQL database quickly solved this problem, but also more problems are generated and the prospect is promising in the future development.

References

[1]Gong Chang. Technical Analysis on NoSQL Database under Big Data[J]. Information Recording Materials, 2018, 019(006):118-119.

[2]Liu Yangna. Analysis on distributed storage strategy of spatial big data based on NoSQL database[J]. Digital Technology and Application, 2018, 036(002):77, 79.

[3]Song Junsu. Research and Application of Query Technology based on NoSQL Database in big Data Environment [J]. Computer Programming Skills and Maintenance, 2019, 404(02):78-79+118.

[4]Xu Ming. Research on storage technology of NoSQL database in big data environment[J]. Digital Users, 2018, 024(048):128,131.

[5]Tan Junkai. The Distributed Storage Strategy of Spatial Big data based on NoSQL Database[J]. Information Weekly, 2019(15):0083-0083.

[6]Meng Xiaofeng, Chen Shimin, Luo Qiong. Introduction to the topic of data Management for new Hardware in 2018[J]. Journal of Computer Research and Development, 2018, 55(2):227-228.

[7]Xue Tao. Practical Exploration of Big Data Query Technology based on NoSQL Database [J]. Computer Programming Skills and Maintenance, 2018, 000(011):89-90, 13

[8]Shi Xiaofeng. Research on Archival Big Data Storage and Retrieval Scheme based on Distributed NoSQL Database[J]. Computer Application and Software, 2019, 036(005):15-20.

[9]Yang Peng, Lin Junhui. A Massive Unstructured Compounds Network Data Processing scheme based on MongoDB and Hadoop[J]. Microelectronics & Computer, 2018.68-72

[10]Yangpeng, Lin Junhui. A Scheme for Massive Unstructured Iot Data Processing Based on MongoDB and Hadoop [J].Microelectronics & Computer, 2018, 035(004):68-72, 78.