IJCSN - International Journal of Computer Science and Network, Volume 3, Issue 6, December 2014
ISSN (Online): 2277-5420 **www.IJCSN.org**
Impact Factor: 0.274

NoSQL: A Database for Cloud Computing

Heena Khan

Faculty of Information Technology, Government Polytechnic,
Pune, Maharashtra, India.

Abstract - With various advancements in the field of computing, scalability, resource utilization and power savings is being given higher priorities. NoSQL database and cloud computing goes hand in hand when compared with SQL databases. The main advantage of NoSQL databases is agility. The demand for NoSQL database is growing because most of the cloud applications built , demands high availability , speed , failover options , fault tolerance and consistency which a traditional relational database fails to offer the modern web applications.

Keywords - *Big users, Big data , Scalability ,Elasticity.*

1. Introduction

In recent years, advances in Web technology and the proliferation of sensors and mobile devices connected to the Internet have resulted in the generation of immense data sets that need to be processed and stored. Just on Facebook, 2.4 billion content items are shared among friends every day . Today, businesses generate massive volume of data which has grown too big to be managed and analyzed by traditional data processing tools. Indeed, traditional relational database management systems (RDBMS) were designed in an era when the available hardware, as well as the storage and processing requirements, were very different than they are today . Therefore, these solutions have been encountering many challenges in meeting the performance and scaling requirements of this "Big Data" reality.

2. Big Data and Cloud Computing

Big Data is a term used to refer to massive and complex datasets made up of a variety of data structures, including structured, semi-structured, and unstructured data. According to the Gartner group, Big Data can be defined by 3Vs: volume, velocity, and variety. Today, businesses are aware that this huge volume of data can be used to generate new opportunities and process improvements through their processing and analysis. At about the same time, cloud computing has also emerged as a computational paradigm for on-demand network access to a shared pool of computing resources (e.g., network, servers, storage, applications, and services) that can be rapidly provisioned with minimal management effort. Cloud computing is associated with service provisioning, in which service providers offer computer-based services to consumers over the network. Often these services are based on a pay-per-use model where the consumer pays only for the resources used. Overall, a cloud computing model aims to provide benefits in terms of lesser up-front investment, lower operating costs, higher scalability, and elasticity, easy access through the Web, and reduced business risks and maintenance expenses.

Due to such characteristics of cloud computing, many applications have been created in or migrated to cloud environments over the last few years. In fact, it is interesting to notice the extent of synergy between the processing requirements of Big Data applications, and the availability and scalability of computational resources offered by cloud services. Nevertheless, the effective leveraging of cloud infrastructure requires careful design and implementation of applications and data management systems. Cloud environments impose new requirements to data management; specifically, a cloud data management system needs to have:

- Scalability and high performance, because today's applications are experiencing continuous growth in terms of the data they need to store, the users they must serve, and the throughput they should provide;
- Elasticity, as cloud applications can be subjected to enormous fluctuations in their access patterns;
- Ability to run on commodity heterogeneous servers, as most cloud environments are based on them;
- Fault tolerance, given that commodity machines are much more prone to fail than high-end servers;
- Security and privacy features, because the data may now be stored on third-party premises on resources shared among different tenants;

- Availability, as critical applications have also been moving to the cloud and cannot afford extended periods of downtime.

Faced with the challenges that traditional RDBMSs encounter in handling Big Data and in satisfying the cloud requirements described above, a number of specialized solutions have emerged in the last few years in an attempt to address these concerns. The so-called NoSQL data stores present themselves as data processing alternatives that can handle this huge volume of data and provide the required scalability. Despite the appropriateness of NoSQL data stores as cloud data management systems, the immense number of existing solutions (over 120) and the discrepancies among them make it difficult to formulate a perspective on the domain and even more challenging to select the appropriate solution for a problem at hand. This paper reviews NoSQL data stores with the intent of filling this gap.

3. Migrating Towards NoSQL Databases

Out of the many different data-model architectures, the relational data model architecture has been dominating since the 80s, with the implementations like Oracle database, MySQL and Microsoft SQL Servers. Later, however, the relational databases leads to the problems in many cases because of its data modeling techniques. The exponential growth of complexity of data generated by social networks, sensors, real time systems, and global users etc, and the storage of this huge amount of data on big distributed system, demands evolution of new data management model . Organizations that collect large amount of unstructured and ever changing data are increasingly turning to non-relational or NoSQL databases.

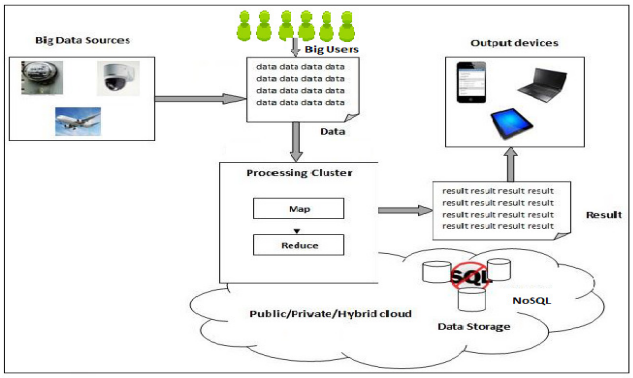

Fig -1: Organizations migrating towards NoSQL database

NoSQL databases focus on analytical processing of large scale datasets in warehouses, offering increased scalability over commodity hardware and servers. Computational and storage requirements of applications such as for Big Data Analytics, Business Intelligence and social networking over peta-byte datasets have published SQL-like Centralized database to their limits. This led to the development of non-relational data stores called NoSQL databases which are distributed and horizontally scalable, such as Google's Bigtable and its open source implementation HBase and facebook Cassandra. The emergence of distributed key-value stores, such as Cassandra and Voldemort , proves the efficiency and cost effectiveness of their approaches.

The limitations with non-relational databases are it is hard to scale with Data warehousing, Grid, Web 2.0 and cloud applications. The strict relational schema of relational databases can be a burden for web applications like blogs, which consists of much different kind of attributes. Text, audios, pictures, videos, real time data and other fast changing information have to be stored within multiple tables. Since such web applications are very agile, underlying database have to be flexible and dynamic as well in order to support easy schema evaluation process. NoSQL systems exhibit the ability to store and index arbitrarily Big Data sets while enabling a large amount of concurrent user requests. Main advantages of NoSQL are the following aspects:

- Reading and writing data quickly
- Support mass storage
- Easy to expand
- Low cost.

3.1 Big Data

Capturing and collecting the data becomes easier and can be access via third parties such as D&B, Facebook, and Twitter etc. User related personal information, location dependent data, graph oriented data, user generated data, system logging data, and real time generated data are just a few examples of the ever-changing and expanding blocks of data being collected. It's not amazing that developers feels the increasing value in leveraging this data to improve existing applications and develops new ones made possible by it. The application of the data is continuously changing the nature of web life that includes web communication, online shopping, web advertisement, entertainment hobbies, and relationship management. The Applications that doesn't meet the current big data market trends will quickly fall behind.

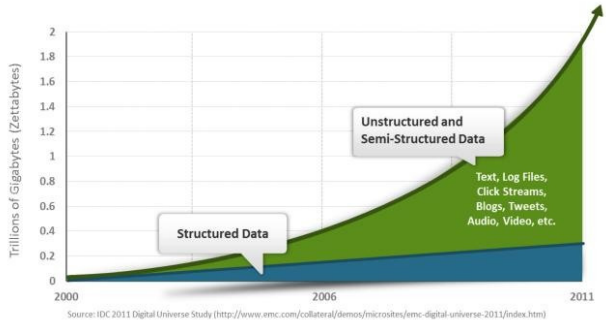

Fig -2: Big Data

The amount of data is growing rapidly, and the nature of data is changing as well. The various kinds of data is collected and it demands for a very different type of database which should be very flexible and easily incorporate any new type of data. So the database must have a capability of efficiently storing and very fast access to the new types of data that includes semi-structured and unstructured data.

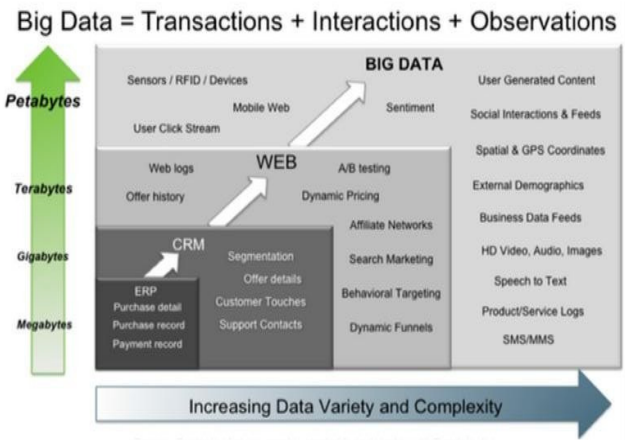

Fig -3: Big Data Transactions with Interactions and Observations

Unfortunately, the relational databases have very poor features to quickly adopt new types of data because of its rigid and static schema based approach, and is not suitable for semi-structured and unstructured data. Finally, the NoSQL meets the growing trends of storage, processing and retrieval of data by providing a flexible, schema-less data model that maps the organization's requirement and simplifies the communication between the application and database, that results in less writing code, debugging and maintenance becomes more easier.

3.2 Big Users

Not then long ago, one thousand users of an application treated as a lot of users and ten thousand users treated as an extreme case. But today with the emerging field of cloud, many applications are hosted on it and it is made available over the internet 24 hours a day and 365 days a year so that it supports many users globally . A survey shows that more than two billion peoples are connected to worldwide and amount of time they spent online per day is gradually increasing and results in increase number of concurrent users. And now many applications have millions of different daily users.

Fig 4: Big Users: With the growth in global Internet use, the number of hours spent online, and increase in Smartphone users, it's not uncommon for apps to have millions of users per day.

Because of huge number of concurrent users, it is very difficult to predict at application usage requirement. It is very much important that an application dynamically support rapidly growing huge number of concurrent users. To achieve this goal, an application must possess following features:

- An application can have features that supports zero to millions of users.
- Application must support frequent active global users while considering those users which access application for some time.
- New applications can be dramatically scalable and provide higher fast access process.

The huge number of global users along with dynamic, flexible usage pattern is the driving force for easily scalable new database technology. Many application developers find very much complication to get scalability and faster access rate with relational database technologies. To overcome this limitation of relational database technology many application developers are turning toward NoSQL for help.

4. Cloud Computing and Internet Architecture

Cloud Computing was initially proposed by Google, Amazon and IBM. There are many definitions, and each described cloud computing from a different point of view. A comprehensive definition is ―Cloud computing is a platform (system) or a type of application. In a cloud computing environment, the server can be physical server or virtual server. Cloud computing describes a scalable

application which can access through the internet.‖ Not long ago, most applications are used by single user that runs on a single system and these applications uses a two-tier client server architecture supported by a limited number of users. Today, with the emergence of cloud, applications uses a three-tier internet architecture that run in a public or private cloud that support a huge number of global concurrent users. With this shift in software architecture, cloud provides many data intensive business services like platform-as-a-service, software-as-a-service and infrastructure-as-a-service and these service models have become more prevalent.

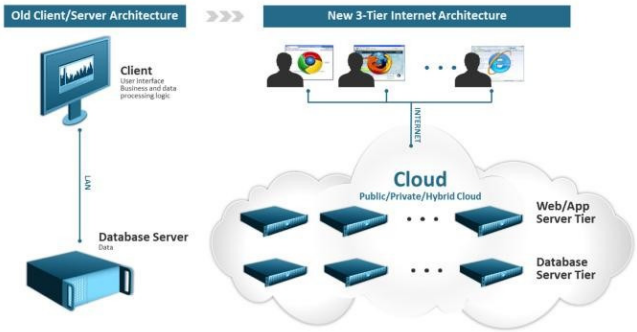

Fig -5: Applications today are increasingly developed using three-tier internet architecture, requiring a horizontally scalable database tier that easily scales with the number of users and amount of data that an application has.

In the three-tier architecture, users interact with the applications through a web browser or by using mobile apps that is connected to the internet. In the cloud, scale-out approach is used. If the number of global concurrent user increases then another commodity server is added to the web/application tier to manage the incoming traffic and this work will be done by a load balancer very beautifully. When we compared relational databases and NoSQL databases, relational databases are problematic because they are centralized, share-everything technology and scale-up rather than scale-out. NoSQL databases are emerged with scale-out approach and better fit for the three-tier internet architecture and cloud services.

5. Conclusion

Big Users, Big Data, and cloud computing are changing the way that many applications are being developed. The relational databases have dominated industries for many years, but NoSQL databases are now getting attention of application developers due to the following reasons:

- NoSQL databases provides schema-less dynamic flexible data model, that is most suitable for the big users and big data.
- NoSQL databases have an ability to scale dramatically to support global users and big data.
- NoSQL database provide an improved performance to satisfy big users expectations without compromising scalability.

References

[1] Abadi, D. J. (2009). Data management in the cloud: Limitations and opportunities. IEEE Data Eng. Bull, 32(1), 3-12.

[2] Oracle Databases from web: http://www.oracle.com/us/products/database/overview/index.html

[3] A B M Moniruzzaman and syed Akhtar Hossain, –NoSQL database: New Era of databases for Big data Analytics- Classification, Characteristics and camparision

[4] Konstantinou,I.,Angelou, E. Boumpouka,C., Tsoumakos,D., andKoziris,N(2011) October. –On the elasticity of NoSQL databases over cloud management platforms http://www.slideshare.net/adorepump/voldemort-nosql

[5] Jing Han, Haihong E, Guan Le, and jian DU, –Survey on NoSQL Databases‖. IEEE C 2011,978-1-4577-0208-2/11

[6] Chen Kang, Zheng Weimin Cloud Computing: System Instances and Current Research[J] Journal of Software, Vol.20, No.5, May 2009,pp.1337–1348.

[7] P. Mell and T. Grance, "Draft nist working definition of cloud computing - v15," 21. Aug 2009, 2009.

[8] Use relational DBMS, N. (2009). Saying good-bye to DBMSs, designing effective interfaces. Communications of the ACM, 52(9).

[9] Find Source: nosqltips.blogspot.com

First Author : I am **Heena F.Khan**.I am B.E.IT.
I have worked as a .net developer invZain Information System then as a lecturer in Bharti Vidyapeeth COE,Pune.Now I am working as a lecturer in Government Polytechnic ,Pune