A REVIEW ON CLOUD DATABASES AND THEIR CHALLENGES

Libin M Joseph[1], Dr. E.J. Thomson Fredrik[2]

ABSTRACT

Traditional databases have been used by the Information Technology industry for four decades. But, there has been a huge change in the IT industry regarding commercial applications during the last few years. Web-based applications have taken the place of standalone applications with dedicated servers and distributed servers with network storage. Nowadays Cloud computing is popular because of its low cost, scalability, flexibility etc. Cloud Databases such as Firebase Real-Time Database, Sherpa and Microsoft Azure are in common use. They have overcome the barriers of existing databases regarding scalability, ease of use, flexibility and dynamic provisioning. Data-critical applications related to data storage, and mining use the main features of cloud databases. These applications are scalable, multipurpose and elastic in nature. Databases used in such real time applications need to absorb four significant highlights. They are ACID (Atomicity, Consistency, Isolation and Durability) properties. But there are some difficulties to deploy them to the cloud. The aim of this article is to highlight the pros and cons of the cloud databases and to assess the limitations in creating them.

Index Terms—Cloud, Databases, Cloud Service Providers, Cloud Service

I INTRODUCTION

Cloud services are part of the new trends in networking and highly efficient hardware and software. A booming trend can be seen among PC users to replace the burden of tremendous storing of data and slow and complex calculations with servers in the cloud which save space, time and power. Subsequently, a growing service providing business, run by Cloud Service Providers (CSPs), has recently been started, particularly by major companies such as Amazon, Google and Microsoft. A CSPs' customers pay for the type of service and the usage provided [1].The volume of digital data increases at a phenomenal rate rendering many organizations' storage capacity inadequate. The outsourcing of data to cloud servers is a solution for storing more data on efficient centralized servers. Data management becomes one of CSP's tasks by storing data in the cloud, and the organization may concentrate on its primary duties. This reduces the cost of various resources like software, hardware and even hired specialists to support the system [2].Under Infrastructure service, a cloud computing atmosphere provides storage as a service through various CSPs, which enables clients or organisations to store information on remote servers. Cloud databases are considered by software engineers to be an intelligent response to store their application data in a flexible and extremely reachable backend. Database outsourcing presents another worldview, called “databases as a service” (DBaaS).Database providers have the framework for hosting outsourced databases on conveyed servers providing their customers with effective services to perform CRUD operations on the Internet.

[1]*Research Scholar, Department of CS,CA & IT, Karpagam Academy of Higher Education, Coimbatore, India*

[2]*Professor, Department of CS, CA & IT, Karpagam Academy of Higher Education Coimbatore, India*

II LITERATURE REVIEW

A lot of research study has been published in recent years on cloud databases and their problems. Some of them are be outlined here. Vodomin and Androcec [3], portrayed a realistic model of a migration tool that can be used for Microsoft SQL, PostgreSQL and MySQL databases. In the context of moving databases to cloud storage, this work has primarily led to the analysis of problems.

Abourezq and A.Idrissi [4], provided a standard for the principal database strategies outlined by CSPs as DBaaS (DataBase as a Service). They assessed features of solutions and ability to adapt Big Data apps.

In the article drafted by Arora and Gupta [5], "cloud databases and various frameworks" studied the best in class.The central goal of their paper was to survey and discuss the recent trends to explore and analyze the barriers and issues in the development of cloud databases technologies.Their paper's central objective was to survey and evaluate current trends and discover and assess the obstacles and challenges in cloud database development.

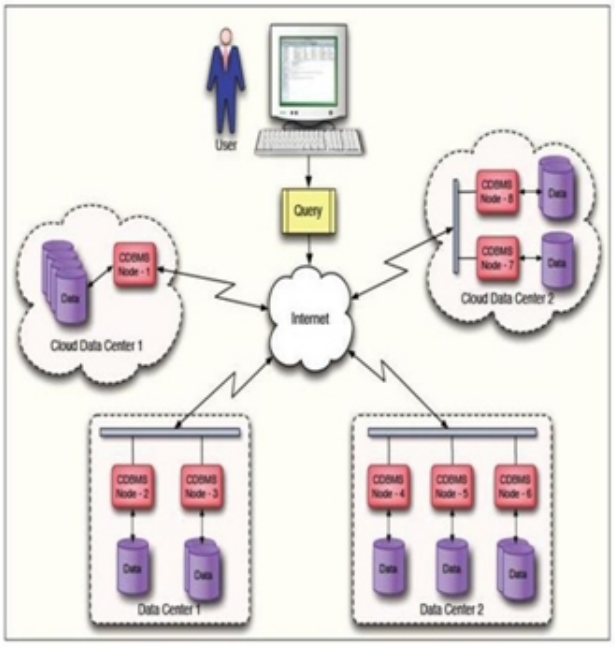
Alomari et al.[6] suggested a cloud storage data model and API for the cutting-edge NoSQL database. Implementing their suggested architecture involved three well-known databases for NoSQL, namelyAmazon SimpleDB, MongoDB and Google Datastore.The proposed system was set up with a high degree of adaptability and could be linked similarly to other frameworks of NoSQL.

Divyakant Agrawal etal. [7] provided a tutorial that extensively portrayed the difficulties faced by the programmers and database designers to develop and deploy apps on the Internet.

In [8]Ferretti et al. provided another design to avoid intermediate parts suggested to attain some sort of accessibility and extensibility proportionate to decoded cloud database services. This architecture gave an environment where various users could execute queries simultaneously.

III CLOUD DATABASES

Data are stored at data centers rather than on dedicated servers such as conventional data storage on various dynamic servers in cloud environments. When a database is stored, the user sees a simulated server, but the information is really stored on more than one information center servers. The actual location of the storage can vary as and when the cloud dynamically manages the available storage space. However, while the location is imaginary, the user sees fixed location for their data and can handle their storage space as if they were attached to their own PC. The cloud database maintains data in various fields on unique server farms. Today, many key-value storage techniques for data management such as Google's AppEngine, Amzon's AWS,Microsoft's Azure are used to keep databases on the cloud platform. To understand the cloud database infrastructure, cloud database structure is shown in Fig.1



A cloud database is a scalable content database running on a cloud computing platform that could be public, private or hybrid. There are two models of cloud databases, Traditional model cloud and Database-as - a-service (DBaaS). In the conventional model of cloud, the server database function on the company's infrastructure and monitoring will be carried out by the company's IT staff.

The cloud database connects on-demand clients with cloud service providers' servers via the Internet. Large availability, scalability, multitenancy and highly optimized allocation of resources are provided by cloud databases. The user can install, configure and maintain these databases on a Cloud server. This option is commonly referred to as "Do-it-Yourself" approach (DIY).Conventional DBMS deals with structured information kept together with its meta-data in databases. Although it is possible to use cloud databases both for unstructured and structured data. Each cloud provider has an alternative strategy to control and access the database, make the data convenient, and it is a testing task to perform between these frameworks.

IV CHALLENGESTO DEVELOP CLOUD DATABASES

A cloud database management system supports cloud figure components and extra databases to enhance the value of Hercules. The potential problems related to cloud databases are shown in Fig. 2.

1. Scalability

The most significant characteristic of cloud database is that the users usually expect to increase and decrease the size of data in databases easily to assure the needs of their business goals. It defines the facilities that are provided by the cloud databases to increase or decrease data significantly without any burden. So, the cloud systems should provide extensible database services to satisfy the needs of its users.

2. Internet Speed

As compared to the speed of internet used to obtain data from the data center, the speed of data recovery from data center is very high. This factor affects the efficiency of cloud databases even if the queries sent to databases are very fast. The data access from the data center depends on the speed of the internet.

3. Fault Tolerance and High Availability

Replicating data across broad geographic places is very important to ensure that the data are easily available, powerful, as well as highly stretchable in adapting inner problems in database. In general, the system availability is described as the availability of resource usability for all users [9].If there is an interruption due to a cloud service failure, the availability of databases may be affected, momentarily or permanently, resulting in severe data loss, in part or in full. Deficiencies in equipment, attacks like DOS and security deficiencies are severe threats to cloud database accessibility.

4. Integrity and Data Consistency

To ensure the integrity of data, it is important to control and assess the database users including the database administrators and authenticated users. It is a very difficult to ensure data consistency, and more so if data change very rapidly, especially in the case of value-based data. Cloud database follows BASE(Basically,Available,Soft State, Eventually Consistent) rather than ACID. The developers should not compromise integrity of data when they are moving

to database in the cloud. So, they should be careful in using BASE properties.

5. ***Interface for Query***\

Querying on a cloud distributed database is a significant challenge faced by cloud developers. Multiple cloud database nodes must be accessed by a distributed query. A streamlined and standardized query interface should be available to query database.

6. ***Database Security and Privacy***

If information is encrypted by a non-host key, it is not very safe. In order to impede unauthorized access, encrypted sensitive data are transferred to cloud. Any cloud-based application should not be capable of decrypting the data directly before accessing it. It is also a major challenge to provide safety and privacy to separate databases on the same hardware.

7. ***Data Portability and Interoperability***

Information Portability is the capability to perform implementation in the environments and systems of another cloud provider prepared for a particular cloud provider. Interoperability is the ability to supply some codes, regardless of their variations, that are adaptable enough to operate with different cloud providers.

V CONCLUSION

The whole database scenario has been altered by massive information produced by web-based apps. Cloud databases seem to be a great way to handle such information. Moreover, it is impossible for all organizations to set up costly data center infrastructure to manage their own databases. The increasing popularity of cloud databases marks the start of a fresh age of databases. Although cloud databases do not comply with ACID, they can manage huge web-based application tasks which do not need such kinds of assurances. Different Cloud databases are available in the market. But, each one has a distinctive functionality of the query interface, such as API, data model and database. For their better development, these ideas need to be standardized. Once the constraints are overcome, the next decade will be governed by cloud computing and cloud databases.

REFERENCES

[1] Attas D, Batrafi O. Efficient integrity checking technique for securing client data in cloud computing. Int J Comput Sci Issues (IJCSI) 2011;11.

[2] Ma D, Deng RH, Pang H, Zhou J. Authenticating Query Results in Data Publishing. The International Conference on Information and Communications Security, Berlin 2005.

[3] Vodomin, G. and D. Androcec. Problems during Database Migration to the Cloud. in Central European Conference on Information and Intelligent Systems. 2015. Faculty of Organization and Informatics Varazdin.

[4] Abourezq, M. and A. Idrissi, Database-as-a-service for big data: An overview. International Journal of Advanced Computer Science and Applications (IJACSA), 2016. 7(1)

[5] Arora, I. and A. Gupta, Cloud databases: a paradigm shift in databases. International J. of Computer Science Issues, 2012. 9(4): p. 77-83

[6] Alomari, E., A. Barnawi, and S. Sakr. CDPort: a framework of data portability in cloud platforms. in Proceedings of the 16th International Conference on Information Integration and Web-based Applications & Services. 2014. ACM.

[7] DivyakantAgrwa l, Suditpto Das, Amr El Abbadi, " Big Data and Cloud Co mputing: Current State and Future Opportunities", EDBT 2011, Mar 22-24, 2011 ACM.

[8] Ferretti, L., M. Colajanni, and M. Marchetti, Supporting security and consistency for cloud database, in Cyberspace Safety and Security. 2012, Springer. p. 179-193.

[9] Sakhi, I., Database security in the cloud. 2012.

[10] Al Shehri, W., Cloud Database Database as a Service. International Journal of Database Management Systems, 2013. 5(2): p. 1.