

Database Management System as a Cloud Service - A Review

Akshay Shankar

Computer Science and Engineering, RV College of Engineering, Bangalore, India.

Alle Naga Rishikesh Reddy

Computer Science and Engineering, RV College of Engineering, Bangalore, India.

Aryn Barman

Computer Science and Engineering, RV College of Engineering, Bangalore, India.

ABSTRACT

A Cloud Database Management System (CDBMS) is a distributed database system that helps to deliver computing as a service rather than a product. Cloud Database Management System involves resource sharing and allows multiple users to access a single software concurrently. The scope of the services available keeps increasing as technology in this field improves with time. A cloud-based Database Management System is extremely measurable. It handles huge amounts of information, processes, and data that would overload a traditional Database Management System. The important characteristics of a Cloud Database Management System and the challenges faced by Cloud Databases such as multitenancy, privacy, security, and internet speed have been discussed in this paper along with the various advantages of a CDBMS over a traditional Database Management System. The purpose of this review is to establish that moving to the cloud is an excellent option for a company because it reduces the cost involved in network resources, storage, and designates responsibility to the cloud service provider to handle huge amounts of data while providing reliability, availability, and scalability at a low cost.

Keywords: Cloud Service, Migration, Cloud-Native, IaaS, Big Data

I. INTRODUCTION

A Cloud Database Management System (CDBMS) can be defined as a database management system which can be accessed through an internet network by hosting it on a remote server maintained by the service provider. It allows the users to modify and access all the applications or documents uploaded to the cloud through any device connected to the internet from anywhere across the world. There are multiple enterprises that commercially provide a cloud service for DBMS. Some of them include Microsoft Azure, MongoLab, Amazon EC2, Google, etc. Generally, these enterprises provide this service in two ways: users can use a virtual machine to access their databases independently, or users could purchase a cloud database that is retained by the provided cloud company. These cloud services are user-friendly and economical as these services are relatively cheaper and could be expanded or reduced if their requirements change [1].

By simultaneously running several machines virtually at a time, a single server could be converted to multiple servers, and a single data center could be converted into a data center host that would be capable of serving multiple organizations. Hence, cloud service providers could offer services to several users at a time compared to traditional ways, and this would result in providing it at a lower cost.

Cloud servers, in today's world, are always online and ready to use, even if some of them fail. This is achieved by backing up its services on many servers in various locations. Customers can access these cloud services via an app or a web-browser, depending upon the service provider, by connecting to the cloud through various interconnected networks over the internet [2].

In this paper, the highlights will be why enterprises need cloud DBMS, the services which CDBMS offers, and how using CDBMS is more advantageous than traditional DBMS. It also focuses on how a company can migrate to CDBMS from traditional DBMS and also provide a checklist for the migration. There are always certain challenges with all technologies. Those are mentioned including the challenges and disadvantages of CDBMS. There are many advances that are being made to CDBMS along with the integration of CDBMS and Big Data that is also accentuated in this paper.

II. NEED FOR DBMS IN THE CLOUD

Cloud Infrastructure avails wide services for Cloud DBMS services to be developed, these services are not necessarily relational. DBMS engines are not specifically made to take advantage of that cloud on which it runs. This is what led to the change in the name of the service to DBMS as a cloud from DBMS in a cloud.

A. Ease of access

Virtual access of the cloud by the use of interface of the web or API's vendor for users to access the cloud

B. Scalability

Storage spaces of the databases that are on the cloud can be altered during run time according to the organization's requirements. The cost is applied only to the resources used by the organization.

C. Disaster recovery

Remote servers are maintained that keeps the data secure during natural calamities, power failure, and issues faced with types of equipment [3].

III. DBMS AS A CLOUD SERVICE

DBMS has transformed into a totally different type of service with its own advantages and benefits after the integration of Cloud Computing with DBMS. It involves organizing, retrieving data along with making it feasible to scan data from hard drives and networks. DBMS is important to maintain communications between the database and the operating systems, regardless of whether traditional or cloud-based DBMS [4].

Cloud DBMS are exceptionally measurable. Despite the measurability, extremely large processes are still not able to scale up to expected levels. In future advancements of CDBMS, it is expected to overcome this shortcoming. Development of new cloud services and processes and testing is the ideal usage of Cloud Database Management System. Pursuing in capitalizing the disproportion between the old and outdated models of DBMS and their full cloud functionality absence [5].

New strategies like combining one or more components such as data query languages or data structures or elements may be utilized in Cloud DBMS. Society is anxious about the usage of cloud DBMS despite the numerous advantages because of the unsolved security issues. Cloud DBMS runs on multiple hardware and servers which is where the security concerns arise. Accessing databases via numerous applications without being noticed is where the security issues are extremely serious. It is possible that a person who wishes to do harm may access the data in CDBMS without authorization and can cause major problems to the database. The database's integral structure will be put in danger. All suspicious database accesses are recorded and analyzed for dealing with this type of issue and defending against malicious attacks from the individual [6].

IV. ADVANTAGES OF CLOUD DBMS OVER DBMS

There are many important factors to be considered while comparing a Traditional Database vs a Cloud Database. In a cloud database there are two ways to host it, first is Infrastructure as a service (IaaS). In IaaS, there is a user's virtual machine where the cloud provider will host the database and the user will also be able to run it. This is similar to the user doing the same tasks on their on-premises servers. The second one is Platform as a Service (PaaS). In PaaS, the user's application is provided with a database, operating systems, or servers that are absent. For many businesses to migrate from the traditional database, migrating to IaaS is easier and more straightforward as compared to PaaS. Initially, companies will first migrate to IaaS as a first step then move onto PaaS as the second step into the cloud. The company must evaluate their

server size, based on that they can decide between IaaS and PaaS, in this scenario PaaS is more economical than IaaS. PaaS may require more planning but it is more economical [7].

In Cloud DBMS, the database is hosted in the cloud service provided, there are no overhead costs of hardware, network devices, servers, etc in the service provider's data centers. These are the costs that CDBMS avoids as compared to the traditional model but there are other overheads. The user must pay a monthly rental fee to the providers to avail of the service.

The other difference of Cloud DBMS is the storage and the computing size which can be altered to support the needs of the user which gives it flexibility, they can alter it in case the workload changes which cannot be done in the traditional DBMS.

In CDBMS the network and security, management and prevention, underlying infrastructure, patching, and corruption are responsibilities that are to be managed by the providers. Although more maintenance is required in traditional DBMS, the whole infrastructure and the stack are more controlled in traditional DBMS. The enterprises behind the scenes are not allowed the same access levels in CDBMS [8].

There are many advantages and disadvantages to CDBMS which make companies and businesses want to migrate towards it. Backup, one of the outstanding features of CDBMS is that there is a corruption check, backups, and patching are all automatically done by the cloud service provider. These services are available and highly durable.

CDBMS is very cost-efficient and is scalable. Increasing the server performance during busy times and the computing power is reduced during quiet times, these are the measures taken in order to reduce the cost. These are a few of the major benefits of CDBMS, the pliability guarantees service up times with the hardware, costs of setup is lower, and managing infrastructure without any requirements should be added too [9].

V. KEY FEATURES OF MIGRATION TO CLOUD DBMS

A. Usability

Database Migration Service offers a detailed experience which facilitates migration with customized configuration data, setting up various network connectivity modules for a secure connection and other built-in modules for ease of execution. It facilitates success in migration by running validation tests before migrations.

B. Server-less service

By using this service, we would be able to eliminate migration servers which require the provision, management, or constant monitoring. The feature of auto-scaling facilitates continuous data replication at high efficiency for a large set of data. Doing so removes the burden of migrating databases.

C. Secure migration of data

The data during migration is safely protected by several private, guaranteed

connectivity techniques. After the migration of data, encryption is performed on the entire set of data by default, and furthermore, Cloud SQL offers different levels of security to meet the user's requirements.

D. Open Source

Cloud SQL is completely open-source compatible with PostgreSQL and MySQL, so the data migration would be fast and smooth without any disturbances. Additionally, it offers the native replication features of other SQL migrations at no extra charges [10].

VI. CLOUD MIGRATION CHECKLIST

When a business wishes to switch from traditional to CDBMS, these are the recommendations for transitioning from traditional to cloud database management systems. The first step for migration is that the features must be supported, then there must be clarity on the tier that must be used in the cloud database, this is important because the computing power of the database is determined by the tier. The business must ensure that the connectivity and the performance must work with the apps and must be tested. The business must make plans for the downtimes which come with the migration.

High availability, monitoring, alerting and disaster recovery must be configured. The access requirements must be identified, such as who must access the database and what level of access is also required to be identified. The costs are identified by understanding the clear usage of data, writing, and reading along with any accounted growth including costs of ongoing support. Data Confidentiality Compliance feature requires proof. The baseline metrics of performance and other KPIs must be stated confidently. There are many cloud providers which have been discussed in this paper [11].

VII. CHALLENGES FACED IN CLOUD DBMS

Between choosing a DBaaS or managing a cloud based database built by the user, there are a few key factors to be considered before determining whether the selected method will satisfy the user's requirements.

A. Size Limitations

This is one of the main factors that differentiate DBaaS and managed storage options. DBaaS systems offer autoscaling features, but their storage still has some hard limits. Problems arise when the user wants to scale past the level that DBaaS can handle. Databases used at an Organizational level require the functionality to expand at some later point in the future. These limits do not apply to managed storage and are overcome by using layered data storage with greater scaling potential.

B. Protection of Data

Though data is stored in the cloud, it still needs protection. DBaaS is equipped with

inbuilt backups and regular snapshots but these functionalities are inadequate for Organizational level databases.

C. Performance of Storage layer

The speed and efficiency at which a database operates is an important factor. In public cloud databases, the speed and efficiency of the storage layer depends on the capacity provisioned by the user.

D. Cloning of a Database

Cloning of databases is a frequently performed activity for testing and development reasons. This might lead to reduced efficiency and increased expenditure based on the different database type opted by the organization. In DBaaS, cloud based clones are complete duplicates of the database, resulting in twice the total magnitude of the database.

E. Hybrid Cloud/Multicloud Operations

Another challenge is the facilitation of hybrid cloud and multi-cloud features on the different types of cloud databases. Despite the fact that public cloud DBaaS are highly capable, they are incompatible with each other. This makes the transfer of database contents from one cloud to another difficult and requires user intervention. Using some managed storage options, this process is simplified using the data management layer as the data management layer which facilitates data transfer between distinct storage environments [12][13].

VIII. DISADVANTAGES OF CLOUD DATABASE

Even though cloud technologies are highly developed compared to where they were in the last few years, there are several limitations as well. Administrators of the database should prioritize understanding and managing the costs of a cloud database. The initial capital investment of a cloud database might be low but over time and usage, the cost of several databases might spiral, due to the user's requirements, resulting in more cost expenditures. The users of the database will be risking the confidentiality of their data as it will be hosted on the service provider's servers and their security systems which leads to loss of control over the data by the user [14].

Alongside these risks, the users have to constantly keep up with upgrades to the latest patches to fix bugs, performance issues, etc for a seamless experience. These patches are deployed according to the convenience of developers which could be inconvenient for the user. The most general way of accessing a cloud database by a user is through the internet. If their internet connection isn't secure, it might result in a loss of data accessibility [15].

IX. ADVANCEMENTS OF CLOUD DBMS

A. Cloud-Native plus Distributed

Cloud infrastructures are being adopted by more and more organizations. The organizations are moving towards applications that are more business generic and data critical applications. There are numerous researches that indicate that over 75% of the industry will move towards cloud platforms by 2022. Cloud-Native Architectures will be supported by cloud service required organizations. In order to be more dynamic, business organizations are sourcing or developing cloud-native applications. Cloud-Native Databases make users enjoy the high availability and elasticity without heavy infrastructures combined with distributed processing techniques.

B. Self Driving and Intelligent

Due to the growth of data, cloud computing has changed everything, but it is not yet close to reaching real AI standards. In order to be useful in large-scale data, there needs to be the use of deep neural networks. AI has started making a change in databases. In the future, there will be databases that are self-driving and fully automating a roadmap to the database. There is quite a bit of complication which is faced in this process because the requirements change from customer to customer making it difficult to automate the entire process. AI can help in improving latency and scalability, the use of ML algorithms ensures without anomaly that the databases are secure and running [16].

C. Multi-model

In this world where there is constant improvement in technology. As a business seeks to derive newer advances in technologies like the internet of things, AI, and many more it continues to remain a challenge to arrive at insights from heterogeneous data. Open-source standard interfaces are to be compatible in order for databases to compute and decouple storage using cloud-native design.

Switching between analytical engines and multiple computing are being seamlessly connected with open source systems that will be supported by the databases. Various models including the time series and wide column table will be supported with the analysis of semi-structured, unstructured, and structured data, and storage. IoT performance will be improved mainly for the storage of multi-model data, device logs, analysis and computing of device metadata, and data of the device operation.

D. Built-in security

Security has now become one of the primary requirements of any business, interface, product, or application. Securing data in all forms is very important while maintaining high levels of transparency and trust in the cloud era. This will be a prime concern for CDBMS now and in the future. Flexible Key Management, Transparent Data Encryption (TDE), SSL encryption, auditing, and traceability are a few techniques that are in motion and are constantly evolving and improving efficiency as well. The database systems can be made significantly better by combining it with blockchain technology and encrypting processes and data in new hardware without decryption

inside the database [17-19].

X. BIG DATA WITH CLOUD DBMS

Accessing information any time, anywhere is mobility. New business capabilities are being developed from emerging technologies like Big Data, mobility and cloud are creating new platforms by such powerful trends for enterprises. Database systems involve centers of data that provide service, the unique point of access can be seen for all the using the cloud for clients' requests which are coming from around the world.

Big Data comes with its own pros and cons, one of the main issues with Big Data is the sheer size. Greater flexibility, handling storage growth requirements at high speeds and controlling costs. These challenges are addressed when Big Data is combined with the Cloud providing a new paradigm. Big Data along with Cloud helps in offering efficient solutions for the problems above. In order to solve these issues, new techniques, deployments, and development efforts in running workloads that are data-intensive computing, are the steps taken by the DBMS designers to deploy and to develop applications on the internet scale.

Cloud Computing provides several solutions that help in overcoming the shortcomings such as scalability issues by changing the machine storage numbers either by increasing or decreasing so that Big Data can handle the load that is actually placed. On the platform of cloud computing, the relational database is Cloud Database. It provides a CDBMS (Cloud Database Management System).

The use of this technology is to allow the users to store their data in the data centers of the service providers on the cloud. There are many benefits, there is high availability because the data is replicated multiple times (three times) and all the data is stored in the Cloud Database.

There is also scalability, there is more reliability in data storage even during network failures or issues with the system. Big Data is beginning to be impacted greatly by Cloud Computing in the way that it is deployed, used, and managed. It covers the downfalls of Big Data and also provides flexibility, reliability, self-regulated, cost-efficient, and guaranteed quality of service infrastructure. Using CDBMS in order to manage the user's Big Data makes it more scalable, productive, available, and improves the performance [20-22].

XI. CONCLUSION

From the above information, we can conclude that CDBMS can be used for managing and maintaining large quantities of data by acting as a distributed database making it scalable and reliable at a lower cost than a database with dedicated hardware. Also, CDBMS provides more security features as the service provider is the one responsible for keeping it secure and backing it up. So, transitioning towards a cloud based database system would be the better option for an enterprise as it requires lesser investments for networking resources and data storage if planned efficiently. The service provider can

share the resources effectively among multiple customers, as the customers can just invest more if their resource requirements increase over time. This way, the customers purchase only the resources they need at that time, making it more economically viable.

REFERENCES

- [1] Harris, D, “Who builds 'em and what they do”, Cloud Databases 101, November 2012.
- [2] Bloor, R “What is a Cloud Database?”, <http://www.algebraixdata.com/wordpress/wp-content/uploads/2010/01/AlgebraixWP2011v06.pdf>, November 2012.
- [3] Deepak Panthri, Chaman Singh, Abhishek Vats, “Database Management System as a Cloud Service”, International Journal of Innovative Research in Technology, February 2017.
- [4] Curino, C., Madden, S, “Relational Cloud: A Database as Service for the Cloud”, <https://people.csail.mit.edu/nickolai/papers/curino-relcloud-cidr.pdf>, 24th November 2012.
- [5] Finley, K. “7 Cloud-Based Database Services”, Read Write, November 2012.
- [6] Waleed Al Shehri, “Cloud Database Database as a Service”, International Journal of Database Management Systems, April 2013.
- [7] Mykletun, E. and Tsudik, G., “Aggregation Queries in the Database-As-a-Service Model”, Lectures in Computer Science Book, November 2012.
- [8] James, “In these stormy days, could cloud be the answer?”, Well Data, November 2020
- [9] Yaniv Yehuda, “Top 7 Cloud Databases for 2020”, DBmaestro, 13 February 2020.
- [10] Pizzete, L. and Cabot, T.2012. Database as a Service: A Marketplace Assessment, November 2012.
- [11] Hogan. Cloud Computing & Databases: How databases can meet the demands of cloud computing, November 2012.
- [12] B. Vandiver, S. Prasad, P. Rana, E. Zik, A. Saeidi, P. Parimal, S. Pantela, and J. Dave. Eon mode: Bringing the Vertica columnar database to the cloud. In Proc. 2018 Int. Conf. Management of Data, SIGMOD '18, pages 797–809, 2018.
- [13] B. Varghese, O. Akgun, I. Miguel, L. Thai, and A. Barker. Cloud benchmarking for performance. In IEEE 6th Int. Conf. Cloud Computing Technology and Science, pages 535–540, 2014.
- [14] Junjay Tan1, Thanaa Ghanem, Matthew Perron, Xiangyao Yu, Michael Stonebraker, David DeWitt, Marco Serafini, Ashraf Aboulnaga, Tim Kraska,

- “Choosing A Cloud DBMS: Architectures and Tradeoffs”, *Choosing A Cloud DBMS: Architectures and Tradeoffs*, 2019.
- [15] M. Stonebraker, A. Pavlo, R. Taft, and M. L. Brodie. Enterprise database applications and the cloud: A difficult road ahead. In *2014 IEEE Int. Conf. Cloud Eng.*, pages 1–6, 2014.
- [16] Dr. Feifei Li, “Five leading trends of database innovation: Cloud-native, self-driving, and more”, April 2021.
- [17] C. Binnig, D. Kossmann, T. Kraska, and S. Loesing. How is the weather tomorrow?: Towards a benchmark for the cloud. In *Proc. Second Int. Workshop on Testing Database Systems, DBTest’09*, pages 9:1–9:6, 2009.
- [18] B. F. Cooper, A. Silberstein, E. Tam, R. Ramakrishnan, and R. Sears. Benchmarking cloud serving systems with YCSB. In *Proc. 1st ACM Sym. on Cloud Computing, SoCC’10*, pages 143–154, 2010.
- [19] B. Dageville, T. Cruanes, M. Zukowski, V. Antonov, A. Avanes, J. Bock, J. Claybaugh, D. Engovatov, 2181 M. Hentschel, J. Huang, A. W. Lee, A. Motivala, A. Q. Munir, S. Pelley, P. Povinec, G. Rahn, S. Triantafyllis, and P. Unterbrunner. The snowflake elastic data warehouse. In *Proc. 2016 Int. Conf. on Management of Data, SIGMOD ’16*, pages 215–226, 2016.
- [20] Mona Mohamed Nasr, Shima Ouf, “Cloud Computing: The Future of Big Data Management”, April 2015.
- [21] Agrawal, D., Das, S., & El Abbadi, A, “Big data and cloud computing: current state and future opportunities”, *Proceedings of the 14th International Conference on Extending Database Technology. ACM*, March 2011.
- [22] Ahuja, S. P., & Moore, B. (2013). “State of Big Data Analysis in the Cloud. *Network and Communication Technologies*”, *Canadian Center of Science and Education*, 2(1), p62, 22 May, 2013.