IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 5, Issue 11, November 2016

Incorporating Database Management System in Cloud

Abhi Vora[1], Disha Jain[2], Gayatri Mavani[3], Sakshi Shah[4]

Student, Department of Information Technology, K.J. Somaiya, Mumbai, India [1]

Student, Department of Computer Science and Information Technology, VJTI, Mumbai, India [2, 3, 4]

Abstract: Cloud computing is the practice of using a network of remote servers hosted on the internet to store, manage and process data rather than a local server or a personal computer. A database management system is a system software for creating and managing databases. In this paper, we are incorporating Database Management System in Cloud. A Cloud Database Management System is a service provided over the internet which provides means for sharing of software, resources and data between multiple nodes in a network. The motivation for this paper is the growing need for handling massive amounts data which can be easily achieved by integrating Cloud with DBMS. In this paper, we discuss the history of Database Management System and the recent shift of DBMS to Cloud. We have also deliberated on the merits and drawbacks of using Cloud Services to manage data. In the end, we propose architecture of DBMS in Cloud.

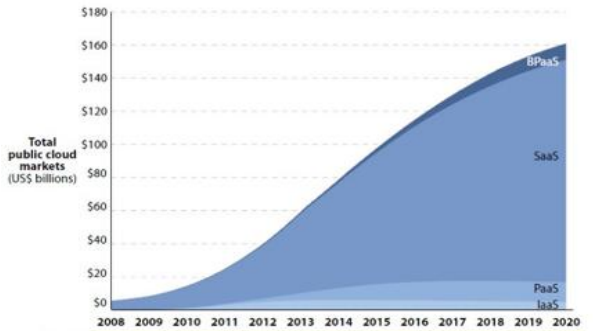
Keywords: DBMS, Database Management System, Cloud, CDBMS

1. INTRODUCTION

Relational Database Management System is a program that lets you create, update and administer a relational database. For many years, this esteemed system, running in corporate companies has handled the bulk of the world's data. The velocity, volume and variety of data being created has been increasing at an exponential rate and hence RDBMS has not been able to keep pace with it. Its clear that the database services is critical to successfully manage the explosion of data. Due to this concern, there was a growing need for a shift of paradigm to a more robust technology which could manage this outbreak of data. The new technology we are talking about is Cloud Database Management System. A Cloud Database Management System is a system that is hosted by a third party service provider on a remote server and accessed over the Internet. A cloud service reduces the cost and complexity of owning and operating computer networks and provides scalability, reliability and efficiency. Cloud Management caters to a plethora of users due to its adaptability and flexibility as a public, private and hybrid cloud.

The database administrator finds it particularly difficult when there is a change in business applications due to the inflexibility of RDBMS. To deal with this problem, the database administrator has to perform complex joins. This challenge can be easily solved with the use of cloud computing as they better fit modern application development due to its agility and on-demand self service. The most important characteristic for a Database Management System is scalability. With this increased scalability in cloud, the DBMS is able to handle and process larger amounts of data which earlier was very difficult the traditional system.

This is an important research area because as we stated above, Cloud DBMS are designed to process thousands and millions of nodes and are able to serve a wide range of data. The benefits of Cloud DBMS have essentially paved the way for this research.



2. BACKGROUND

What is Database Management System?

Database Management System refers to the technology of storing and retrieving user's data with utmost efficiency along with appropriate security measures. The DBMS truly provides an interface between the database and the clients ensuring the integrity and ease of accessibility.

Characteristics of DBMS include:

1. Self-describing: This characteristic of the system contains not only the database itself but also complete depiction of its structure and constraints.

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 5, Issue 11, November 2016

2. Supporting multiple views of data: The DBMS provides access to multiple users based on their authorization level.
3. Data sharing and independence: In the system, data is shared between multiple users over a network and to avoid data inconsistency locking mechanisms are used to prevent two or more users to concurrently make changes to the data.
4. Backup and Recovery Facilities: This involves recovery of the system to the last committed state in case of a server failure during a transaction.
5. Restriction of Unauthorized Access: It implies that security should be provided when multiple users have access to a single database then only some should be granted access to the whole database whereas some should have only restricted access.

There are various types of DBMS, which include Relational, Hierarchical, Network and Object-Oriented. The most widely used is the Relational DBMS, the backbone of many multinational companies.

DBMS has the following defining properties:

1. Atomicity: Transaction must be treated as a single unit, which is indivisible in which either all database operations occur or nothing occurs.
2. Consistency: It is a state in which only valid data is written into the database. Any transaction which violates the consistency rules and the system will be rolled back to consistent state.
3. Isolation: It is a property which allows for concurrent transactions to take place such that one transaction is not affected by the other.
4. Durability: The transactions which have been committed successfully to the database will survive permanently even in case of system failure.

What is cloud computing?

Cloud computing, often referred to as simply 'The cloud', is the delivery of on-demand computing resources-everything from application to data centers -over the internet on a pay-per use basis.

Salient features of cloud are:

1. On-demand self-service: This refers to service provided by vendors that empowers the user to obtain the demanded resources as and when they are required.
2. Broad network access: There is a wide range of services available over the network which can be accessed by a plethora of devices, such as mobile phones, PDA, tablets or laptops.
3. Resource pooling: A multi-tenant model is used to pool several computing resources, like storage, network bandwidth and storage processing, to multiple clients or tenants. These services can be adjusted to suit each client's needs without being apparent to the end user.
4. Rapid Elasticity: This is another term for scalable provisioning which means that the capabilities available for provisioning often appear to be limitless. These can be allocated and re-allocated in any quantity and anytime to meet the user's requirements.
5. Measured Service: The client can utilize this service for various reasons including billing, effective use of resources or overall predictive planning.

Types of cloud are:

Public Cloud is an infrastructure owned and managed by third party vendors selling cloud services which are made available to the general public. The users don't need to purchase hardware, software or supporting infrastructure which reduces the user's expenses. Other benefits include pay-per-use billing, on demand allocation of resources.

Private Cloud is an infrastructure operated solely for a single organization which maybe administered by the organization itself or by a third party. It enables virtualization such as high availability and dynamic resource scheduling to meet the requirements of that organization.

Hybrid Cloud is the integration of private cloud foundation with the use of public cloud services. Private cloud is used for applications which require high level security where as public cloud is used for less sensitive data. The benefits include portability of data, services, more choices for deployment models and data can also be shared between different types of clouds.

3. INCORPORATING DBMS IN CLOUD AND ITS ARCHITECTURE

Database Management System refers to the technology of storing and retrieving user's data with utmost efficiency along with appropriate security measures. But when data is accessed by the client from Cloud and delivered to the user On-Demand it is referred to as Cloud Database also known as DBaaS (Database as a Service). Incorporating the database in the cloud can be an efficacious way to support and cloud-enable business. This can be achieved by abridging the processes needed to avail information through web based connections. Cloud databases can have several assets. These include the ease for bandwidth and storage capacity on per-use standard. It also insures high availability and scalability as and when required.

The notion of managing data has been around since decades such a Relational Database Management Systems of the 1970s. DBMS makes it possible to analyze, query and retrieve and organize data on networks and physical storage devices such as hard drives. Whether it is a traditional or a Cloud database management system, they essentially function as mediators between the Operating Systems and the database systems. Undoubtedly, traditional databases like MySQL, Microsoft Access have been instrumental in handling information in the past but fail to keep up with the current data boom.

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 5, Issue 11, November 2016

In today's world, we are building applications that involve high volume data fields and incorporate predictive analytics. These modern applications do not fit the RDBMS framework. The cost of managing data is five to ten times higher than the initial cost of acquisition. Also sometimes an on site traditional database system typically comes with a large, untenable up- front cost, both in hardware and in software. To ensure cost efficiency there is a growing demand of redistributing DBMS tasks to third party vendors that provide assistance for these services.

On the other hand, Cloud DBMS can easily overcome the challenges faced by traditional databases. Cloud databases can offer tremendous benefits over their traditional equivalents, including increased accessibility, automated scaling, minimal investment and safeguarding of in-house hardware, potentially better performance and fast automated recovery from failures. Vendors who desire a less expensive platform will be initially affected by Cloud DBMSs. But as the maturity in scalability and reliability for Cloud DBMSs increases, the market expenses for applications and rapid development platforms will experience a cost reduction.

This merit of the Cloud DBMS is backed up by lack of need of IT professionals. Usually, a lot of planning and gathering of requirements is necessary for IT projects, but this system is driven to achieve a high speed of set up which leads to rapid deployment of the application. Short duration and short notice projects can achieve faster response, thus reducing overall expenses in the IT departments.

Data management applications are potential candidates for implementation in the cloud. For many start-ups and medium-sized businesses, the pay-per-use model used in cloud computing as well as having a third party worrying about maintenance of hardware is very attractive.

A few of the many companies which are currently using the cloud DBMS are, Amazon Web Services, Microsoft Azure and Google Cloud SQL. Microsoft utilizes an SQL Server technology that provides a traditional database, permitting clients to either access a SQL database on its cloud, or using virtual machines which host SQL server instances. Microsoft also places high importance on hybrid databases that combine data both on a customer's premise and with the Azure cloud through SQL Data Sync.

Despite all of the advantages of cloud-based DBMS, many people still have doubts about them. The most likely reason for this is the variety of security issues that haven't been dealt with yet. The root of these issues is that cloud DBMSs are difficult to monitor because they extend multiple hardware devices, like stacks and servers.

Another example is MongoDB Atlas which is a cloud hosted Database which facilitates users to build applications on their platform.

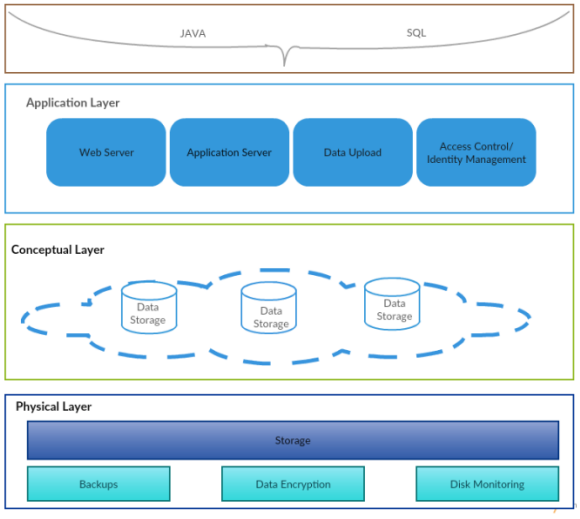
When multiple Virtual Machines are present that might be able to access a database being unnoticed or setting off any warnings or alerts, security becomes a major issue for cloud DBMS. In such a situation, any malevolent person can probably access significant data or seriously harm the integral framework of a database and put the system in peril. However, there is a recommended method to deal with these issues.

A solution that is obvious is the employment of an independent network agent that meticulously monitors and defends all actions referring to database access. The drawback is that this agent may not be able to maintain excessively large and dense volumes of traffic.
It can be argued that the best solution to security problems encountered in a cloud DBMS, is to use continuous auditing of the database. A system must be set up to rigorously record, scrutinize and report all actions relating to database access, especially suspicious ones.

The information regarding these actions is noted and saved in an extremely inaccessible and secure location which sends alerts out to management (or any designated individual) in the case of a breach. The people who are in charge of security will thus have all the necessary information to find out who is the culprit, the specifics of their hardware as well as their location.

Even though a comprehensive cloud DBMS has not been deployed till date, its development has definitely begun. The flood gates will open to a new generation of cloud computing when a thorough solution for all cloud service models emerges.

The cloud databases that are devised can serve petabytes of data as they will be running on a cluster of thousands of nodes. These cloud databases may face the difficulty of reduced querying capacity and less consistent guarantees but they will also have a major advantage. These databases will provide in-built support for elasticity, load balancing and availability.



DOI 10.17148/IJARCCE.2016.51112

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 5, Issue 11, November 2016

Many a time's business analysts do not have technical proficiency and are not restful when it comes to directly interfacing with the low-level database software. This makes data management tools an essential piece of analytical and relational data management.
Since these tools use ODBC and JDBC to interface with the database, SQL queries must be accepted by the software. This makes a unique technology which combines the scalability of Cloud with DBMS capability indispensable. The proposed architecture consists of three layers as described below:

Application Layer: This layer provides an interface to users. The main objective of the service provider is to provide a platform with complete manageability, security and transparency.

Conceptual Layer: This layer provides interoperability and abstraction of conceptual level heterogeneity. The logical structure of the complete database is represented by this layer; it also deals with the internal processing of data.

Physical Layer: The function of this layer is similar to Cache Memory in CPU, where regularly accessed SQL queries are saved for faster access thus reducing the usage of bandwidth on runtime.

4. MERITS AND DRAWBACKS OF DBMS AS A CLOUD SERVICE

The acute drive of any business to lower operational expenses and IT intricacies has fueled a new breed of technology – Cloud Computing Database. Undoubtedly, businesses can draw huge benefits from Cloud DBMS services. However, with the various advantages, in some areas this service has a few shortcomings as well.

The Merits of using DMBS in a Cloud service are: -
1. Storage Virtualization: It is the merging of several physical storage devices over a network into a single centrally managed storage device. This ability provides the operators more independence in how they manage the manage the resources as Virtualization isn't directly visible to the application.
2. Scalability: Traditional DBMS required to built infrastructures to withstand the highest point of usage, but Cloud provides On-Demand Scalability which takes care of this issue efficiently.
3. Reduce administration burden: A Cloud Hosted Database eliminates the unnecessary features that typically consume database administrator's time which could be focused on more essential tasks.
4. Economics: The main focus of businesses is keeping operational expenses to a minimum. Cloud significantly lowers the cost in two ways. First by reducing the cost of hiring IT experts. Second by eliminating the purchase of hardware equipment and software licenses since it uses Multi-Tenant cloud environment.
5. Reliability: Cloud Computing is more dependable and consistent than in-house IT framework since it provides Service Level agreement which guarantees 24x7 availability.

Some of the drawbacks are: -
1. Security: In cloud computing, every unit is potentially vulnerable to attacks from Internet. When a business employs cloud services, it exposes itself to malicious attacks since it discloses business information to service vendors.
2. Downtime: With the increase of cloud computing database technology, the number of clients has increased exponentially which has resulted in services becoming overwhelmed and coming up against technological blackouts.
3. Limited Control: The service vendor imposes limits on what clients can do with their applications. Thee customers can control and manage applications and data operated but not the backend infrastructure itself.

From this it is clear the advantages of Cloud DBMS outweigh the disadvantages. As the technology matures the costs will fall further and security norms will improve.

5. CONCLUSION

As suggested in this paper Cloud Database Management Systems has played an instrumental role in storing and managing large volumes of data which is the centre of most applications in the modern word. It enhances scalability, flexibility, reliability and many such capabilities at a low cost with superior performance compared to traditional Database infrastructures. In this paper we presented the idea of DMBS as a cloud service, its advantages, disadvantages and an architecture for the same.

REFERENCES

[1] Keiko Hashizume, David G Rosado, Eduardo Fernández-Medina and Eduardo B Fernandez,"An analysis of security issues for cloud computing", Hashizume et al. Journal of Internet Services and Applications 2013.
[2] Adrienne Watt, Characteristics and Benefits of database, Database design ,2012.
[3] Abadi D, "Data Management in the Cloud: Limitations and Opportunities", Bulletin of the IEEE Computer Society Technical Committee on Data Engineering, (2009).
[4] Kai Hwang, Distributed and Cloud Computing in Mk Publication, 2012.
[5] Kellogg D, "DBMS in the Cloud: Amazon SimpleDB", http://kellblog.com/2007/12/18/dbms-in-thecloud-amazon-simpledb/
[6] Hsieh M, Chang C, Ho L, Wu J and Liu P, "SQLMR: A Scalable Database Management System for Cloud Computing", In Proceedings of ICPP, (2011), pp. 315-324.
[7] Jurg Van Vliet, Flavia Paganelli, Programming Amazon EC2 in O'Reilly Media, 2011.