Database Management System as a Cloud Service

Yvette E. Gelogo[1] and Sunguk Lee[2]*

[1]*Society of Science and Engineering Research Support, Korea*
vette_mis@yahoo.com

[2]*Research Institute of Industrial Science and Technology*
Pohang, Gyeongbuk, Korea
sunguk@rist.re.kr
**Correspondent Author: Sunguk Lee* (sunguk@rist.re.kr)*

Abstract

A Cloud database management system is a distributed database that delivers computing as a service instead of a product. It is the sharing of resources, software, and information between multiple devices over a network which is mostly the internet. It is expected that this number will grow significantly in the future. As a result, there is a growing interest in outsourcing database management tasks to third parties that can provide these tasks for much lower cost due to the economy of scale just like putting it into the cloud. In this paper, we discuss the recent trend in database management system and the possibilities of making it as one of the services offered in the cloud. We also proposed an architecture of database management system in the cloud.

Keywords: *DBMS, Database Management System, Cloud computing*

1. Introduction

In recent years, database outsourcing has become an important component of cloud computing. Due to the rapid advancements in a network technology, the cost of transmitting a terabyte of data over long distances has decreased significantly in the past decade. In addition, the total cost of data management is five to ten times higher than the initial acquisition cost. As a result, there is a growing interest in outsourcing database management tasks to third parties that can provide these tasks for much lower cost due to the economy of scale. This new outsourcing model has the benefits of reducing the cost for running Database Management System (DBMS) independently [1]. Cloud computing economics leveraging the power of multi-tenancy delivers extremely fast shared storage at a dramatically reduced cost. Virtualization then compounds these advantages by enabling users to scale elastically and to pay only for the resources they use. The cost/performance advantages have decisively shifted in favor of the shared-disk DBMS. It is just a matter of time before the shared-disk DBMS establishes dominance in the cloud.

A Cloud database management system (CDBMS) is a distributed database that delivers computing as a service instead of a product. It is the sharing of resources, software, and information between multiply devices over a network which is mostly the internet. It is expected that this number will grow significantly in the future. An example of this is Software as a Service, or SaaS, which is an application that is delivered through the browser to customers. Cloud applications connect to a database that is being run on the cloud and have varying degrees of efficiency. Some are manually configured, some are preconfigured, and some are native. Native cloud databases are traditionally better equipped and more stable that those that are modified to adapt to the cloud.

2. Background

2.1 Database Management System (DBMS)

A database management system (DBMS) is a software package with computer programs that control the creation, maintenance, and use of a database. It allows organizations to conveniently develop databases for various applications by database administrators (DBAs) and other specialists. A database is an integrated collection of data records, files, and other objects. A DBMS allows different user application programs to concurrently access the same database. DBMSs may use a variety of database models, such as the relational model or object model, to conveniently describe and support applications. It typically supports query languages, which are in fact high-level programming languages, dedicated database languages that considerably simplify writing database application programs.

2.2 Cloud Characteristics

One of the oft-cited advantages of cloud computing is its elasticity in the face of changing conditions. For example, during seasonal or unexpected spikes in demand for a product retailed by an e-commerce company, or during an exponential growth phase for a social networking Website, additional computational resources can be allocated on the fly to handle the increased demand in mere minutes (instead of the many days it can take to procure the space and capital equipment needed to expand the computational resources in-house). Similarly, in this environment, one only pays for what one needs, so increased resources can be obtained to handle spikes in load and then released once the spike has subsided. However, getting additional computational resources is not as simple as a magic upgrade to a bigger, more powerful machine on the fly; rather, the additional resources are typically obtained by allocating additional server instances to a task [3]. Having DBMS in the cloud will give advantage in fast and elastic computing.

3. DBMS as a Cloud Service

Most DBMS or database management systems are simply software packages that users can acquire to create, maintain or use a database. However, since the introduction of cloud computing, DBMS has morphed into an entirely new type of service with its own unique benefits and task specific advantages. For one thing, any type of cloud service model will have to employ a dedicated cloud DBMS in order to truly provide customers with excellent access to data and databases. Traditional DBMS's are simply not set up or equipped to deal with the demands of cloud computing. And of course, if DBMS was deployed as a service as part of a larger package provided, it would likely be much more efficient in its duties and therefore cheaper in the long run.

The concept of the DBMS has been around since the beginning of commercial computing; such as the navigational DBMS of the1960's. Database management systems are one of the oldest integral components of computing, essentially making it possible to scan, retrieve and organize data on hard drives and networks. All DBMS, despite whether traditional or cloud-based, are essentially communicators that function as middlemen between the operating system and the database.

How is a cloud DBMS different a traditional one? For one thing, cloud-based DBMS are extremely scalable. They are able to handle volumes of data and processes that would exhaust a typical DBMS. Despite their scalability however, cloud DBMS are still somewhat lacking in their ability to scale up to extremely large processes; this is expected to be remedied in the coming months and years however. Currently, the use of cloud DBMS's are principally used in the testing and development of new cloud applications and processes. But while a stand-alone DBMS can be used on a cloud infrastructure;

most are not designed to take full advantage of cloud resources. DBMS as a cloud service-type models seek to capitalize on the disparity between antiquated DBMS models and their lack of full cloud functionality.

Cloud DBMS may utilize all of these components or may have devised new strategies that combine one or more elements (like combining data structures and the data query language, for example). Many organizations are exploring the option of utilizing pre-existing modeling languages as a basis for expansion in a cloud model. This strategy ultimately saves on the time spent developing cloud DBMS's as well as enhances their overall effectiveness, since traditional modeling languages are more than adequate for handling data.

Despite the benefits offered by cloud-based DBMS, many people still have apprehensions about them. This is most likely due to the various security issues that have yet to be dealt with. These security issues stem from the fact that cloud DBMS are hard to monitor since they often span across multiple hardware stacks and/or servers. Security becomes a serious issue with cloud DBMS when there's multiple Virtual Machines (which might be accessing databases via any number of applications) that might be able to access a database without being noticed or setting off any alerts. In this type of situation a malicious person could potentially access pertinent data or cause serious harm to the integral structure of a database, putting the entire system in jeopardy.

There is however a proposed method for dealing with these types of incongruence. An obvious solution is the deployment of an autonomous network agent, which rigorously monitor and defends all activities related to database access. The limitation of this method however, is that a network agent may be unable to handle extremely large and dense volumes of activity / traffic.

Arguably, the best solution for dealing with security issues is to employ continuous database auditing. This involves setting up a system that meticulously records, analyze and report on all activities regarding database access, especially suspicious database access. All information regarding these activities is logged and stored in an extremely remote and secure location with alerts being sent out to cloud management (or including any other individuals they might have designated to receive this information) in the event of a breach. This will provide those in charge of security with the information necessary to determine who is responsible, where they are located as well as the specifics of their machine / hardware.

While deployment of a dedicated and thorough cloud DBMS hasn't occurred yet, it is certainly under development. The emergence of a comprehensive solution for all cloud service models regarding database management will open the door to a new era of cloud computing.

Many of these cloud databases are designed to run on a cluster of hundreds to thousands of nodes, and are capable of serving data ranging from hundreds of terabytes to petabytes. Compared with traditional relational database servers, such cloud databases may offer less querying capability and often weaker consistency guarantees, but scale much better by providing built-in support on availability, elasticity, and load balancing.

On the other hand, data management tools are an important part of relational and analytical data management business since business analysts are often not technically advanced and do not feel comfortable interfacing with low-level database software directly. These tools typically interface with the database using ODBC or JDBC, so database software that want to work these products must accept SQL queries. Therefore, a novel technology to combine DBMS capability with Cloud scale scalability is highly desirable.

4. Why DBMS in Cloud?

Database Management Systems as a cloud service are engineered to run as a scalable, elastic service available on a cloud infrastructure. These DBMS are available only as a cloud offering and are not necessarily relational. For example, Microsoft's SQL Azure is fully relational DBMS, while Microsoft's SQL services, Amazon's simpleDB and Google's Big Table are not relational and have different persistence models. Cloud-based DBMS services are provided in a multi-tenancy environment with elastic resources allocation, for use in simple to complex transactions. DBMS as a cloud service excludes those DBMS that will run on the cloud infrastructure, but are not purpose-built as a cloud service. Most of the currently available DBMS engines will run on cloud infrastructure, but are not specifically engineered to take advantage of the cloud. This differentiation is the reason for the change in name from "DBMS in the Cloud" to "DBMS as a cloud Service"; running on cloud infrastructure does not define a DBMS as a cloud service [2].

All currently available cloud DBMS are relatively new. SQL azure, the only fully relational DBMS available, began full production at the beginning of 2012 and still has some size limitations; Microsoft plans to reduce, and eventually lift, these restrictions.

Today, DBMS as a cloud service are used primarily for development and testing of applications- where database sizes are small and issues of security and collocation with multiple users are not concern. One big advantages of cloud DBMS is their elasticity: the more you use, the more you pay; the less you use, the less you pay [2].

Initially, cloud DBMSs will have an impact for vendors desiring a less expensive platform for development. As cloud infrastructure with DBMSs gains maturity especially in scalability, reliability and security, cloud implementations used for short-term projects such as small departmental applications and rapid development platforms will show marked cost reductions compared with implementations within the IT department. This advantages reinforced by the ability to set up a cloud DBMS environment without the use of expensive IT personnel. The speed of setup will be a primary driver to rapid deployment of systems without the usual requirements and planning necessary for IT projects within the IT department. This will also reduce the necessity for IT to respond to short notice and short duration projects, reducing overall costs in IT. Data management applications are potential candidates for deployment in the cloud. This is because an on premises enterprise database system typically comes with a large, sometimes prohibitive up-front cost, both in hardware and in software. For many companies (especially for start-ups and medium-sized businesses), the pay as- you-go cloud computing model, along with having someone else worrying about maintaining the hardware, is very attractive. Due to the ever-increasing need for more analysis over more data in today's corporate world, along with an architectural match in currently available deployment options, we conclude that read-mostly analytical data management applications are better suited for deployment in the cloud than transactional data management applications. We thus outline a research agenda for large scale data analysis in the cloud, showing why currently available systems are not ideally-suited for cloud deployment, and arguing that there is a need for a newly designed DBMS, architected specifically for cloud computing platforms [3].

5. DBMS in Cloud Architecture

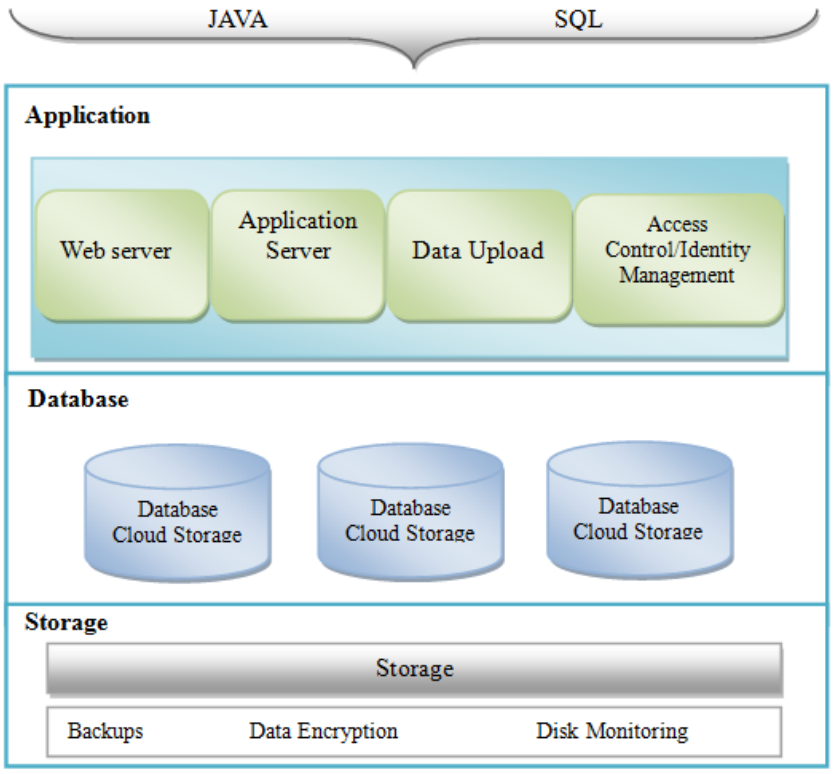

Figure 1. DBMS in the Cloud Architecture

Above is a proposed DBMS in Cloud Architecture, first layer is the storage, followed by databases and the upper layer is application layer. In terms of performance, it provides efficient data access with a better distribution of values for some data. Stores frequently used SQL statements in memory, avoiding the need for time-consuming recompilation at run-time. Produces a detailed report on each step used for data access, allowing you to accurately implement performance enhancements. Data is encrypted when stored or backed up, without any need for programming to encrypt and decrypt.

6. Conclusion

Database Management Systems as a cloud service are engineered to run as a scalable, elastic service available on a cloud infrastructure. Cloud DBMSs will have an impact for vendors desiring a less expensive platform for development. In this paper, we presented the idea of DBMS in the cloud, the possibilities to be offered as one of the services offered by promising capability of cloud computing, that is to be a DBMS as a Service. In this paper we proposed an architecture of DBMS in the cloud.

References

[1] Buyya R, Broberg J and Goscinski A, "Cloud computing Principles and Paradigms", A Jon Wiley & Sons, Inc. Publication, **(2011).**
[2] Feinberg D, "DBMS as a Cloud Service", **(2010)**, Gartner, Inc. and/or its Affiliates.
[3] Abadi D, "Data Management in the Cloud: Limitations and Opportunities", Bulletin of the IEEE Computer Society Technical Committee on Data Engineering, **(2009).**
[4] Kellogg D, "DBMS in the Cloud: Amazon SimpleDB", http://kellblog.com/2007/12/18/dbms-in-the-cloud-amazon-simpledb/.

[5] Gravelle R, "Should You Move Your MySQL Database to the Cloud?", http://www.databasejournal.com/features/mssql/should-you-move-your-mysql-database-to-the-cloud.html.
[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] Hogan M, "Database Virtualization and the Cloud", ScaleDB Inc., **(2009)**.