International Journal of Research (IJR)
e-ISSN: 2348-6848, p- ISSN: 2348-795X Volume 2, Issue 09, September 2015
Available at http://internationaljournalofresearch.org

Efficiency and fee evaluation of an adaptive encryption architecture for cloud databases

[1]G.Bargavi& [2]J.Ravikumar

1M.Tech (CSE), Chadalawada Ramaanamma Engineering College, bhargaviganapaneni@gmail.com

2Assistant professor, Chadalawada Ramaanamma Engineering College, ravikumar509@gmail.com

***Abstract*—**

The cloud database as a service is a novel paradigm that can support several Internet-based applications, but its adoption requires the solution of information confidentiality problems. We propose a novel architecture for adaptive encryption of public cloud databases that offers an interesting alternative to the trade-off between the required data confidentiality level and the flexibility of the cloud database structures at design time. We demonstrate the feasibility and performance of the proposed solution through a software prototype. Moreover, we propose an original cost model that is oriented to the evaluation of cloud database services in plain and encrypted instances and that takes into account the variability of cloud prices and tenant workload during a medium-term period.

Keywords— Cloud Database; Confidentiality; Encryption; Adaptivity; Cost Estimation Model

1. INTRODUCTION

The cloud computing paradigm is successfully converging as the fifth utility, but this positive trend is partially limited by concerns about information confidentiality and unclear costs over a medium-long term. We are interested in the Database as a Service paradigm (DBaaS) that poses several research challenges in terms of security and cost evaluation from a tenant's point of view. Most results concerning encryption for cloud-based services are inapplicable to the database paradigm. Other encryption schemes, which allow the execution of SQL operations over encrypted data, either suffer from performance limits or they require the choice of which encryption scheme must be adopted for each database column and SQL operations. These latter proposals are fine when the set of queries can be statically determined at design time, while in this paper we are interested to other common scenarios where the workload may change after the database design.

In this paper, we propose a novel architecture for adaptive encryption of public cloud databases that offers a proxy-free alternative to the system proposed . The proposed architecture guarantees in an adaptive way the best level of data confidentiality for any database workload, even when the set of SQL queries dynamically changes. The adaptive encryption scheme, which was initially proposed for applications not referring to the cloud, encrypts each plain column into multiple encrypted columns, and each value is encapsulated into different layers of encryption, so that the outer layers guarantee higher confidentiality but support fewer computation capabilities with respect to the inner layers. The outer layers are dynamically adapted at runtime when new SQL operations are added to the workload. Although this adaptive encryption architecture is attractive because it does not require defining at design time which database operations are allowed on each column, it poses novel issues in terms of feasibility in a cloud context, and storage and network costs estimation.

In this paper, we investigate each of these issues and we reach original conclusions in terms

International Journal of Research (IJR)
e-ISSN: 2348-6848, p- ISSN: 2348-795X Volume 2, Issue 09, September 2015
Available at http://internationaljournalofresearch.org

of prototype implementation, performance evaluation, and cost evaluation.
We implement the first proxy-free architecture for adaptive encryption of cloud databases. It does not limit the availability, elasticity and scalability of a plain cloud database, because concurrent clients can issue parallel operations without passing through some centralized component as in alternative architectures. We evaluate the performance through this prototype implementation by assuming the standard TPC-C benchmark as the workload and different network latencies. Thanks to this test bed, we show that most performance overheads of adaptively encrypted cloud databases are masked by network latency values that are quite typical of a cloud scenario. Other performance evaluations carried out in assumed a LAN scenario and no network latency.

2. EXISTING SYSTEM

Improving the confidentiality of information stored in cloud databases represents an important contribution to the adoption of the cloud as the fifth utility because it addresses most user concerns. Our proposal is characterized by two main contributions to the state of the art: architecture and cost model. Although data encryption seems the most intuitive solution for confidentiality, its application to cloud database services is not trivial, because the cloud database must be able to execute SQL operations directly over encrypted data without accessing any decryption key.

Na¨ive solutions encrypt the whole database through some standard encryption algorithms that do not allow any SQL operation directly on the cloud. As a consequence, the tenant has two alternatives for any SQL operation: downloading the entire database, decrypting it, executing the query and, if the operation modifies the databases, encrypting and uploading the new data; decrypting temporarily the cloud database, executing the query, and re-encrypting it. The former solution is affected by huge communication and computation overheads, and costs that would make the cloud database services quite inconvenient; the latter solution does
not guarantee data confidentiality because the cloud provider obtains decryption keys.

The right alternative is to execute SQL operations directly on the cloud database, but avoiding that the provider obtains the decryption key. This proposal is based on data aggregation techniques that associate plaintext metadata to sets of encrypted data to allow data retrieval. However, plaintext metadata may leak sensitive information and data aggregation introduces unnecessary network overheads.

3. PROPOSED SYSTEM

The proposed architecture guarantees in an adaptive way the best level of data confidentiality for any database workload, even when the set of SQL queries dynamically changes. The adaptive encryption scheme, which was initially proposed for applications not referring to the cloud, encrypts each plain column into multiple encrypted columns, and each value is encapsulated into different layers of encryption, so that the outer layers guarantee higher confidentiality but support fewer computation capabilities with respect to the inner layers. we propose the first analytical cost estimation model for evaluating cloud database costs in plain and encrypted instances from a tenant's point of view in a medium-term period. It takes also into account the variability of cloud prices and the possibility that the database workload may change during the evaluation period. This model is instanced with respect to several cloud provider offers and related real prices. As expected, adaptive encryption influences the costs related to storage size and network usage of a database service. However, it is important that a tenant can anticipate the final costs in its period of interest, and can choose the

International Journal of Research (IJR)
e-ISSN: 2348-6848, p- ISSN: 2348-795X Volume 2, Issue 09, September 2015
Available at http://internationaljournalofresearch.org

best compromise between data confidentiality and expenses.

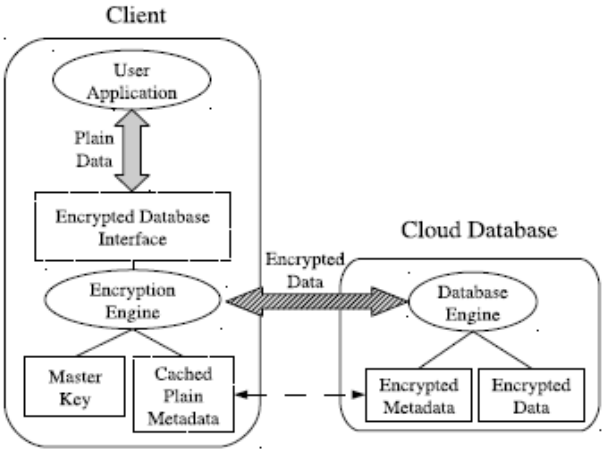

FIG 1: PROPOSED SYSTEM ARCHITECTURE

ADVANTAGES OF PROPOSED SYSTEM

- Data confidentiality is provided through our proposed system.
- The costs related to storage size and network usage of a database service are influenced.

4. IMPLEMENTATION OF PROPOSED SYETM

The framework of our proposed system has the accompanying modules alongside the following prerequisites.

- Adaptive encryption
- Metadata structure
- Encrypted database management
- Cost Estimation of cloud database services
- Cost model
- Cloud pricing models
- Usage Estimation

ADAPTIVE ENCRYPTION

The proposed system supports adaptive encryption methods for public cloud database service, where distributed and concurrent clients can issue direct SQL operations. By avoiding an architecture based on one [or] multiple intermediate servers between the clients and the cloud database, the proposed solution guarantees the same level of scalability and availability of the cloud service. Figure 1 shows a scheme of the proposed architecture where each client executes an encryption engine that manages encryption operations. This software module is accessed by external user applications through the encrypted database interface. The proposed architecture manages five types of information.

• Plain data is the tenant information;
• Encrypted data is stored in the cloud database;
• Plain metadata represent the additional information that is necessary to execute SQL operations on encrypted data;
• Encrypted metadata is the encrypted version of the metadata that are stored in the cloud database;
• master key is the encryption key of the encrypted metadata that is distributed to legitimate clients.

METADATA STRUCTURE

Metadata include all information that allows a legitimate client knowing the master key to execute SQL operations over an encrypted database. They are organized and
stored at a table-level granularity to reduce communication overhead for retrieval, and to improve management of concurrent SQL operations. We define all metadata in formation associated to a table as *table metadata*. Let us describe the structure of a table metadata .Table metadata includes the correspondence between the *plain table name* and the *encrypted table name* because each encrypted table name is randomly generated. Moreover, for each column of the original plain tableit also includes a *column metadata* parameter containing the name and the data type of the corresponding plain column (e.g., integer, string, timestamp). Each column metadata is associated to one or more

International Journal of Research (IJR)
e-ISSN: 2348-6848, p- ISSN: 2348-795X Volume 2, Issue 09, September 2015
Available at http://internationaljournalofresearch.org

onion metadata, as many as the number of onions related to the column.

ENCRYPTED DATABASE MANAGEMENT

The database administrator generates a *master key*, and uses it to initialize the architecture metadata. The master key is then distributed to legitimate clients. Each table creation requires the insertion of a new row in the metadata table. For each table creation, the administrator adds a column by specifying the column *name*, *data type* and *confidentiality parameters*. These last are the most important for this paper because they include the *set of onions* to be associated with the column, the *starting layer* (denoting the actual layer at creation time) and the *field confidentiality* of each onion. If the administrator does not specify the confidentiality parameters of a column, then they are automatically chosen by the client with respect to a tenant's policy. Typically, the default policy assumes that the starting layer of each onion is set to its strongest encryption algorithm.

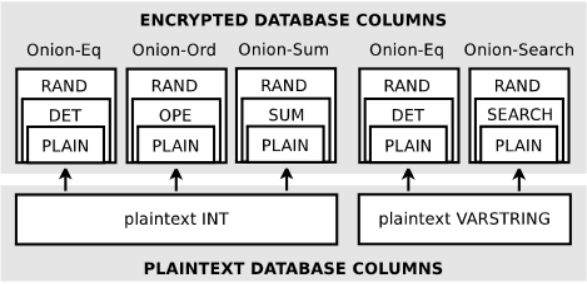

Fig. 2: Example of onion structures

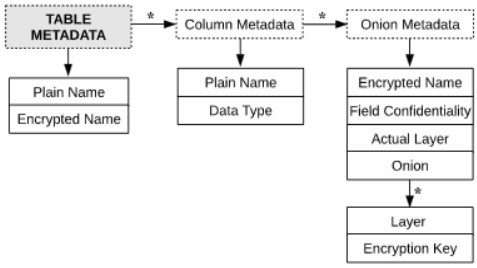

Fig 2: Metadata Structure

COST ESTIMATION OF CLOUD DATABASE SERVICES

A tenant that is interested in estimating the cost of porting its database to a cloud platform. This porting is a strategic decision that must evaluate confidentiality issues and the related costs over a medium-long term. For these reasons, we propose a model that includes the overhead of encryption schemes and variability of database workload and cloud prices. The proposed model is general enough to be applied to the most popular cloud database services, such as *Amazon Relational Database Service.*

COST MODEL

The cost of a cloud database service can be estimated as a function of three main parameters:
Cost = f (Time, Pricing ,Usage) where:

- *Time*: identifies the time interval T for which the tenant requires the service.
- *Pricing*: refers to the prices of the cloud provider for subscription and resource usage; they typically tend to diminish during T.
- *Usage*: denotes the total amount of resources used by the tenant; it typically increases during T .In order to detail the *pricing* attribute, it is important to specify that cloud providers adopt two subscription.
- *Policies:* the *on-demand* policy allows a tenant to payper-use and to withdraw its subscription anytime; the *reservation* policy requires the tenant to commit in advance for a *reservation period*. Hence, we distinguish between *billing costs* depending on resource usage and *reservation costs* denoting additional fees for commitment in exchange for lower pay-per-use prices. Billing costs are billed periodically to the tenant every *billing period.*

International Journal of Research (IJR)
e-ISSN: 2348-6848, p- ISSN: 2348-795X Volume 2, Issue 09, September 2015
Available at http://internationaljournalofresearch.org

CLOUD PRICING MODELS

Popular cloud database providers adopt two different billing functions, that we call *linear* L and *tiered* T . Let us consider a generic resource x, we define as xb its usage at the *b-th* billing period and px b its price. If the billing function is tiered, the cloud provider uses different prices for different ranges of resource usage. Let
us define Z as the number of tiers, and [ˆx1, . . . , ˆxZ−1] as the set of thresholds that define all the tiers. The uptime and the storage billing functions of *Amazon RDS* are linear, while the network usage is a tiered billing function. On the other hand, the uptime billing functions of *Azure SQL* is linear, while the
storage and network billing functions are tiered.

USAGE ESTIMATION

The uptime is easily measurable, it is more difficult to estimate accurately the usage of storage and network , since they depend on the database structure, the workload and the use of encryption. We now propose a methodology for the estimation of storage and network usage due to encryption. For clarity, we define sp, se, sa as the storage usage in the plaintext, encrypted, and adaptively encrypted databases for one billing period. Similarly, np, ne, na represent network usage of the three configurations. We assume that the tenant knows the database structure and the query workload and we assume that each column a A stores ra values. By denoting as vp a the average storage size of each plaintext value stored in column a, we estimate the storage of the plaintext database.

5. CONCLUSION

There are two main tenant concerns that may prevent the adoption of the cloud as the fifth utility: data confidentiality and costs. This paper addresses both issues in the case of cloud database services. These applications have not yet received adequate attention by the academic literature, but they are of utmost importance if we consider that almost all important services are based on one or multiple databases. We address the data confidentiality concerns by proposing a novel cloud database architecture that uses adaptive encryption techniques with no intermediate servers. This scheme provides tenants with the best level of confidentiality for any database workload that is likely to change in a medium-term period. We investigate the feasibility and performance of the proposed architecture through a large set of experiments based on a software prototype subject to the TPC-C standard benchmark.

REFERENCES

[1] R. Buyya, C. S. Yeo, S. Venugopal, J. Broberg, and I. Brand ic, “Cloud computing and emerging it platforms: Vision, hype, and reality for delivering computing as the 5th utility,” Future Generation Computer Systems, vol. 25, no. 6, pp. 599–616, 2009.

[2] T. Mather, S. Kumaraswamy, and S. Latif, Cloud security and privacy: an enterprise perspective on risks and compliance . O’Reilly Media, Incorporated, 2009.

[3] H.-L. Truong and S. Dustdar, “Composable cost estimation and monitoring for computational applications in cloud computing environments,” Procedia Computer Science, vol. 1, no. 1, pp. 2175 – 2184, 2010, iCCS 2010.

[4] E. Deelman, G. Singh, M. Livny, B.erriman, and J. Good, “The cost of doing science on the cloud: the montage example,” in Proc.2008 ACM/IEEE Conf. Supercomputing, ser. SC ’08. Piscataway, NJ, USA: IEEE Press, 2008, pp. 50:1–50:12.

[5] H. Hacig¨um¨us¸, B. Iyer, and S. Mehrotra, “Providing database as a service,” in Proc. 18th IEEE Int’l Conf. Data Engineering , Feb. 2002.

[6] G. Wang, Q. Liu, and J. Wu, "Hierarchical attribute-base d encryption for fine-grained access control in cloud storage services," in Proc. 17th ACM Conf. Computer and communications security . ACM, 2010, pp. 735–737.

[7] Google, "Google Cloud Platform Storage with server-side encryption,"http://googlecloudplatform.blogspot.it/ 2013/08/google-cloud-storage-now-provides.html, Mar. 2014.

[8] H. Hacig¨um¨us¸, B. Iyer, C. Li, and S. Mehrotra, "Executing sql over encrypted data in the database-service-provider model," in Proc. ACM SIGMOD Int'l Conf. Management of data , June 2002.

[9] L. Ferretti, M. Colajanni, and M. Marchetti, "Distributed, concurrent, and independent access to encrypted cloud databases,"IEEE Trans. Parallel and Distributed Systems, vol. 25, no. 2, Feb. 2014.

[10] R. A. Popa, C. M. S. Redfield, N. Zeldovich, and H. Balakrishnan, "CryptDB: protecting confidentiality with encrypted query processing," in Proc. 23rd ACM Symp. Operating Systems Principles,Oct. 2011.

[11] C. Gentry, "Fully homomorphic encryption using ideal lattices," in Proc. 41st ACM Symp. Theory of computing, May 2009.

[12] A. Boldyreva, N. Chenette, and A. O'Neill, "Order-preserving encryption revisited: Improved security analysis and alternative solutions," in Proc. Advances in Cryptology – CRYPTO 2011. Springer, Aug. 2011.