

Measuring Trust and Selecting Cloud Database Services

W.P. Eureka Priyadarshani¹, Gihan N. Wikramanayake² and E.M. Piyal Ekanayake³

¹Information Communication Technology Center, Wayamba University of Sri Lanka Kuliyapitiya, Sri Lanka e.priyadarshani@deakin.edu.au

> ²University of Colombo School of Computing Colombo, Sri Lanka gnw@ucsc.cmb.ac.lk

³Department of Mathematics, Faculty of Applied Sciences, Wayamba University of Sri Lanka Kuliyapitiya, Sri Lanka piyalekanayake@gmail.com

Abstract

In a cloud environment, third party controlled asset sharing cloud databases is the motivation for search of trust. Lack of trust appears to be a probable major cause for fear when outsourcing databases. Research literature reviews indicate that the architecture of cloud databases has a potential to mitigate user distrust. In this paper, we focus on measuring trust in cloud database services based on user selected relative or direct factors. Our study proposes a formal trust mechanism in cloud databases where the user can select his most trusted Cloud Service Provider.

Keywords: Cloud Computing, Database as a Service (DaaS), Cloud Trust

1. Introduction

One of the basic issues that a cloud customer faces is maintaining control over data, particularly if he has to leave or change the hosted cloud service provider (CSP). In such situations, most of the time, the control lies with the cloud service provider and hence the data are subject to loss of trust. In this regard, questions have been raised about the protection of data in terms of tracking customer information, cross border transfer of personal data, data theft (personal and confidential data) and data misuse especially for advertising purposes etc. In this context, the outsourced data can be observed in three aspects namely User Control [1], Transparency [2] and Trust [3]. User control accords authority to the user to decide the storage location, CSP and the ability to move easily from one CSP to another. However, with Transparency, the user knows where data are really stored, computed and which privacy legislations are applicable for the identified cross border transfer of data. Comparatively, the trust in a cloud database can be considered as a combination of factors such as cloud security (security from data misuse, hacker attacks and data damage), data recovery due to a failure on the part of the CSP, confidential computation and the certainty of the CSP.

Moreover, many countries are of the view that the only system they trust is the one operated within their own authority [4]. However, a unified data system with more people accessing it and more diverse types of data coming through more applications can actually make it harder to appropriately limit access and detect misuse.

The certainty of the cloud service provider is also an important factor in respect of building trust in cloud data storages. The cloud database model [5] ensures certainty of data by keeping backups with the user, disregarding the certainty of the CSP. Multi-cloud architectures ([6]; [7]) provide better answers for uncertainty of the CSP as they maintain duplicate copies with other CSPs. The multi-sharing method is effectively linked with multi-cloud databases. Considering Johnston's argument, "A cloud service provider holds data in a purposefully designed fragmentation across servers, it will help to protect information from misuse"[8], various solutions have been introduced by the researchers for cloud data integrity ([9]; [10]). However, the research material to measure the trust based on integrity solutions was not available.

In this study, existing cloud database architectures are considered along with measuring cloud database trust according to the necessities of various users. Finally, a cloud trust measuring mechanism is introduced to select the best cloud service provider based on the user-requested trust factors.

A cloud database architecture that supports measuring trust is explained briefly in Section 2. In Section 3, we define the cloud database trust and describe how it could be achieved with data integrity



solutions. A cloud trust measuring mechanism for the user to select the best cloud DaaS according to their requirements is explained in Section 4.

2. Cloud Database Architecture

In most situations, multi tenant cloud database architecture can be either shared database shared schema architecture or shared database separate schema architecture [11]. For example in *salesforce.com*, multiple tenants use this shared database shared schema architecture and customers (tenants) are much concerned about their data as they are in shared schemas. The structure of many cloud databases is based on a shared database architecture which can moderate some of the distrust points. As Molnar and Schechter point out, the shared database architecture puts cloud users at risk from other cloud users [12].

According to the layered model introduced by Grossman [13], most of the existing cloud architectures run as a collection of services. In these services there are four layers, namely Storage Cloud (provides storage services), Data Cloud (provides data management services on records, columns or objects), Compute Cloud (provides computational services) and Application Layer. The layers used by Grossman over cloud computing give a positive link to the researchers to think about trust in cloud databases.

3. Trust and Evidence for Trust in Cloud Databases

3.1 What is Trust and Cloud Trust?

Trust is an emotional form of thoughts and it is not clear whether integrity, good will and sincerity can be segregated since they are related. Many researchers have explained trust as an act of faith, confidence and reliance on something that is expected to behave or deliver as promised ([14]; [15]). It is obvious that these definitions cannot be applied directly to cloud trust.

In today's cloud society, all services are accessed remotely over the internet. Therefore, definitions related to traditional face-to-face or human-to-human interactions cannot be directly applied to cloud trust. Most researches in cloud computing have considered trust as a social phenomenon based on social science definitions. "Trust is a mental state comprising: (1) expectancy the trustor expects a specific behavior from the trustee (such as providing valid information or effectively performing cooperative actions); (2) belief - the trustor believes that the expected behavior occurs, based on the evidence of the trustee's competence, integrity, and goodwill; (3) willingness to take risk -the trustor is willing to take risk for that belief" [16].

The reputation of a company has a great effect on trust. For instance, In [16] Huang and Nicol mention that the trust and reputation of a company are corelated. According to their method, the trust level of the trustee is measured by estimating the reputation of the company. These measuring mechanisms are not yet completed mathematically and the comparison of trust across entities is not addressed in their paper. Cloud computing capabilities and the intentions of the CSP to create the necessity of trust in cloud computing are explained by Khan and Malluhi [17], and it is obvious that the intentions of the specific CSP have to be compared against other CSPs.

Previous studies have suggested that cloud trust is a social phenomenon [18]. However, cloud database capabilities can be detectable and can be measured by values. As explained in previous studies on trust, relative and absolute measures can be identified in the same way in cloud database trust. Khan and Malluhi further argue that trust can be established by improving transparency, control and security assurances, which imply that it is not merely a social phenomenon [17]. Such expositions direct our cloud database trust measuring mechanism into two ways (relative and direct) and so we define the cloud database trust as follows:

Cloud database trust is empirical and extrinsic evidence based positive expectancy, a trustor expects from the trustee. In the cloud environment, the trustor is the user and the trustee is a cloud service provider or a cloud database. The positive empirical evidence based expectancy obtained from the trustee is called direct trust and the positive extrinsic evidence based expectancy obtained from the trustee is called relative trust.

3.2 Cloud Data Integrity Solutions

Recently, researchers have shown an increased interest in finding solutions for cloud data integrity. On the other hand, and perhaps more importantly, there appears to be evidence of trust factors. In the literature review, distinct solutions for cloud security,



data recovery due to a failure of the CSP and confidential computation have been identified.

Data Recovery: Damaged data can be recovered easily if the same share is stored in another location. In cloud databases, the method called multi-sharing [6] gives adequate answers to this problem. Bowers demonstrated a method called HAIL (High Availability and Integrity Layer) to manage file redundancy across cloud storage providers [9]. It detects and resets the faulty server with the correct share with the help of the cross-server redundancy built in the encoded file. This method gives a better answer for data recovery of static files in a failure of the share of a third party.

Cloud Data Storage Security: Cloud data security spreads across a large area of user requirements. Cloud users are basically willing to outsource destruction and de-identification data as they are invisible to a third party [19] and then it can be easily secured from hacker attacks, data modification and data misuse. In 2011, Popa et al. demonstrated a cloud security storage system called CloudProof which helps customers detect violations of integrity, write-serializability and freshness. Further, it proves the occurrence of these violations to a third party. CloudProof can be built on top of conventional cloud storage services like Amazon S3 or Azure Blob Storage [20]. The Data Coloring and Software Watermarking Technique introduced by Hwang and Li in 2010 is a more secure solution for relational databases and virtual storages. It guarantees that data damage, stealing, altering and deleting cannot be done. A confidential storage has been introduced by Jaatun [21].

Confidential Computation: In a situation where a complete database (database instance) is outsourced, secured computation is also a significant factor from the user's perspective. Database instance includes RDBMS (Relational Database Management Systems) software, table structure, stored procedures and other functionalities. Santos, Gummadi & Rodrigues ensure confidentiality and integrity of computations that are outsourced to IaaS services. According to their proposed trusted cloud computing platform privileged (TCCP), the cloud provider's administrator cannot inspect or tamper with its content and so it allows a customer to reliably and remotely determine whether the service backend is running a trusted TCCP implementation [10].

4. Approach

4.1 System of Trust Measuring Mechanism

First, a general system is introduced which maintains records on known cloud service providers and data integrity solutions they provide with their DaaS. Particularly, it keeps a database of evidence of trust in each and every cloud service that a provider provides with his Database as a Service. This system can be owned by a globally accredited association.

Steps for User,

- First, the user must decide what type of trust he expects from the DaaS and CSP.
- According to his expected trust, he has to identify and select the trust factors (explained in section 4.2) from the system.
- Two types of trust factors (direct & relative), explained in section 4.2, have to be selected separately by the user.

In the meantime, the system will measure the trust level for each and every DaaS and CSP based on the user-selected trust factors. Finally the system reveals the most suitable CSPs according to the userrequested trust factors. Then the user is able to measure the most suitable CSPs for his desired requirements. However, the user is not transparent to all the details the cloud service provides and what the system maintains. This preserves the privacy of the CSP while providing the user with a trusted DaaS.

4.2 Trust Factors

In our study trust factors are considered in two main ways. Firstly, as a Direct Trust Factor (D) which is valued from evidence a user can be identified directly. Second type of trust factor is called Relative Trust Factor (R) which is valued from past user experiences, CSP ratings and indexes. Direct Trust (DT) and Relative Trust (RT) are measured using the values of Direct Trust Factors and Relative Trust Factors respectively.

Direct Trust Factor (D): According to our definition of cloud trust, positive empirical evidence based expectancy can be expected at different levels of the database. For instance, it can be at storage level, computation level or management level of the database. Therefore, the direct trust factor is positive empirical evidence based expectancy on a particular level (or part of) of the database, expected by a trustor. The trust value of a trust factor is decided by its variables.

Relative Trust Factor (R): Relative trust factor is a positive extrinsic evidence based expectation of the



trustor on his trustee. It can be a belief, behavior, agreement or a law expected from the trustee. The factors which help to make them positive are called sub factors of relative trust factors.

Relative Trust (RT): Cloud trust is comparable if it is considered as a social phenomenon. The reputation based trust explained by Huang [18] can help in the cloud selection. In a similar way, relative trust is obtained by comparing relative trust factors with other cloud service providers. The number of relative trust factors considered by the user for measuring trust can vary with their expected level of trust and necessity.

Lemma 1: The relative trust measurement on the trustee is obtained by taking the average of the sum of relative trust factors considered by the user. If the number of trust factors equals τ then the relative trust *RT* is equal $(R_1 + R_2 + \dots + R_\tau / \tau)$ to where $R_1, R_2, ..., R_\tau$ are trust factors that may be comparably measured with other cloud service providers. If R_1 consists of γ number of sub factors such as $R_{11}, R_{12}, \dots, R_{1\gamma}$ and if $v_{111}, v_{112}, \dots, v_{11\alpha}$ are variables of R_{11} and $v_{121}, v_{122}, \dots, v_{12\beta}$ are variables of R_{12} etc. and considered for Likert Scale calculation then

$$|R_{1}| = \frac{(|R_{11}| + |R_{12}| + \dots + |R_{1\gamma}|)}{|R_{1}|} + \frac{|R_{11}| + \dots + |R_{1\gamma}|}{|R_{1}|} + \frac{(v_{121} + v_{122} + \dots + v_{12\beta})}{|R_{\beta}|} + \dots \Big) \Big/_{\gamma}$$
(1)

,where α , β , ... are the numbers of variables of each and every sub factor considered in the Likert Scale calculation.

Example-1: Assume that professional behavior (*Pb*) is considered as a relative factor similar to the source of trust in evidence based trust explained by Huang & Nicol [18]. According to their analysis, the competency of the CSP (capability), its integrity (consistency in performance and principles), and its goodwill (motivation or intention) are measured. Let's consider integrity as a factor having provable evidence which is taken into account in the next section.

Therefore, $Pb = |R_1|$ is derived from the sub factors named competency and goodwill of the CSP.

- Competency Measures $(|R_{11}| = |\text{cmp}|)$.
 - ✓ $v_{111} = cmp_1$: How best a CSP can respond to the client's requirement in the cloud services (SaaS, PaaS, IaaS). It reflects the CSP's scope and complexity of services.
 - ✓ $v_{112} = cmp_2$: What levels of cost, quality, robustness and flexibility are the CSP able to meet?
- Goodwill Measures $(|R_{12}| = |gdw|)$.
 - ✓ $v_{121} = gdw_1$: How long the CSP exists in the cloud market.
 - $\checkmark v_{122} = gdw_2$: Existing client ratio compared to other CSPs.

Assume that cmp_1 , cmp_2 , gdw_1 , gdw_2 are varied in 1-5 Likert scale.

Then,
$$Pb = \frac{\left(\frac{(cmp_1+cmp_2)}{5\alpha} + \frac{(gdw_1+gdw_2)}{5\beta}\right)}{\gamma}$$

where $cmp, gdw \ll 1$, $\theta = 5$ and two variables ($\alpha = 2$) are considered under competency and two variables ($\beta = 2$) are considered under goodwill. The total number of sub factors is also two ($\gamma = 2$).

Direct Trust (DT): In this case, direct trust is not compared with others. If the trustee has empirical evidence for cloud trust then it is called direct trust. What is surprising is that there are theoretically proven solutions for trust factors such as loss of data (e.g. data theft, data misuse, data damage), uncertainty of backup and restoration (i.e. non transparent data storage, inability of owners to access data), uncertainty of confidential computation (i.e. CSP can inspect computation, CSP interference with computation, inability of remote users to inspect computation, remote users cannot interfere with computation) ([22]; [20]; [23]). However, the researchers working on measuring trust are less concerned by those findings. In our study, direct trust is measured on the basis of availability of solutions for direct trust factors in cloud databases. The number of direct trust factors considered for measurement can vary according to the requirements of the user. That means, users can be concerned about storage or computation or management of the database or all of them.



For example, according to the requirement of the user, a selection can be made only for storage or else only for computation or all of them. In other words, cloud data integrity can be measured using these direct values.

Lemma 2: The direct trust measurement on the cloud database (*DB*) is taken from the sum of available direct trust values. If there are σ number of direct trust factors $D_1, D_2, ..., D_{\sigma}$ and if D_1 has k number of variables ($v_{11}, v_{12}, ..., v_{1k}$) then for all k, trust value of $v_{1k} \in \{0,1\}$.

Similarly, the direct trust $DT \text{ on } D_{\sigma} = |D_{\sigma}| = \frac{\sum_{k=1}^{n} v_{\tau k}}{n}$ where the trust value $|D_{\sigma}| \rightarrow [0, 1]$ and n

is the number of variables in the considered factor D_{σ} . The level of trust in D_{σ} returns a real number between 0 &1. Therefore total

$$DT = \frac{|D_1| + |D_2| + \dots + |D_{\sigma}|}{\sigma}$$
$$= \left(\sum_{\sigma=1}^{\varphi} \left(\frac{(\sum_{k=1}^n v_{\sigma k})}{n}\right)\right) / \sigma$$
(2)

where , $\pmb{\varphi}$ is the number of trust factors considered for measuring direct trust.

Example-2: Assume that, confidential computation is a direct trust factor which exists within the selected CSP on database outsourcing. Then the function gives 1 for each variable that provides confidentiality and otherwise 0. Assume that, a particular user is concerned with direct trust factors such as data loss guarantee and confidential computation. Then the number of selected direct trust factors (σ) is two.

5. Total Trust Measures on Cloud Database

Trust is based on two directions in the cloud database as a service (DaaS). They are the trust on database service provider and the trust on the outsourced database instance.

The generalisability of equations (1) & (2) is precise because more relative measures and direct measures can be considered for high accuracy and also with various requirements of the user. It is more important to note that relative and direct trust factors are selected by the user. For example a particular DB user may want to consider the trust on his data storage and some other user may be interested in trust on his database running platform. In that case, a fewer number of direct trust factors can be selected. The other most important advantage in this system is that the user has the authority to assign a weight (λ) for the total relative and direct trust values to obtain the final trust. If the user has more sensitive data and is more concerned with direct trust, he could assign a higher weight to direct trust. Then the total trust is calculated as follows:

$$T = (1 - \lambda) RT + \lambda DT$$
, where $0 < \lambda < 1$ (3)

The other most interesting thing in this mechanism is that the cloud service providers who use this analysis model do not know the exact variables being evaluated by the system. That means, the variables used by the user for selecting the best trustee according to his requirements are different from the variables used by the model to analyze the trust level of the service provider.

6. Sample Case Study

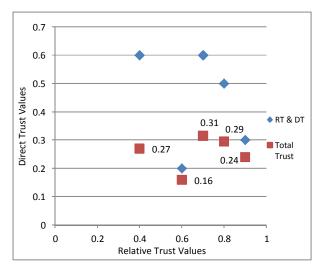
Considering the previous Example 1 and 2, assume that a user X gets the following values for his factors. According to Example-1, there is only one relative factor and according to Example-2 user X has selected two direct trust factors and λ is assigned to 0.7, then the total trust is calculated as shown in the following Table 1.

Cloud Service Provider	Relative Trust Value (RT)	Direct Trust Value (DT)	Total Trust (T) T=(1-λ) RT+λ DT
CS-A	0.6	0.2	0.16
CS-B	0.9	0.3	0.24
CS-C	0.4	0.6	0.27
CS-D	0.8	0.5	0.29
CS-E	0.7	0.6	0.31

Table 1: Top 5 CSP's Relative and Direct Trust values based on user X's selection factors

Graph 1 represents relative trust values against direct trust values and the total trust points. According to the graph, CS-E gets the highest total trust even though CS-C and CS-E have similar direct trust values. User X finally achieves a clear result based on his selected trust factors and the weights he assigned to the RT and DT. According to this example, user X will select CS-E for outsourcing his database.





Graph 1: Total Trust points based on RT and DT values

7. Conclusions

A number of research papers have dealt with trust in data. So far, however, there has been little discussion about measuring trust in cloud databases. With multicloud database architecture, building trust is an important factor of trust as various users share common databases or schema or both in the cloud environment. The trust mechanism is introduced taking into consideration a number of problems which cause a negative impact on the user's trust in the database, including data loss, data theft, and data misuse etc. We argue in this paper for the need of a trust-building mechanism in cloud databases and we introduce a mechanism to measure trust in the cloud database as well as in the CSP.

The theoretical implication of this research is that a standard equation for measuring trust will add to a growing body of literature on cloud trust. The methods used for measuring trust may be applied to other cloud services elsewhere in the cloud world.

It is suggested that the collection of these trust factors is investigated in future studies with a survey made on CSPs and cloud users. We suggest that before this trust mechanism is implemented, a study similar to this case study should be carried out on real world cloud users and CSPs.

REFERENCES

 W.P.E. Priyadarshani, G.N. Wikramanayake, and L.M. Batten, "Enhancement of user level controls in cloud databases", *30th National Information Technology Conference* (NITC), Computer Society of Sri Lanka, Colombo, Sri Lanka, 2012, pp. 46–53.

- [2] A.A. Friedman, and D.M. West, "Privacy and security in cloud computing", *Center for Technology Innovation at Brookings*, 2010.
- [3] S.S.E. Thorpe, "Modeling a trust cloud context", 3rd workshop on Ph.D. students in information and knowledge management, *ACM New York*, USA, 2010, pp. 95–98.
- M.S. Blumenthal, "Is Security Lost in the Clouds?", *Communications and Strategies*, 2011, 81(1), pp. 69– 86.
- [5] V.A. Talasila and Peruri, A Secure Privacy Preserving Storage Architecture of Cloud Database, *International Journal of Computer Science Engineering and Technology* (IJCSET), 2011, 1(9), pp. 587-595.
- [6] M.A. Alzain, and E. Pardede, "Using Multi Shares for Ensuring Privacy in Database-as-a-Service", 44th Hawaii International Conference on System Sciences (HICSS),2011, pp. 1–9.
- [7] A. Bessani, M. Correia, B. Quaresma, F. Andr'e, and P. Sousa, "DEPSKY: Dependable and Secure Storage in a Cloud-of-Clouds", 6th Conference on Computer Systems, 2011, 31–46.
- [8] S. Johnston, Cloud Computing and Privacy, Retrieved from www.circleid.com/posts/89163, 2008.
- [9] K.D. Bowers, A. Juels, and A. Oprea, "HAIL: a high availability and integrity layer for cloud storage", 16th ACM Conference on Computer and Communications Security, ACM New York, USA, 2009, pp. 187–198.
- [10] N. Santos, K.P. Gummadi, and R. Rodrigues, "Towards trusted cloud computing", Conference on Hot topics in cloud computing", Article 3, USENIX Association Berkeley, CA, USA, 2009.
- [11] J. Namjoshi, and A. Gupte, "Service oriented architecture for cloud based travel reservation software as a service", In Cloud Computing, CLOUD'09. *IEEE International Conference on IEEE*, 2009, pp. 147-150.
- [12] D. Molnar, and S. Schechter, "Self Hosting vs. Cloud Hosting: Accounting for the security impact of Hosting in the Cloud", *Proceedings of the Ninth Workshop on the Economics of Information Security* (WEIS), 2010.
- [13] R.L. Grossman, Y. Gu, M. Sabala, and W. Zhang, "Compute and storage clouds using wide area high performance networks", *Future Generation Computer Systems*, 2009, 25(2), pp. 179-183.
- [14] Costa, C. Ana, and K. Bijlsma-Frankema, "Trust and Control Interrelations New Perspectives on the Trust—Control Nexus", *Group & Organization Management*, 2007, 32(4), pp. 392-406.
- [15] Lund, M. Soldal, B. Solhaug, and K. Stølen, "Evolution in Relation to Risk and Trust Management.", *IEEE Computer*, 2010, 43(5), pp. 49-55.
- [16] J. Huang, and D. Nicol, "A formal-semantics-based calculus of trust", *Internet Compute IEEE*, 2010, 14(5), pp. 38–46.
- [17] K.M. Khan and Q. Malluhi, "Establishing Trust in Cloud Computing", *IEEE Computer Society, IT* professional,2010, 12(5), pp. 20-27.



- [18] J. Huang, and D. Nicol, "Trust mechanisms for cloud computing", *Journal of Cloud Computing: Advances, Systems and Applications*, 2013, 2(1), 9.
- [19] Department of Finance and Deregulation (DFD), "Privacy and Cloud Computing for Australian Government Agencies", *Better Practice Guide*, Nov. Australian Government Information Management Office, 2012.
- [20] R.A. Popa, J.R. Lorch, D. Molnar, H.J. Wang, and L. Zhuang, "Enabling security in cloud storage SLAs with CloudProof", USENIX Annual Technical Conference, Portland, 2011, pp. 31–31.
- [21] M.G. Jaatun, G. Zhao, A.V. Vasilakos, A.A. Nyre, S. Alapnes, and Y. Tang, "The design of a redundant array of independent net-storages for improved confidentiality in cloud computing", *Journal of Cloud Computing: Advances, Systems and Applications*, 2012, 1(1), 13.
- [22] K. Hwang and D. Li, "Trusted cloud computing with secure resources and data coloring", *Internet Computing, IEEE*, 2010, 14(5), pp. 14–22.
- [23] G. Tajadod, L. Batten, and K. Govinda, Microsoft and Amazon "A comparison of approaches to cloud security", *IEEE 4th International Conference on in*

Cloud Computing Technology and Science (CloudCom), Taipei, Taiwan, 2012, pp. 539–544.

Mrs. W.P. Eureka Priyadarshani is a Lecturer at the Information & Communication Technology Center, Wayamba University of Sri Lanka. She is a Science graduate of the University of Peradeniya and obtained her Masters in Computer Science also from the University of Peradeniya. Currently she is reading for her PhD studies on Cloud Database Trust and working as a visiting academic of University of Deakin, Australia.

Prof. Gihan N Wikramanayake is a Science graduate from the University of Colombo, Sri Lanka and obtained his M.Sc. and PhD in Computer Science from the University of Wales Cardiff, UK. He is a Professor, Director and Chairman of the Board of Management of the University of Colombo School of Computing (UCSC). He has been awarded the University of Colombo School of Computing research award twice.

Dr. E.M. Piyal Ekanayake is a Science Graduate from University Kelaniya, Sri Lanka. He obtained his M.Sc. and PhD from university of Kyushu, Japan and was a fellow in Mathematics, Atmospheric Dynamics at Oxford University, UK. He is a senior lecturer at the Department of Mathematical Sciences, Faculty of Applied Sciences, Wayamba University of Sri Lanka.