DETERMINING STEPS TO ENHANCE THE SECURITY OF DATA STORED IN THE CLOUD

[a]GABRIELA MACÁKOVÁ, [b]ZUZANA PRIŠČÁKOVÁ

Mendel University in Brno, Zemědělská 1, Czech republic
email:[a]macakova.gabriela@gmail.com,
[b]zuzana.priscakova@hotmail.com

IGA 02/2015 Klasifikace cloud computingu na základě zabezpečení dat.

Abstract: Current trend in the information technologies is the use of cloud computing as a storage sites. Data storage is also associated the data security. The contents of this article are methods for the data security and the data verification. The data security is based on the type of data. To determine the type of data we use the data classification by Kuyoro, Poynter and Rocha. However, the main element of data security is the data integrity in the cloud. For verify the data integrity we can use protocol call – response and verify the correctness of the data stored in the cloud. The aim of this article is summarize the above methods and determine the phases of data storage. Within the phases, we have determined the safety rules of the data storage.

Keywords: cloud computing, petri nets, data storage, data protection, data security

1 Introduction

Storing data in the cloud can be seen as relatively attractive form of outsourcing focused on daily data management. The real responsibility for data management falls under the company that owns the data. With view of this fact, it is important to understand some of the causes of data corruption. Such causes advise keeping the big responsibility of cloud services, some basic best practices for the use of secure data storage to the cloud, as well as the methods and standards for monitoring the integrity of the data regardless of data storage (Velte et al, 2009).

One of the main advantages of storing data in the cloud is unlimited access to the data without limitations in the time and place of access (Pearson et al, 2013). This property is important for the company, whose work activity takes place in various remote locations. For such companies it pays to enter the cloud solutions and on the ground, thus eliminating the load on physical storage devices, use the same computer and multiple access data in real time (real-time reporting) (Winkler, 2011).

In this case, it is important to create storage cloud to think about the specialty store. Although there are several hundreds of cloud storage, each storage site is geared to other requirements, such as storing communication via e-mail, store employee profiles, documentation storage projects and the like (Winkler, 2011). Of course requirement may be also store all types of documents.

In this article we have determined the sequence of storing different types of data with respect to data security. During storing data are formed dead-lock situations. For identifying critical situations we have created a Petri net. Through simulation of Petri net, we determined states, which data acquired. This article is solution for deducing the behavior of the cloud as data storage.

2 Materials and methods

For the compilation of Petri net is necessary to know the data communication between server and client. With the increasing number of data increases the number of required servers. For basic operational running cloud storage is needed one data server connected to the Internet. Communication between the data server and the client (customer of the cloud) is performed on the basis of requirements for storage in the cloud (Nielsen et al, 2013).

This communication is shown in the following figure (Fig. 1) through the UML sequence diagram. Communication interface is a web browser. Parameters for this communication are files that have been selected for placement on the server.

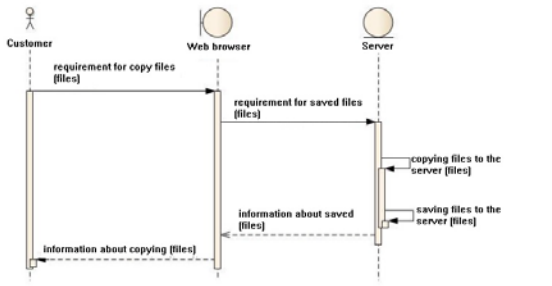

Fig. 1 Communication between customer and server through web browser

Model situations storing and accessing data in the cloud, we introduced the basic principle demonstration, but in reality cloud storage sites used by hundreds of data servers, and included the servers to maintain redundancy. It is for this reason that servers need to maintain, repair, and therefore it is necessary to store the data on multiple computers.

LAN redundancy is meant addition one (or two) back-up servers in a data center in case of problem (Lim, 2013). Since the actual use of virtualization capabilities, so we understand under redundancy cloning a virtual server on the same device, or all virtual servers cloning one device to another physical server, thus achieving the creation of shadow copies. Eswaran (2012) provides a simple algorithm to store data in the cloud.

If the cloud did not satisfy the condition of redundancy would also denied the basic characteristics and therefore client access to the cloud stored data at any time. As the most systems stores the same data on servers with different sources of power, so customers can seamlessly access their data while maintaining redundancy. Thus, the use of cloud as a data storage company is not simple, and therefore no direct hosting company's servers in a data center provider cloud solutions, but the data are divided into several parts (Mather, 2009), the individual parts can be scattered even within the whole earth.

When the cloud provider implements a redundant system, the data is again scattered throughout the cloud. Provider for these reasons does not provide redundant services directly connecting another server, but changing the allocation of resources to achieve a redundant system (Lim, 2013). It is also important to mention that while maintaining those conditions provides cloud and data security against theft (Erl et al, 2013), since the error on the local network could mean permanent loss, while the cloud companies will not lose your data.

A simple method for securing data include commonly used authentication client. Authentication system (process) allows access to data based on a correctly specified username and password (Mather, 2009). This method is one of the more simple and because it increased the possibility of breaking (Erl et al, 2013). In addition to authentication and authorization processes known procedures also provide user access to data, but in this case, the client will provide a list of persons authorized to access the data stored in the cloud (Rhoton and Haukioja, 2013).

For the phrase Storage as a Service is worth recalling that the client again works through software, which selects data backup, and then transfer them over WAN connections (Marsh, 2011). If any data loss, the client may at any time withdraw their data provider. Some people use providers and data archiving, and copying to DVD. In the event of data loss and data output requirements of the customer, the provider will send him only the drive with the copied data. For data security is also important to mention the fact that there are general principles which could drive the company to ensure data protection. However, there are standards for privacy policy (Poynter, 2008).

In the context of data security is often used in a corporate environment scheme classification data (Rhoton et al, 2013). Its aim is to highlight the need to further controls various data types that are processed businesses. The data classification scheme was developed on the basis of legal, regulatory and commercial requirements that the company must follow (Winkler, 2011). In the current corporate environment, there are three modes (or four) (Kuyoro et al, 2011), (Poynter, 2008), (Rocha et al, 2013):

- public/unclassified (e.g. marketing materials),
- internal use (e.g. information shared within the organization or with suppliers, such as the intranet),
- confidential/private (e.g. sensitive information such as credit card information),
- secret (e.g. market sensitive information, such as results at end year).

All data using cloud-based technology in the software are stored on servers at a remote location - data centers. Environment the data center allows enterprises to run applications faster with simplified management and less effort to maintain, and much faster scaling of resources (servers, storage, networks) relative to needs. The data center in the cloud environment contains information that is usually stored on end-user computers. By Chen and Zhao (2012) this type of data center cause for great concern with regard to the protection of user privacy.

By Ning (2014) use of virtualized infrastructure as a springboard may introduce new attacks on the integrity of user data. The data integrity is defined as accurate and consistent data stored in the absence of any modifications of data between two updates a file or record (Barsoum, 2013). Cloud services should ensure data integrity and to ensure confidence in the user privacy. Although data outsourcing to the cloud is economically attractive due to the cost and complexity of large-scale long-term data storage is still missing offer strong guarantees data integrity and availability of independence possible within a wide coverage of business and individual users of the cloud (Sun, 2014).

For monitoring the data integrity is important to consider these protocols:

- Service Level Agreement (Winkler, 2011), (Marsh, 2011), (Rhoton, Clercq a Graves, 2013),
- Proof of Retrievability (Eswaran, 2012), (Juel, Kaliski, 2012), (Neha, 2012),
- Protocol based on inserting random guardian in the data file (Sravan, Saxena, 2012).
- To verify the data integrity:
- Protocol call – response (Thuraisingham, 2013), (Cong et al, 2013) – If the data file is distributed to the cloud, then the system calculates forward by a number of short individual vector authentication tokens $G^{(j)}(j \in \{1,\ldots,n\})$, each token represents a random subset of data blocks that will be distributed to various cloud servers,
- Verify the correctness of the data stored in the cloud (Ming, 2013), (Yang, 2014) – If the user wishes to challenge cloud server t to verify the data integrity, then the user needs forward to calculate x authentication tokens for each $G^{(j)}(j \in \{1,\ldots,n\})$, change the key k_{chal} and the major permutation key K_{PRP}. For generating tokens i^{th} for server j, the user shall proceed follows:
  - creating a derivative of random values of the key challenges and permutations $k^{(i)}_{prp}$ based on K_{PRP},
  - calculates a set of randomly selected indices r,
  - counts tokens $v^{(j)}_i$ with a random challenge value α_i.

After generating tokens, the user can either manage pre-calculated tokens locally or encrypted values stored in the cloud. Summary of response from the server for each call (task) determines not only the accuracy of the repository, but also includes information on potential data errors.

3 Results

Based on these facts we have created a security model data stored in the cloud. Processing of this request, we modeled using a sequence diagram. The end user, which is logged in, may stores data. Login is repetitive operation condition checks credentials. After successful login, the client broadcasts a request for data storage, which is prepared based on the following steps:

- identify the input data-based on grade input data the system determines what data to be stored,
- require encryption of data – when the input data is necessary to require data encryption,
- determine the level of data security – each imputed data to authorized data security due to their sensitivity. On the basis of the authorized security level is assigned to data security,
- determine the minimum permissions – minimum permissions are an extension of the authorized data security. Minimum permissions determine what minimal operations are allowed for the implementation of data taking operations do not cause data leakage or new risks,
- encrypt data – on the basis of the previous steps 3 and 4, the data is encrypted,
- mark the data – to maintain data integrity is attributed data code (hash value), flagging data consists of the following steps,
- identify risks – for labeling data is needed to determine the risk procedure that can be performed with the data and could lead to the theft or total loss of data,
- create the security policy – when the risk is compiled security policy, which shows the way to address risk procedures and means of prevention,
- determine the boundary of protection – boundary of data protection remains at the base of the identified risks and security policy. Defining the boundaries of data protection we separate risky operations to implement with data from safe operations,
- ensure of data – secure data is encrypted and assigned code. Secured data contain a defined risk, security policy and data protection threshold,
- determine a type of security – secure data assigned the type of security, i.e. inclusion in the security level under level data contained in the system,
- confirm the security – unless they contain protected data type of security, we do not consider the data as secure. Once the type of system perceives data for secure,
- confirm the save data – after confirmation of proper data security, data is stored and the end user receives information about saving data.

Sequence of steps 1-13 is strictly defined and the output of one step is the input for the second step. Procedure cannot be performed simultaneously. These steps are handled operand input data. Steps 3 and 4 loop, which is a condition determining eligible security. To ensure the operation 4 is important to check the identification number and the minimum level of permissions. Incorrectness occurs when data breach data security system. Procedures 7-11 represent the loop systems, because without them the correct output system do the following steps 12 and 13.

For methodology of determining the security data stored in the cloud we transformed sequence diagram into the Petri nets (Fig. 2). The Petri nets we set the step value to 1 ms, the time for completion of the simulation to 1000 ms and the maximum number of simulations is 1000. The Petri net consists of 22 transitions and 33 places. Weight value in each case is 1. Seating capacity is mostly 1, only the points that serve as terminal configurations of the model are determined by the capacity to 100. Transitions are defined without waiting a time instant model.

In the Petri net apply following definitions:

Denote by p(t) the number of tokens in place p. Then, for each transition rule for all points of entry when this transition.

The definition of Petri nets model, increasing safety data stored in the cloud:

The Petri Net model increases the security of data stored in the cloud is safe, bounded and conservative, because for each of the assessment is valid $z(p) \le 1$, *while* $\exists k \in N_0; \forall z, p; z(p) \le k$ *and for each of its condition is true, the total number of tokens is constant:*

$$\forall i; \sum_j z^{(i)}(p_j) = k$$

This definition implies the transitory characteristics, their definitions are as follows:

Provide input data to the end user can only occur if the administration verifying the user logs into the system and assigns the data to the user's account.

If the system wants to determine the level of security the user must then made available to the authorized security file located on the server and perform verification of authorized security.

To encrypt the data you need to load at least authorizations and changes done in the minimum permissions.

Identified risks are compiled on the basis of previously identified risks imposed on the system and identified new risk procedures with the new risk operation cannot be already stored in the system.

For allocation the boundary of data protection, it is necessary to load the previously determined threshold data protection and security policies in place for this type of data.

If granted re-entry into the system, and the data were stored data or are not allowed to the user's account, or the client confirmed the continuation of the work.

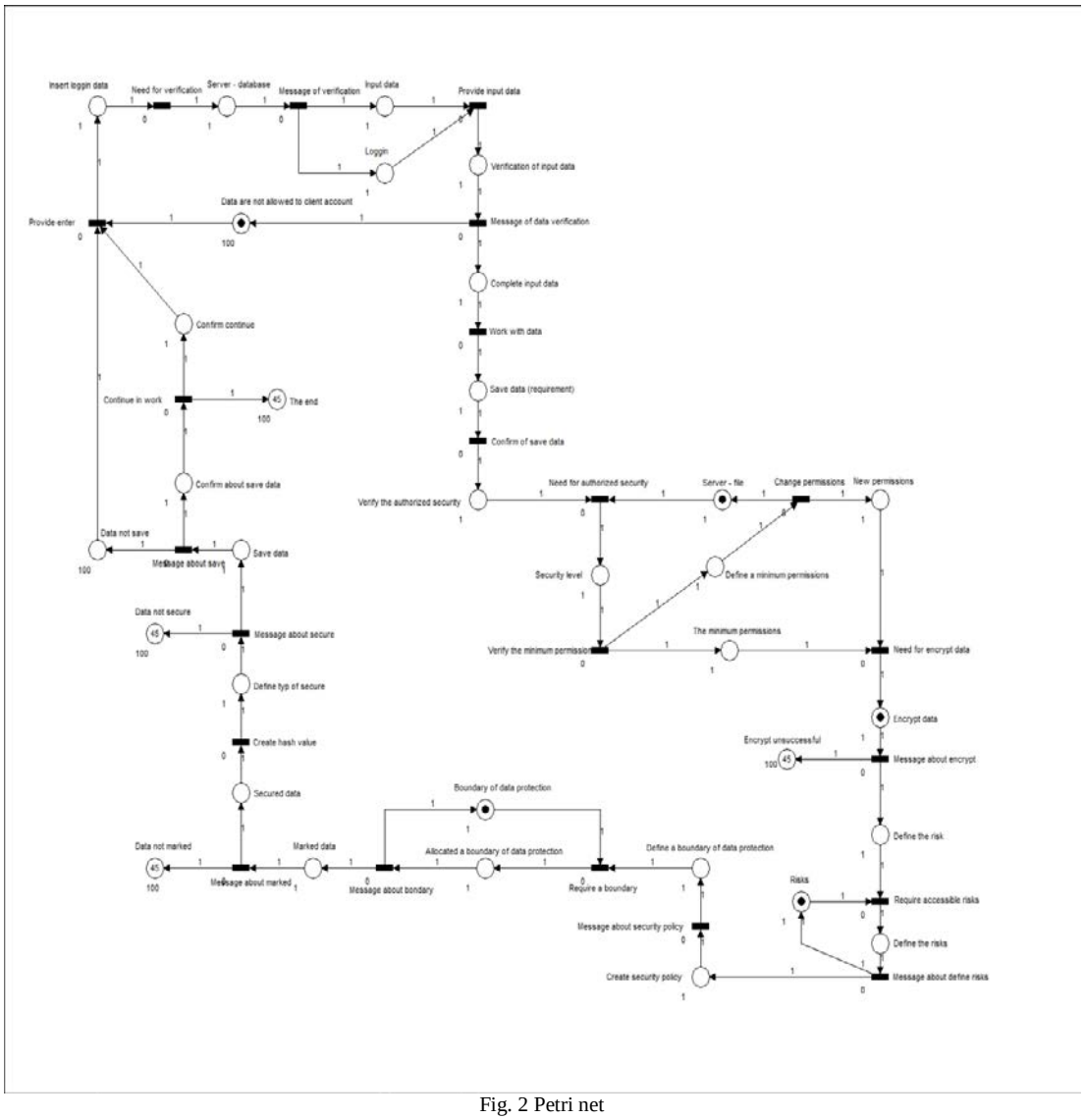

Fig. 2 Petri net

4 Conclusion

The current trends include the use of cloud as the cloud data storage. When we store data in the cloud, we need determine the safety for storing data. We determine the safety through the classification data scheme by Kuyoro, Poynter and Rocha. The next step is verify of the data integrity through the protocol call – response by Thuraisingham (2013) and by Cong (2013), and

verify the correctness of the data stored in the cloud by Ming (2013) and by Yang (2014). The referred verification does not include process for the storing data.

In this article, we defined the process of storing data in order to enhance to the data security. The process consists of the following steps: identify the input data, require encryption of data, determine the level of data security, determine the minimum permissions, encrypt data, mark the data, identify risks, create the security policy, determine the boundary of protection, ensure of data, determine a type of security, confirm the security, confirm the save data. The above steps are visualized in the Petri net. From the Petri net, we establish the basic rules of the process of storing data in the cloud.

Although we defined the process of storing data, it still remains an open question whether it is not appropriate to use other means of verification of security.

Literature:

1. BARSOUM, A. Data Integrity in Cloud Computing Systems: Challenges and Solutions. New York: LAMP LAMBERT, 2013.
2. CONG, W., SHERMAN, S. M. CH., QIAN, W., KUI, R., WENJING, L. Privacy-Preserving Public Auditing for Secure Cloud Storage. IEEE Trans.Computers, 2013.
3. ERL, T., PUTTINI, R., MAHMOOD, Z. Cloud Computing: Concepts, Technology Architecture. New York: Prentice Hall, 2013.
4. ESWARAN, A., ABBURU, S. Identifying Data Integrity in the Cloud Storage. [online]. [cit. 2013-08-06] Available: http://ijcsi.org/papers/IJCSI-9-2-1-403-408.pdf
5. CHEN, D., ZHAO, H. Data Security and Privacy Protection Issues in Cloud Computing. International Conference on Computer Science and Electronics Engineering, 2012.
6. JUELS, A., KALISKI, B. S. Poors: proofs of retrievability for large files. New York: ACM, 2007.
7. KUYORO, S. O., IBIKUNLE, F., AWODELE, O. Cloud Computing Security Issues and Challenges. International Journal of Computer Networks, 2011.
8. LIM, I., COOLIDGE, E., HOURANI, P. Securing Cloud and Mobility: A Practitioner´s Guide. New York: CRC Press, 2013.
9. MARSH, CH. Data Integrity In The Cloud. [online]. [cit. 2013-12-05] Available: http://www.wwpi.com/index.php?option=comcontentview = articlecatid = 99 : cover storyid = 12800 : data integrityin
10. MING, L., SHUCHENG, Y., KUI, R., WENJING, L., THOMAS, H. Toward privacy assured and searchable cloud data storage services. IEEE Network, 2013.
11. NEHA, T., MURTHY, P.S. A novel approach to data integrity proofs in cloud storage. [online]. [cit. 2013-01-05] Available: http://www.ijarcsse.com/docs/papers/10October2012/V olume2issue10October2012
12. NIELSEN, L. The Little Book of Cloud Computing SECURITY: 2013 Edition. Rhode Island: New Street Communications, 2013.
13. NING, C., CONG, W., MING, L., KUI, R., WENJING, L. Privacy-Preserving Multi-Keyword Ranked Search over Encrypted Cloud Data. IEEE Trans, 2014.
14. PEARSON, S., YEE, G. Privacy and Security for Cloud Computing. London: Springer, 2013.
15. POYNTER, K. Review of information security at HM Revenue and Customs. London: HM, 2008.
16. RHOTON, J., CLERCG, J., GRAVES, D. Cloud Computing Protected. New York: Recursive Press, 2013.
17. RHOTON, J., HAUKIOJA, R. Cloud Computing Architected. New York: Recursive Press, 2013.
18. ROCHA, F., ABREU, S., CORREIA, M. The Next Frontier: Managing Data Confidentiality and Integrity in the Cloud. IEEE, 2013.
19. SOSINSKY, B. Cloud Computing Bible. New York: John Wiley and Sons, 2011.
20. SUN. Sun Cloud Architecture Introduction White Paper (in Chinese). [online]. [cit. 2014-01-15] Available: http://developers.sun.com.cn/blog/functionalca/resource/sun353cloudcomputingchinese.pdf
21. THURAISINGHAM, B. Developing and Securing the Cloud. New York: Auerbach Publications, 2013.
22. VELTE, T., VELTE, A., ELSENPETER, R. Cloud Computing, A Practical Approach. New York: McGraw Hill, 2009.
23. WINKLER, V. Securing the Cloud. New York: Syngress, 2011.
24. YANG, K. Security for Cloud Storage Systems. London: Springer, 2014.

Primary Paper Section: I

Secondary Paper Section: N