A Mechanism for Dependability Processing of Information Flow in Cloud Computing

Hao Liu, Linpeng Huang, Chen Li, Luxi Chen
Department of Computer Science and Engineering
Shanghai Jiao Tong University
Shanghai, China
e-mail:{liuhaosjtu, lphuang, lcroy, leen1988}@sjtu.edu.cn

Abstract—As cloud computing becomes more and more common among enterprises and users, the dependability problem on cloud environment, especially in terms of correctness, reliability, security and performance, attracts an ever increasing attention. It is therefore necessary to provide a comprehensive mechanism to solve the dependability problem. Through this work, we propose a new comprehensive mechanism that dependability processes the information flow on a cloud platform. More precisely our approach includes information obfuscation and conversion method. It also allows, the monitoring and verification methods of the information flow processing. realizes the security of information flow transmission, the dependable hierarchical storage of information, the service provided by cloud safety and correctly, enhanced the dependability of the information flow processing in cloud computing.

Keywords-information flow; dependability; cloud computing; decentralized label model

I. INTRODUCTION

At present, a large number of services are provided by cloud platforms. The interaction between these services, such as the delivery of data transmission or the communications between applications, is using the information flow to deliver messages. Compared to previous computing models, the different points of services provided by a cloud platform are listed as follows:

a) Shared control mode has changed;
b) Data storage mode has changed;
c) Data protection requirement as increased;
d) Boundary control of network requirement has increased.

These changes have brought new challenges to information flow dependability, especially regarding the following three aspects:

a) **Network transmission**. When a user access the service resource provided by the cloud platform, data transmission through the network is essential. How to ensure that sensitive data in the transmission process is not illegally leaked, attacked, modified or destroyed is a major concern to address.

b) **Information flow processing**. To date, no suitable mechanism that could guarantee the dependability of the information flow processing on the cloud platform exists. In deed there is no solution to really ensure the cloud services will response to the user quickly, without the stealing any private information. This is a major issue as information and computing resources are shared among cloud services which do not trust each other.

c) **Cloud storage**. Centralized data storage and physical resource sharing in cloud computing introduce major data security and privacy concerns. Security cannot be guaranteed by any physical boundary defined for the machine or the network, hence increasing the difficulty of storing data in the cloud.

Based on the analysis of a cloud platform architecture, this work mainly focuses on four aspects of dependability, namely correctness, reliability, security and performance. We discuss the problem of the information flow in a cloud environment, that is (1) how to ensure the security of the transmitted information; (2) how to render cloud services fast, effective in the information processing; (3) how to guarantee the security and reliability of the information stored in the cloud. We propose an innovative solution, which can provide theoretical and technical support for widespread applications in cloud computing.

II. RELATED WORK

Cloud computing has received much attention over the past few years. In fact a wide variety of topics, encompassing both the technologies and their applications, was covered. Among others, it includes cloud computing platform architecture[1,2,3],dynamic information flow tracking technology, hardware expansion on a cloud platform [4], storage security of the cloud data[5,6,7] and service model architecture [8].

In the literature an additional layer composed of a software Virtual Machine Monitor (VMM),which introduces system virtualization, was presented [9].The VMM is completely isolated from the runtime environment, providing services such as security log, intrusion detection and virus protection. Binary obfuscation of the control and data flows using the normal control flow conversion was realized in [10].Unfortunately its concealment is not very strong as an attacker can remove some impossible conditions by only using constrain analysis to the condition code.

III. IMPLEMENTATIONOFTHE MECHANISM

Using the information flow requirements for depend ability on cloud platform, this paper designs an information obfuscation method, and proposes some new dependable monitoring and verification mechanism in information flow. We also realize adependable hierarchical information storage, and illustrate both the feasibility and effectiveness of the method using a typical application example.

(1) Information obfuscation and conversion between the clients and the cloud. From an information life cycle perspective, the key to ensure information flow security is

to guarantee the safe information transmission between the clients and the cloud. A transfer agent between the usersand the cloud is constructed. It offers information obfuscation and conversion in order to confidentially and privately send data to the cloud data flow.

Based on the information transmission between the user and the cloud environment, we analyzedpotential security threatsand privacy issues arising during a transmission process.

(2) The processing and validation mechanism of the information flow on the cloud

In order to fully take advantage of the information and computation resources either in the services or among them, it is vital to control information processing.

Our mechanism combines a decentralized label model together with formal semantics and advanced programming language. It buildsan information processing verification agent on the cloud to guarantee the execution of the information flow in or among services which do not trusted one another. As expectedit complies with the specified reliability strategy.

Analyze the existing differences and problems introduced on a cloud platform compare to a more traditional computing platform.

Study secure programming rules of advanced programming language based on cloud computing applications, static information flow analysis technology. It should also take into account the user oriented program annotations, decentralized label model, formal semantics expression, privacy and integrity strategy and the type checking mechanism during compile-time and run-time.

Studythe data migration and computation migration of services which do not trust with each other on the cloud platform.Ensure the coincident of execution resultsfor the embedded transaction mechanism, information access control with automatic verification and automatic generation of log information.

Study the dependability problems among different combinations of cloud services, including (1) the reliability, such as service copy or tolerance mechanism; (2) the correctness that is behavior consistence, whether the service satisfies the constraint requirements of the system, prevent the loss of important information, or deadlock between service calls; (3) the security i.e. comply with the copied service instance content authentication requirements, assure the absence of leakage of user private data through a hidden interface; (4) the performance, such as the selection of performance index and the evaluation criteria of combined services.

(3) Information dependable hierarchical storage on cloud platform.

Many copies of the data information would be stored on the cloud platform. Private data can be easily cached, replicated and archived by the third party. Furthermore, once the data is stored on the cloud, the control authority of the data is transferred from the user to the cloud storage provider. This process renders easy for a malicious provider or attacker to steal user private information, disseminate or just use it. Thus it is necessary to design an access control agent on the programming platform, offering data storage monitoring which have been converted after obfuscation, ensure the dependability of data storage.

Through an analysis of the risks faced by the information stored on a cloud platform and of the related solutions, wedesigns a new dependable hierarchical storage model, and realize the dependable protection of the data.

This paper mainly studies the solution architecture as shown below:

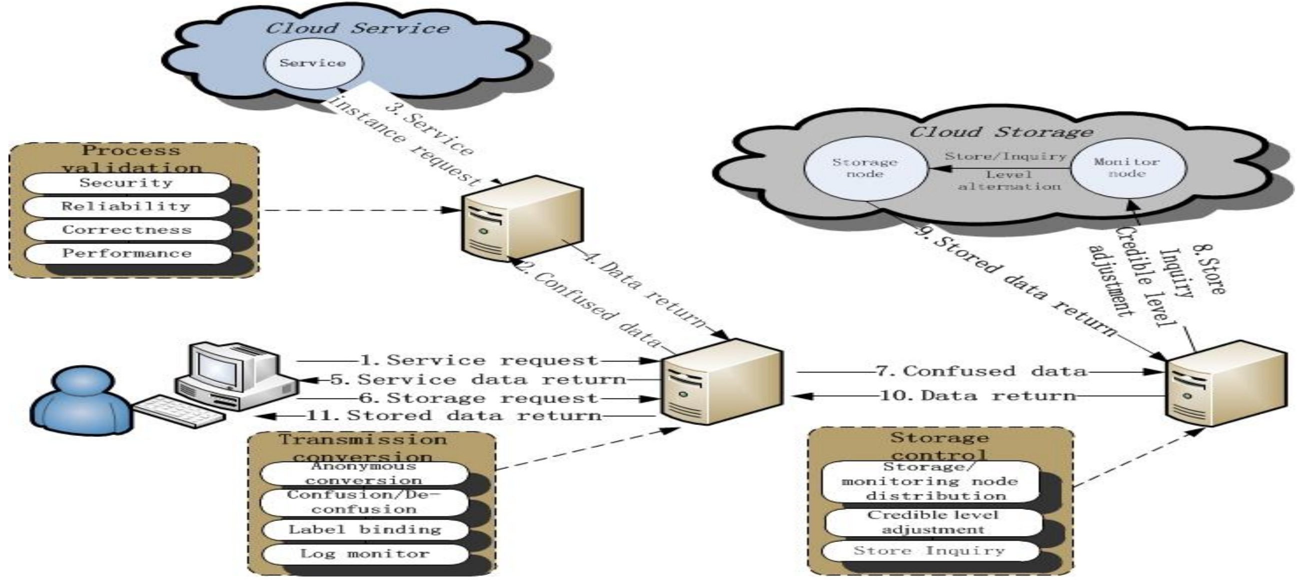

Figure1. Main architecture of the mechanism

This mechanism is based on dynamic information flow on a cloud environment. In order to meet the dependability requirements on a cloud platform, our mechanism designs an information obfuscation method, proposes some dependable monitoring and verification mechanisms of the information flow while also realizing dependable hierarchical storage of the information. Through the obfuscation and conversion to the process of information transmission the validation to process and the control to the storage process, achieves the dependable control of information flow.

1.Using Decentralized Label Model(DLM)to realize label binding, guarantees the privacy and security during data transmission, processing and storage on the cloud platform.

The Decentralized Label Modelis composed of the following parts:

Principal: Principal is the owner, updater and releaser of information. In a cloud environment, it can represent a data owner or a service provider.

Label: The lemma of Decentralized Label Model is using labels to annotate the program and the data in order to represent their security level.

2.Security strategy: Security in the cloud environment is used to specify the access and modify the permissionson the sensitive data, including confidentiality strategy and integrity strategy. The confidentiality strategy is used to identify which subject can access specific information or invoke a service; integrity strategy is used to identify which subject can modify the information or services on the cloud.

By studying typical applications on a cloud platform, such as providing special education services, and analyzing the information transmission that may exist between the clients and the cloud platform, this approach highlights the hidden dangers existing in the process of information transmission. Figure 2 shows the design of the conversion agent in the process of transmission.

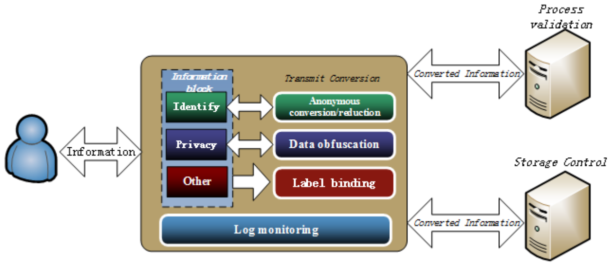

Figure. 2 Information transmission conversion agent

During a transmission process the security of any confidential data is assured by applying a conversion (or reverse conversion) that renders the user anonymous such that the client output (or input) information follows the privacy principles, and sensitive data is obfuscated (or reverse obfuscated), Furthermore the log manager is used to monitor logged data returned from the cloud.

Using data obfuscation technology, realize shiding or obfuscation of confidential data throughout an information flow sent to the cloud. The goal is that no user confidential information leaksatany stage during the transmission process.

Using anonymous conversion technology, the true customer is hidden，Hence using an anonymous identity to access cloud services ensures that the user's personal information will neither be identified nor stolen.

Using the decentralized label model, realizes the label binding to data dependability requirements, and enact the corresponding security strategy. It protects sensitive data in the information flow, and pretreats it for a future hierarchical storage on the cloud.

Using log monitoring technology, allow to track information passing in and out the cloud, check the contents of the information flow, record the information nusage into the log, and return the log to the transmission conversion agent for being reviewed. Monitor the cloud services and how the information sent by clients is used(including whether a large amount of data is copied),is fundamental for detecting the usage and the spread of any malicious information.

3.Take the analysis of cloud service dependability in cloud environment as basis. The dependability guarantees that the mechanism is divided into four aspects: correctness, reliability, security and performance. First, combining the decentralized label technology and the formal method, allow us to define a decentralized label model based on formal semantics, and to build a platform for constraint and monitoring the processing of the information using some advanced programming language. This prevents any attacker from accessing, intercepting or disseminating the private data. Then, using formal methods (e.g.π-calculus, graphic method) define some formal specifications and modeling of the service, verify the consistency of the service behavior, in addition, the authentication of the service instance and description ensuresits safety, The performance evaluation index does the evaluation and selection of the service performance. Details are provided in figure 3.

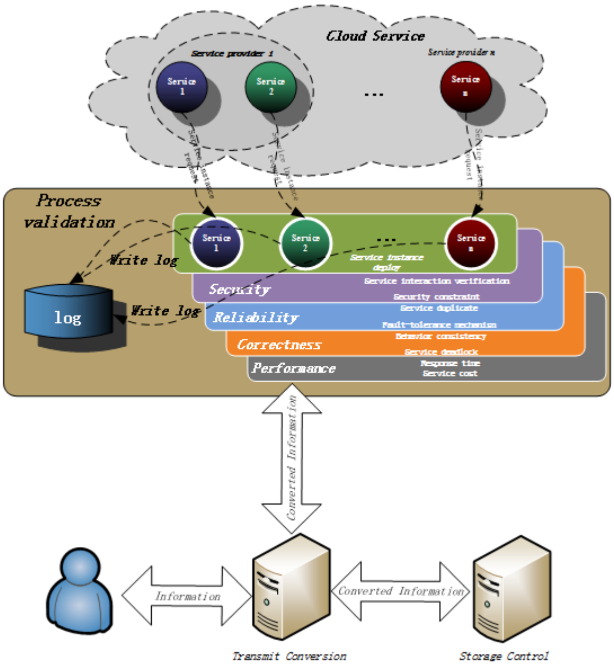

Figure. 3Processing and Verification Process

In services: Extracts related confidential properties involved in the execution process of the information flow. Then, add a formal semantic description of the decentralized labels and execute a process to constraint them. Analyses the executed process using static information flow analysis technology, and verify whether the service satisfies certain constraints and leaks any information.

Among services: As data and computation migration mainly occur among the cloud service, it is necessary to analyze the dependability problem among cloud services which do not trust each other. A formal modeling of the cloud services which satisfies the requirements, helps to analyses the call process and constraint among the services. This ensures the consistency behavior of the cloud services and avoid any mismatch and deadlock problem.

Using access control annotations and the type-checking mechanism at both compile and run times permits to verify the security of the information flow process. Then log any access to the information and return the logged information to the transmission monitoring agents for review.

The performance indices such as the service response time and the service cost, permit to establish the performance evaluation scheme, evaluate and select the services satisfying the requirements, and create the final call strategy for dependable cloud services.

4. Before the client sends the information to the cloud, hierarchical label binds the sensitive data following the security requirements through the transmission monitoring agent. Then the data in the information flow isassigned to the corresponding level storage node by the storage monitoring agent. Details are explained in figure 4.

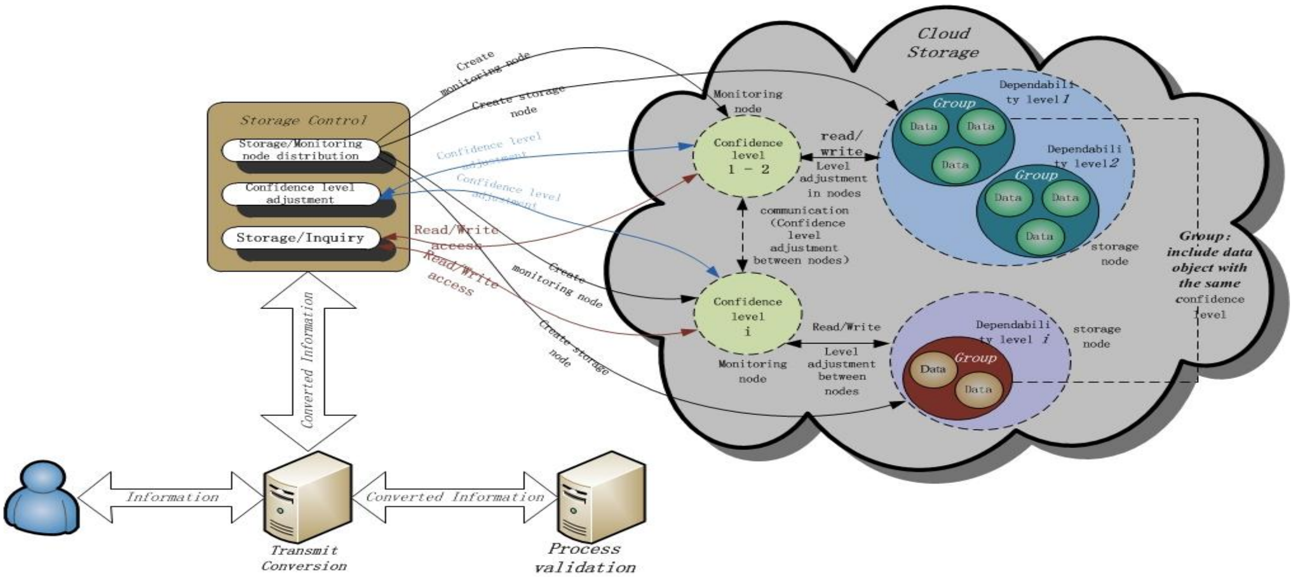

Figure 4. Dependable hierarchical storage of cloud information

According to the dependability requirements on the sensitive data in the information flow, our mechanism designs a storage control agent based on a formal semantic label programming platform, more specifically it defines a hierarchical storage model.

According to the hierarchical storage strategy, our mechanism specifies the level of information dependability. It analyses of the data on the cloud using the information flow analysis technology, and according to the confidentiality and integrity requirements of the sensitive data, separates the data into different levels, as the basis of hierarchical storage.

Following a traditional hierarchical storage strategy, our mechanism sets the monitoring node, uses the tag access control technology to set the monitoring nodes as the global highest security rating label. Then it, analyses the security level of the data object to be stored, check whether the security level of the binding storage node matches the security level of data object. Using a dynamic monitoring technology, it can detect any change in the security level of the data object.

According to the hierarchical storage strategy, it sets the storage node. Using a dynamic binding technology it maps the storage nodes to the designated monitor node, and matches the storage access process for any given data object.

IV. CONCLUSION

In this paper, we present a comprehensive mechanism for dependability processing of information flow in could computing environment. In this approach, we propose an information obfuscation method, a hierarchical storage strategy and realize an information flow control method. In order to verify the designed dependability guarantee mechanism of information flow processing on cloud platform, take the project "Application of cloud platform to assist regional special education development research" as the basis, lay the foundation of high performance cloud computing platform built in Shanghai Jiao Tong University in 2012 as a verification platform, deploy the information flow dependability guarantee mechanisms designed in the paper to special education application. These parts of the mechanism can enhance the dependability of the information flow processing in cloud computing environment independently. And our future work would be to explore the mechanism more deeply in each part of the approach.

ACKNOWLEDGMENT

The work described in this paper was supported by the Innovation Program of Shanghai Municipal Education Commission (No.13ZZ023),and the 973 Program of China (No.2009CB320705).

REFERENCES

[1] Dean J, Ghemawat S. MapReduce: Simplified Data Processing on Large Clusters. Proceedings of Usenix Symposium on Operating System Design and Implementation. 2004: 137-150.

[2] Ghemawat S, Gobioff H, Leung S. The Google file system. ACM SIGOPS Operating Systems Review, 2003, 37(5): 29-43.

[3] Chang F, Dean J, Ghemawat S, Hsieh W, Wallach D, Burrows M, Chandra T, Fikes A, Gruber R. Bigtable: A distributed storage system for structured data. Proceedings of the 7th USENIX Symposium on Operating Systems Design and Implementation. 2006: 205-218.

[4] Crandall J, Chong F. Minos: Control Data Attack Prevention Orthogona to Memory Model. Proceedings of Micro, 2004, 221-232.

[5] Frank Dabek, M. FransKaashoek, David Karger, Robert Morris, and Ion Stoica. Wide-area cooperative storage with CFS. In Proc. 18th ACM Symp. on Operating Systems Principles (SOSP), October 2001.

[6] P. Druschel and A. Rowstron. Past: A large-scale, persistent peer-to-peer storage utility. In In Proc. IEEE Workshop on Hot Topics in Operating Systems, SchossElmau, Germany, May 2001.

[7] John Kubiatowicz, David Bindel, Yan Chen, Steven Czerwinski, Patrick Eaton, Dennis Geels, Ramakrishna Gummadi, Sean Rhea, Hakim Weatherspoon, Westley Weimer, Chris Wells, and Ben Zhao. OceanStore: An architecture for global-scale persistent storage. In Proc. 9th international Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS 2000), November 2000.

[8] Subashini, S. and V. Kavitha : A survey on security issues in service delivery models of cloud computing. Journal of Network and Computer Applications. 2011: 1-11.

[9] Chen P, Noble B. When virtual is better than real. Proceedings of the 2001 Workshop on Hot Topics in Operating Systems (HotOS). 2001 133-138.

[10] Myles G, Collberg C. Software watermarking via opaque predicates: Implementation, analysis, and attacks. Electronic Commerce Research, 2006, 6(2): 155-171.