DATA MOBILITY AS A SERVICE

Thanh Dat Dang
School of Computing and Communications,
University of Technology Sydney, Australia
Thanh.D.Dang@student.uts.edu.au

Doan Hoang
School of Computing and Communications,
University of Technology Sydney, Australia
Doan.Hoang@uts.edu.au

***Abstract*— Cloud computing and cloud services provide an alternative IT infrastructure and service models for users. The users use cloud to store their data, delegate the management of the data, and deploy their services cost-effectively. This usage model, however, raised a number of concerns relating to data control, data protection and data mobility: 1) users may lose control of their resource; 2) data protection schemes are not adequate when data is moved to a new cloud; 3) tracking and tracing changes of data location as well as accountability of data operations are not well supported. To address these issues, this paper proposes a novel cloud service for data mobility from two aspects: data mobility and data protection. A data mobility service is designed and implemented to manage data mobility and data traceability. A Location Register Database (LRD) is also developed to support the service. Furthermore, data is protected by a data security service CPRBAC (Cloud-based Privacy-aware Role Based Access Control) and an Auditing service that are capable of verifying data operations and triggering alarms on data violations in the Cloud environment.**

Keywords— data mobility; cloud data mobility service; data location; traceability location; data protection.

I. INTRODUCTION

Cloud computing has grown into an efficient computing model with lower costs, rapid scaling, easier maintenance, and services availability anywhere, anytime [1]. Many cloud services have been developed to facilitate users shifting their business to cloud infrastructure [2]. Data services [2] have been developed as software as a service to provide tool to access data in cloud. However, few research efforts focus on data mobility issues related to data location and data protection when data is stored at its home cloud or moved to a new cloud. Despite the fact that Cloud Service Providers (CSPs) have established Service Level Agreement (SLA) with users, when user's data is moved to other clouds located at different locations from its home, little has been done to assure the users the whereabouts of their data and what happens the data. Ensuring the users receive the evidence of their data's violation whenever needed also presents a real challenge for cloud service providers.

The recent researches on data services in cloud computing have partially addressed some issues. Data as a service is represented as software as a service for data access across platforms and systems. This type of services is used to share and process a large collection of data in cloud environment [3]. The work in [4] presented the data protection as a service, in which security primitives are supported by a cloud platform within data encryption. Geographic data location [5-7] is one of well discussed approaches which adopt mechanism to detect data movement among clouds and verify data and user access location. However, it focuses little on data protection and relationship between data and users. Another solution [8, 9] focuses on securing data at cloud by offering data traceability service or cryptographic primitives but not on solutions for moving data among clouds. More importantly, the lack of secured components or equivalence protection schemes at foreign locations results in various data violations including location, operation, and illegal content disclosure violations. Hence, a cloud service is clearly needed to track and protect users' data and prevent illegal requests.

To address the aforementioned problems, this paper proposes a new but fundamental cloud computing service, the Data Mobility as a Service (DMaaS). The proposed service focuses on two aspects: data mobility and data protection. Data mobility deals with the location changes of users' data in the cloud, and ensures that locations of data are accounted for and that the destination of a move operation is conformed with the SLA established at the time when the user subscribed to the CSP. Data protection deals with access permissions, auditing data operations and recording changes on data. By integrating data protection and data location considerations, DMaaS achieves tracking data locations efficiency, protecting and recording activities and operations on the data for auditing purposes. DMaaS has the following distinguished features.

- *Location registration support.* Data location storages need to be built to store changes of data operations and locations. Offline recorded locations will be valuable to detect data when losing the connection with location services. To enhance the link between data and users, data locations must be periodically transmitted their status to original cloud by updating new data locations at LRD and notifying users via DMaaS. In doing so, changes of data location are logged at LRD for verifications and users' concerns related to privacy and security when their data is located in the specific location are mitigated.
- *Data accessing support.* The Access transparency on the data enables applications on various cloud platforms using data without time-consuming converting operations that may involve third parties or middleware tools. This helps reduce the costs of cloud deployment through economies of scale for a large-scale platform provider.

- *Verification platform support.* Verification services accept access requests with essential parameters to analyze and grant access permissions or reject requests. Platform providers should build support for confinement and auditing into the platform in a verifiable way. CSPs could communicate with each other for verifications when users send a request to access data. The process can be executed at original side or visited side with a simple deployment of protection service across various platforms.
- *Data and User binding support.* Ensuring that the CSP timely truthfully informs the users data violations is an essential service for cloud mobility management. Data operations are audited and recorded based on users' permissions for tracing purposes. Illegal operations are not only terminated but also immediately notified the data owner and the cloud administration. Therefore, cloud users are able to retrieve the status of their data via the service. If the CSPs are required to report the evidences of the data violations to the users, the users are more inclined to adopt cloud solutions for their businesses as they can build more acceptable SLA with their subscribed CSPs in a more trustworthy relationship.

The rest of the paper is organized as follows. Section II presents our motivation. Section III presents the DMaaS. Section IV describes the service interfaces. Section V presents the analysis of the DMaaS and the implementation The conclusion and future work will be drawn in the section VI.

II. Motivation

Our work is motivated by the lack of location register for the data when it is allocated or moved to a new cloud. In other words, DMaaS has not been developed and supported by CSPs. Even when users' data is located in specific locations that are covered by the scope of the SLA with the CSP, violations against the established SLA may occur and damage the data. Furthermore, despite the fact that third party auditors can be deployed to periodically monitor the data, attackers can exploit the vulnerability of the system to compromise the data and depart without a trace if the data cannot protect itself. Reporting evidences on violation and informing user present real challenges for CSPs. This lack of data mobility service leads to three basic challenges as follows:

Location management challenges. It is difficult to maintain data locations when data is moved away from its original cloud or transferred among clouds frequently since data location information and/or structures for collecting and keeping location information are not available. Cloud providers must provide a location database to record data location when a user first subscribes its data and track the changes in data location subsequently when moved. Furthermore, given the number of data locations could be very large and stored in the distributed environment, tracking and tracing of data locations are time and resource consuming and should be delegated to a specific data mobility service.

Security and privacy challenges. CSPs normally provide appropriate security mechanisms to protect data at their own storage but they lose the control when data moves to another cloud where there is no equivalence protection scheme. Enhancing mobility while maintaining security and privacy becomes a primary challenge for CSPs in building mobility management models that can be deployed at various cloud platforms with rapid development and maintenance.

Coupling between data geographic and data security. Currently, data location and data protection represent separate concerns and independent from each other. There may exist inconsistencies in tracking data location and protecting data in cloud environment.

Our work focuses on designing an important class of cloud platform services and building a generic infrastructure to support the data mobility service. In particular, a data location register database is designed and implemented to store information about data registration, data owner, operation, location and timestamp. Despite the fact that data locations can be changed dynamically, data owner still relies on this transparency scheme to build more acceptable SLAs with their subscribed CSPs to protect their data and business.

III. Data Mobility as a Service (DMaaS)

Currently, users rely primarily on some prior agreements with their CSP for data protection. However, users are still concerned about the integrity of their data once it is entrusted to the CSP that allows data to move about in the cloud environment. The concerns are legitimate due to the lack of location services, protection services and user tracing services in the environment.

This section presents the design of our proposed service. It consists of three components as follows: 1) the mobility component, 2) the trusted security component, and 3) the cloud interface. Figure 1 depicts the design of DMaaS. The Cloud Privacy-aware Role Based Access Control (CPRBAC) and the Active Auditing (AU) are deployed to deal with security; the Mobility Service (MS), the Location Registration Database (LRD) and the Traceability Service (TS) are designed to deal with data mobility. DMaaS is designed to achieve the following goals. It allows data to be moved securely while ensuring users can perform all data operations at various platforms. It allows simple interactions among clouds for communicating data locations and verifications. Protection services can be deployed at both the original cloud and the visited cloud where data is moved to from its home cloud to protect data.

DMaaS enforces location management at both the cloud management level and within the data itself through the location database for either tracking or tracing locations. It uses the LRD to store and update changes of data location. Data protection components employ both cryptographic and role based access control to offer robust logging and auditing to provide accountability.

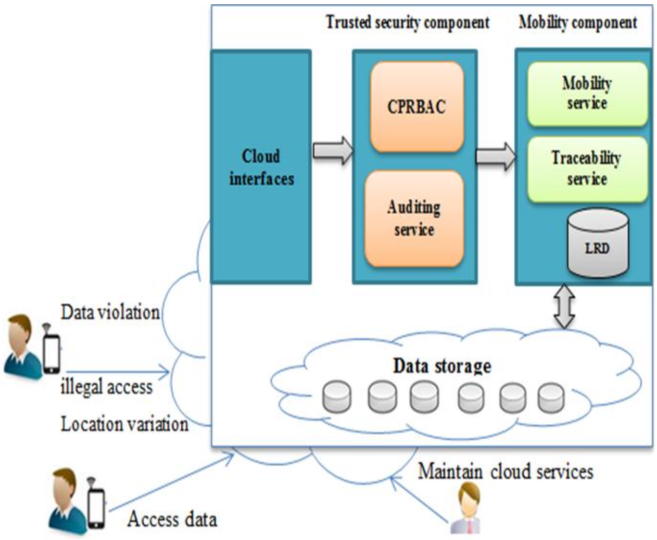


Fig. 1. The architecture of DMaaS

A. *Mobility component*

Traceability Service (TS): data owner will be notified through the messages to follow their data's paths and changes made by various entries on the cloud that access it. Depending on the source of access requests, the messages can be sent to data owner or their parties if there is a trust access at cloud providers for multi users. Often, in a cloud environment, highly complex and dynamic hierarchical service chain may allow the data handling task to be delegated from a CSP to another CSP in a flexible manner. The Trace Service ensures that user still maintain its control over the data even it is stored in a new cloud.

The Location Register Database (LRD) stores information related to user such as data identification and cloud location, etc. When data subscribes to the cloud, it needs to register with a TS which is responsible for tracing and with a notification message if user requests data locations.

Mobility Service (MS) is responsible for creating queries to the LRD. When the data is created at the cloud, the MS queries the LRD to insert a new record including the information about the data. When data is moved to a visited cloud, the MS will invoke a query to update new data location.

B. *Trusted security component*

CPRBAC supports for fine-grained access control requirement when data is deployed in a multi-user environment where one user can assume multiple roles in accessing the data. Normally, the actions on the data must be authorized according to the user authentication process. Only valid users with authorized permission can access and move the data, while unauthorized actions will be rejected by security cloud services.

C. *Cloud interface*

Cloud access interfaces provide a service interfaces to access the DMaaS. It forwards requests with parameters to the security management component to verify access permissions.

IV. DMaaS INTERFACES

This section presents specifications of the DMaaS interface which enables establishing a secure communications between users and the cloud service. Table 1 gives the description of the notations that are used in the service.

TABLE 1 NOTATION USED IN THE SERVICE DESCRIPTION

Notation	Description
SVID, UID, DID, RID	Service ID, User ID, Data ID, Registration ID corresponding with each data.
PrKuser	Private key for user
H (M, Salt)	Hash function to Message and Salt
SIP	The original cloud address of data
CIP	The current cloud address of data
DIP	The destination cloud address of data
OP	The operations on the data
TP	The TimeStamp of operation executed on the data

Since the data is an entity stored in the cloud, it needs to be registered with current cloud at the LRD when user submits it to the cloud storage. Hence, the DMaaS can be established for the user. When the user is interested in verifying the data integrity or data location, the user will be the *verifier*, and the DMaaS will be the *prover*. When the DMaaS accomplishes the integrity verification or location analysis, it will send the response to the user which is the *receiver*. In the stage of establishing DMaaS for users with their data, the user and data will be issued the following set of parameters: *User= {UID, PrKuser, RID, LID, Salt, TP}. Salt* is utilized to increase the difficulty to crack the digest message in the one-way hash function. Then, the LRD updates the corresponding registration information. The *UID* is described in the form of GUID (32-character hexadecimal string). The *RID* is used to identify user's communication device which registers the notification message service with the cloud service. *SIP* and *CIP* are the data's original and current cloud addresses, which are represented by a set of {Web server address, Public VM address, Physical address}. The LRD stores a collection of data set which is depicted as:

DMaaS set= {UID, DID, RID}

The *RID* is used to identify the corresponding *UID* and *DID*. In this set, the cloud does not own any useful information associated with the data and the user. Once the configuration accomplishes, the service for users within data will be established.

Tracking data location between user and cloud

The user sends a request to DMaaS via cloud interface. The requested parameters can be described as the follows:

Request= {Sign (H (M, Salt), PrKuser), M}
M= {UID, RID, DID, OP, SIP, CIP, DIP}

The request message consists of User ID, Registration ID, Salt, and Operation on the data. The message firstly is submitted to the CPRBAC to verify access permission. If it has a valid permission, the request is authorized to continue. If it does not, an error with error information is returned. Following that, the CPRBAC creates a responseRst (response context information of CPRBAC authorization result) as a request and fordwad it to the MS if the request is illegal. The request parameters can be described as follows:

Trigger message = {responseRst, SQL request}
responseRst = {RequestID, M, TimeStamp}

All the requests are logged, both successful and failed. The logged data can be used by both the service provider and storage users for tracing, analysis and diagnostic purposes. The MS will be triggered to search the data's accessible location in the LRD. Once obtaining the data information in LRD, the cloud generates a triggering sequence to access the data. At that time, the TS also generates a message to inform the data owner about the data access operations. When a request to move the data at the original cloud, for example, is approved, the access permission is forwarded to the MS to update information such as new data locations, operation, and TimeStamp, etc. The following is the interface which is used to request the data operations and data locations.

Interface DMaaS {
// Data Mobility as a Service interface
Boolean RegisterData (string *UID, DID, RID*, **func,….arg){}**
// UID, DID, RID are parameters used to store initial data location at cloud.
// func triggers MS function which will peform a new data regisistration processs
Boolean AuditRequest (ResponseRst respRst, func,….arg){}
// func is the Auditing function which will monitor and report functions against illegal operations on a runtime basis.
ResponseRst VerifyRequest (string *UID, DID, RID, SIP, CIP, DIP, OP*, **func,….arg){}**
// UID, DID, RID are parameters verified by CPRBAC service for permissions. OP is the data operation.
// func is the CPRBAC function which will authorize request permission
Boolean AuditRequest (ResponseRst respRst, func,….arg){}
// func is the Auditing function which will monitor and report functions against illegal operations on a runtime basis.
Void LogRequest (string *UID, DID, RID, OP, SIP, CIP, DIP, TP*, **func,….arg){}**
// UID, DID, RID, OP, SIP, CIP, DIP, TP are parameters used to record the requests.
// func is the Log function which will record requests to data
Boolean TriggerSevice(ResponseRst respRst, String SVID, func,….arg){}
// SVID is the service corresponding with request, like MS or TS
// func is the trigger function which will activate relevant service.
}

Interface MS {
// Mobility service interface
List ReadRequest (ResponseRst respRst, String SQLRequest, func,….arg){}
// respRst is the response result from CPRBAC service
// SQLRequest is the normal SQL query sentence
// func is the MS function which will perform relevant operation, like register
Boolean UpdateRequest (ResponseRst respRst, String SQLRequest, func,….arg){}
// func is the MS function which will perform relevant operation, like update data location, user.
}

Interface TS {
// Traceablity service interface
List ReadRequest (ResponseRst respRst, String SQLRequest, func….arg){}
// respRst is the response result from CPRBAC service
// SQLRequest is the normal SQL query sentence
// func is the TS function which will notify data status
String NotifyMessage (String message, func,….arg){}
// message includes UID, DID, RID, TP, SIP, CIP, DIP, data Operations used for notify user about data status
}

V. Analysis of DMaaS

A. Case study

When a customer subscribes to a service from a cloud service provider, the CSP will assign roles associated with the data for the user. The user sends a request including *UID, DID, RID* to DMaaS interfaces to register its data with the service. Once the data is successfully registered, the user will be informed with an acknowledgement message via a mobile device.

When the user needs to move data to a new cloud, he/she will send a request including *UID, DID, RID, SIP, CIP, DIP, OP* to the service. This request is verified by CPRBAC for permission and the MS for data validation. If the request is permitted, the data owner is notified new data location. If the request is denied, data owner is immediately altered with a message reporting data violations. Figure 2 shows the workflow for the DMaaS and figure 3 demonstrates the notification message via the mobile device.

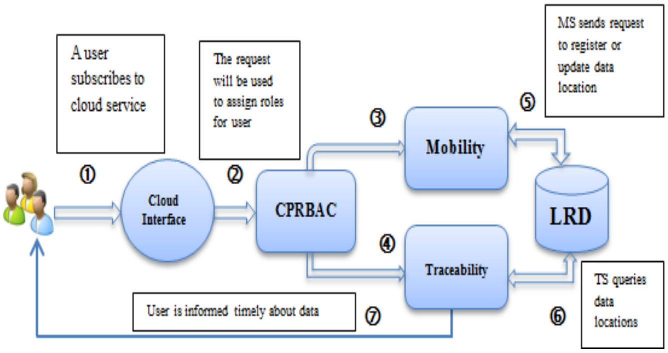
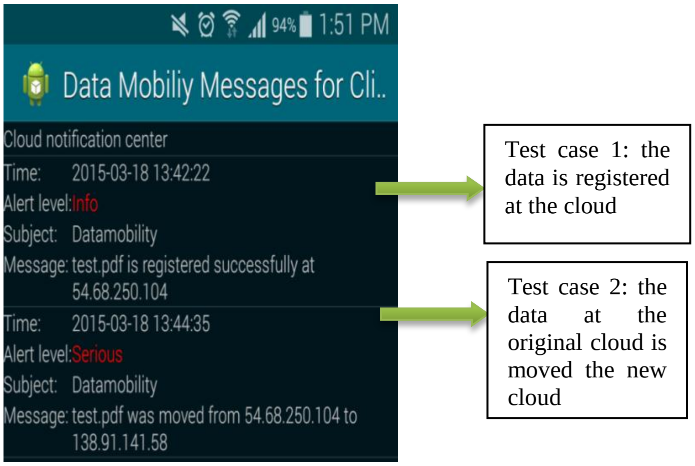

A user subscribes to cloud service
The request will be used to assign roles for user
MS sends request to register or update data location
①
②
③
④
⑤
⑥
⑦
Cloud Interface
CPRBAC
Mobility
Traceability
LRD
TS queries data locations
User is informed timely about data

Fig. 2. DMaaS workflow

B. *Implementation*

We used Amazon EC2 [10] to provide cloud resources and verify requests based on CPRBAC. The LRD was also deployed at Amazon EC2 running MySQL. Packages of the model were deployed on this instance and requests were accessed via RMI interface. Azure cloud [11] was used to send requests for moving data. Google Cloud Message (GCM) [12] was devised as TS to notify users when there are any operations on the data.

We established following test cases: the first one showed that the MS is triggered when data is first registered at the cloud; the second one notified the movement of the data to a new cloud. These requests for service are submitted through the DMaaS interface. The requests are verified at the CPRBAC and queried at the LRD by the MS. Violation messages are sent to the TS and at the same time the alert message is delivered to our mobile device.



Data Mobiliy Messages for Cli..

Cloud notification center

Time: 2015-03-18 13:42:22

Alert level: Info

Subject: Datamobility

Message: test.pdf is registered successfully at 54.68.250.104

Time: 2015-03-18 13:44:35

Alert level: Serious

Subject: Datamobility

Message: test.pdf was moved from 54.68.250.104 to 138.91.141.58

Test case 1: the data is registered at the cloud

Test case 2: the data at the original cloud is moved the new cloud

Fig. 3. Notification messages to data owner view in Galaxy Note 4

VI. Conclusions

In this paper, we addressed the emerging data mobility issues associated with data location, data protection in cloud environments. We proposed a novel DMaaS for the mobility and traceability of data and authorized access to cloud resources. We also proposed a LRD for tracing data locations and notifying messages on moving data or policy violations. Notifying owners of data violation is just for demonstration purposes, the notification mechanism can be deployed in a flexible manner as desired by cloud service providers.

Current implementations are only initial prototypes as a proof-of-concept but constitute the first step towards a comprehensive solution for a novel cloud service for data mobility. Our approach can be extended as a cloud service for a large real cloud platform due to its distributed nature and low-couple architecture design.

References

[1] L. Schubert and K. Jeffery, "Advances in clouds," Report of the Cloud Computing Expert Working Group. European Commission, 2012.

[2] Z. Minqi, Z. Rong, Z. Dadan, and Q. Weining, "Services in the Cloud Computing era: A survey," in Universal Communication Symposium (IUCS), 2010 4th International, 2010, pp. 40-46.

[3] O. Terzo, P. Ruiu, E. Bucci, and F. Xhafa, "Data as a Service (DaaS) for Sharing and Processing of Large Data Collections in the Cloud," in Complex, Intelligent, and Software Intensive Systems (CISIS), 2013 Seventh International Conference on, 2013, pp. 475-480.

[4] D. Song, E. Shi, I. Fischer, and U. Shankar, "Cloud Data Protection for the Masses," Computer, vol. 45, pp. 39-45, 2012.

[5] T. Ries, V. Fusenig, C. Vilbois, and T. Engel, "Verification of Data Location in Cloud Networking," in Fourth IEEE International Conference on Utility and Cloud Computing (UCC), 2011, pp. 439-444.

[6] A. Albeshri, C. Boyd, and J. G. Nieto, "GeoProof: Proofs of Geographic Location for Cloud Computing Environment," in 32nd International Conference on Distributed Computing Systems Workshops (ICDCSW), 2012, pp. 506-514.

[7] E. Androulaki, C. Soriente, L. Malisa, and S. Capkun, "Enforcing Location and Time-Based Access Control on Cloud-Stored Data," in IEEE 34th International Conference on Distributed Computing Systems (ICDCS), 2014, pp. 637-648.

[8] D. Popa, K. Boudaoud, M. Borda, and M. Cremene, "Mobile cloud applications and traceability," in RoEduNet International Conference 12th Edition Networking in Education and Research, 2013, pp. 1-4.

[9] A. Noman and C. Adams, "DLAS: Data Location Assurance Service for cloud computing environments," in Tenth Annual International Conference on Privacy, Security and Trust (PST), 2012, pp. 225-228.

[10] A. EC2. (2014). Amazon Elastic Compute Cloud. Available: http://aws.amazon.com/

[11] Azure. (2014). Microsoft Azure. Available: https://azure.microsoft.com

[12] GCM. (2014). Google Cloud Messaging for Android. Available: https://developer.android.com/google/gcm/index.html