

Meta-Cloud : A Cloud of Clouds

Myung-Hoon Jeon, Dong-Joo Choi, Byoung-Dai Lee*, Namgi Kim

Department of Computer Science, Kyonggi University, Suwon, Korea
jmh@kgu.ac.kr, choidongjoo@kgu.ac.kr, blee@kgu.ac.kr, ngkim@kgu.ac.kr

Abstract— As smart device hardware and software technologies have advanced, performance of devices has improved and their types and categories have diversified. In addition, as smart devices have become popular, the number of multi-devices held by a single person has increased. Following this trend, data production and management through multi-devices of a single user have increased, and various data sharing services have emerged to meet the need for data sharing across multi-devices. Much attention has been given to data sharing services utilizing a cloud computing environment. Cloud computing is a computing technology to provide IT resources (e.g., software, storage, and server) with internet technology. Data sharing services are provided by virtualizing unlimited storage space and computing resources utilizable via the cloud computing environment. Users have access to unlimited storage space for their devices and can share the same data regardless of which device is used. There are a number of pay or free cloud storage services available, but cloud service providers have their own proprietary interfaces so that users may have to install a specific application program for the use of a cloud or data management service in order to access multiple cloud storage services. In the present paper we define and propose the Meta-Cloud, which is an integrated platform for the use of multiple cloud storage services.

Keywords—Multi-devices, Data Sharing, Meta Clouds, Multiple Cloud

I. INTRODUCTION

As hardware and software technologies of smart devices have advanced, performance and the diversification of smart devices have increased. In addition, a wide variety of application programs are now available to utilize smart device for various tasks through application programs. For example, editing, such as documents, photos, and videos, which was previously possible mainly through desktop computers, can now be accomplished in various smart devices. In addition, the popularity of smart devices has resulted in multi-devices being held by a single person. Therefore, users can create various data by utilizing their own smart devices, and this has created a need for efficient data sharing over the multi-device environment.

Data sharing methods can range from a basic one where users transfer data manually to a method that uses Bluetooth, NFC, or Wi-Fi Direct. In particular, cloud storage-based data sharing services have been developed where data is stored by means of storage resources of cloud computing and data sharing via various devices. Thus, users have access to unlimited storage space for their own devices, and data are shared by any device through the cloud environment. To

increase competitiveness, cloud service providers not only provide data sharing services but also provide application services using the stored data. Therefore, when choosing cloud storage services for data sharing of multi-devices, users take service quality into consideration in terms of data sharing as well as the additional application service categories and performance. This can result in a number of cloud storage services being chosen to meet the user's needs. However, cloud service providers have their own proprietary interfaces so that users who want access to multiple cloud storage services have to install specific application programs.

In this paper we define and propose the Meta-Cloud, which is an integrated platform for the use of multiple cloud storage services.

II. RELATED WORK

As the use of multi-devices by a single user has increased so has the importance and the study of data sharing. A representative data sharing method between devices is the device to device (D2D) mode[1][2]. BlueTorrent, which is a P2P-based file sharing application utilizing Bluetooth, has been proposed, and performance improvement and feasibility from simulation results of its contents sharing process have been validated [3]. In addition, [4] proposed a system that can reduce cellular network traffic efficiently via D2D communication between neighboring devices over the 3GPP LTE cellular network. However, such D2D methods are dependent on locality between devices and not efficient for data synchronization. Therefore, a data sharing method utilizing cloud computing has been highlighted to overcome such problems. In particular, a data sharing service based on software as a service (SaaS) [5][6], one of the service models of cloud computing, provides a data sharing feature along with an application program service based on stored data. [7] proposed multimedia cloud computing that provides streaming and viewer services of media data, such as images, music, and video, thereby not only providing a storage and sharing feature but also analyzing data types that can be executed in smart devices. In addition, [8] proposed a system that can process complicated queries in portable devices, such as mobile devices, by providing database service based on BigData in the cloud storage service environment. As types and features of such cloud storage-based service have diversified, users use data sharing services through a number of cloud storages. Therefore, we propose an integrated platform that provides data sharing and application services

through a number of cloud storage services via a general interface.

III. PROPOSED PLATFORM

In this section we explain the proposed Meta-Cloud. First, the overall system structure and components are described; this is followed by an explanation of the overall service process flow and internal running flow.

A. System Architecture

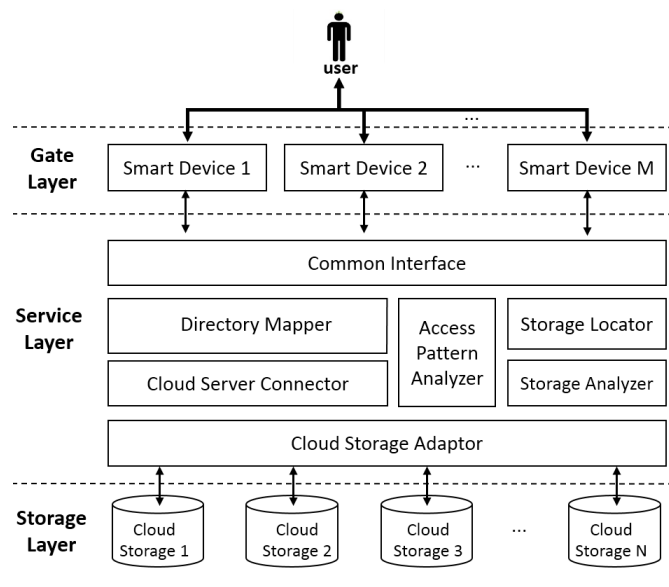


Figure 1. Meta-Cloud system architecture

Figure 1 shows the system architecture of the proposed Meta-Cloud. The system is largely divided into three layers: Gate, Storage, and Service. The Gate Layer consists of multi-devices of individual users in which various types and quantities of devices are present. The Storage Layer consists of cloud storages that store data; cloud storage provides individual supplementary services. The Service Layer is a core layer that performs the Meta-Cloud service, which plays a role of connecting the Gate Layer and Storage Layer. The function of the Service Layer is to integrate the information of the Storage Layer, thereby forming a virtual storage environment and providing this to the Gate Layer. Users who use the Meta-Cloud service can share data through the integrated virtual storage environment provided via the Service Layer without the need to know the status of the cloud storages in the Storage Layer. The role of the Service Layer is performed through the following components:

- Common Interface is an interface, which exists as a general interface form to use Meta-Cloud to communicate with smart devices directly in the Gate Layer,
- Directory Mapper provides users with information of data shared via Meta-Cloud through the virtual directory and file structure. It also maps and manages real and virtual storage information of data.

- Storage Locator is a component that determines cloud storage for user's data. Optimum cloud storage is determined via the Storage Analyzer by considering cloud storage information and size and types of resources to be stored.
- Storage Analyzer analyzes support services, storage space status, and network speed information of cloud storages interconnected to Meta-Cloud thereby delivering this to the Storage Locator.
- Cloud Server Connector—if preprocessing is required before storing data, this can be done by utilizing the computing resources in the cloud. File encryption is a typical example of such jobs.
- Access Pattern Analyzer analyzes execution history information, such as user's data access, download, storage, and service request, to summarize the user's use pattern thereby discovering intelligence, such as storage optimization and decisions on data replication storage.
- Cloud Storage Adaptor is a component that performs a role of storing data to the cloud storages. This requires scalability to include another cloud storage that provides new services.

B. Service Execution Flow

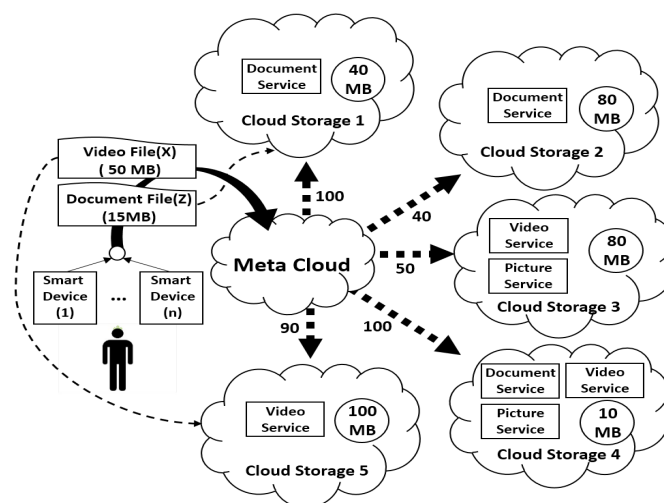


Figure 2. Service Scenario of Meta Cloud

Figure 2 shows the overall service process flow and storage decision process in Meta-Cloud. The Meta-Cloud service flow is as follows: Users transfer their data stored in their own multi-devices to Meta-Cloud to share them. Meta-Cloud analyses a type and size of transferred data and determines the optimal storage among the connected cloud storages to store the data. The most important consideration for such service flow is that a user is not involved in decisions regarding which cloud storage stores the data but only transfers data via the interface of Meta-Cloud to store them. Thus, Meta-Cloud should determine the optimal storage among the connected cloud storages. For this, the following three steps are required for the storage selection process by Meta-Cloud:

- 1) It examines whether storage can be available by checking each storage status over the cloud storages.
- 2) It checks whether suitable service for the type of data can be supported.
- 3) It checks network status and locality of the storage selected via the above two steps, thereby determining a storage that has the optimal transfer performance.

Figure 2 shows an example of the cloud storage determination process. As shown in Fig. 2, five cloud storages can be found for the individual user and these are connected with Meta-Cloud. The inner circle in the cloud storage means available storage size, while the rectangle indicates available supplementary services (e.g., editing and viewer feature for document service). In addition, numbers in the dotted arrow indicate network speed when cloud storage is accessed via the interface of Meta-Cloud by a user (for ease of explanation, in this paper we assumed that a higher number indicates a faster network speed). Thus, if an individual user tries to share through Meta-Cloud a 15-MB document file (Z) and a 50-MB video file (X) created through his/her own smart devices, the following service flow will occur. First, since the size of document file Z is 15 MB, Meta-Cloud eliminates Cloud Storage 4, which has insufficient storage space. Next, service support for the document type is checked so that Cloud Storages 3 and 5 are excluded. Finally, Meta-Cloud selects Cloud Storage 1, which has the optimal network performance for the current user, to store the document file. Similarly, if a user tries to store video file X, Meta-Cloud determines Cloud Storage 5 to store the video file. To summarize, a user can make use of data sharing services utilizing a number of cloud storages through only the Meta-Cloud interface without needing to consider the service flow.

IV. ADAPTIVE DECISION MAKING

In this section, the major functions of Meta-Cloud that improve efficiency of its performance are explained, and we describe the decision-making process to effectively conduct these functions.

A. Availability Guarantee

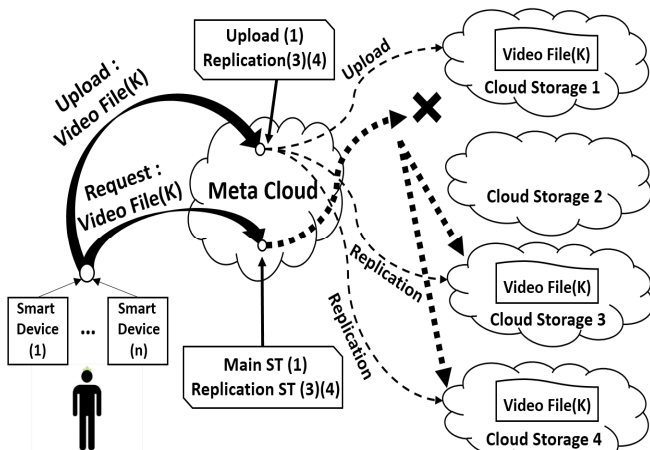


Figure 3. Data replication Mechanism for availability

Availability guarantee is one of the most important elements for determining service quality when providing a data sharing service via the cloud storages. Thus, a guarantee of efficient availability is required to improve overall service quality of data sharing.

Basically, availability guarantee in cloud computing-based data sharing services means providing users with continuous service. For example, let us assume that a user in the situation of Fig. 2 stored a document file in Cloud Storage 1. Then, if an unexpected outage occurred in the Cloud Storage 1 server, a user cannot access the document file until the server is recovered. That is, availability cannot be guaranteed. To solve this problem, we use a data replication mechanism (Fig. 3). As depicted in Fig. 3, a video file (K) transferred from a user is uploaded to Cloud Storage 1 via Meta-Cloud. Here, Meta-Cloud determines Cloud Storages 3 and 4 as suitable replication cloud storage according to the replication mechanism so that video file K is stored in both Cloud Storages 3 and 4. In fact, video file K is stored in Cloud Storages 1, 3, and 4, which means that this video file is replicated in Cloud Storages 3 and 4. Thus, if Cloud Storage 1 server experiences an outage so that a user cannot access the file via the server, Meta-Cloud guarantees the availability of video file K via Cloud Storage 3 or 4. The key to the data replication mechanism that guarantees availability is that a user is not involved with determining service availability of each cloud storage but only requests the required services through the Meta-Cloud interface.

B. Storage Optimization

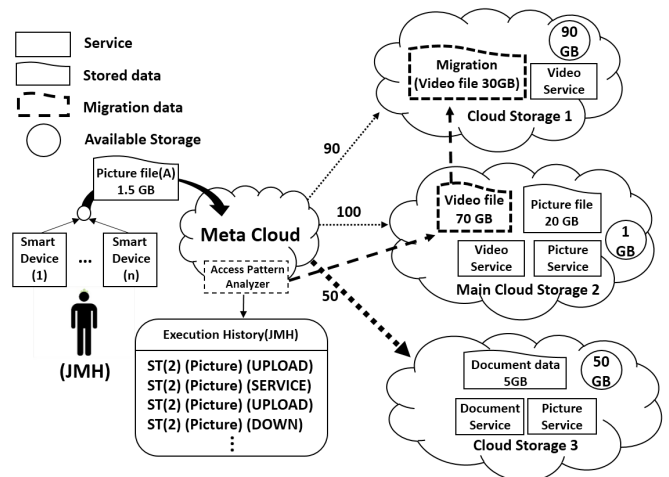


Figure 4. User access pattern based Storage Optimization

The efficient utilization and management of storage is one of the key elements that improves the data sharing service quality based on cloud computing. Thus, we provide a storage optimization feature by utilizing the user access pattern to improve efficiency of cloud storage management. A characteristic of the proposed storage optimization feature is to analyze the recent performing pattern of a user with respect to storage of high utilization frequency, thereby migrating the infrequently used data to neighboring cloud storages. Thus, a user can use cloud storage efficiently, which is utilized

frequently. Figure 4 shows an example of storage optimization in Meta-Cloud. As shown in the figure, user JMH is performing an upload to share his picture file (A) of 1.5 GB via Meta-Cloud. In the process of determining cloud storage for this upload, Cloud Storage 2 is excluded because of lack of storage space, and picture file A is stored in Cloud Storage 3. However, in this process, Cloud Storage 2, which was eliminated earlier, may be considered for storage of user data due to its insufficient data storage availability, and in turn, it can be considered as cloud storage that is utilized highly. In fact, as the execution history of recently performed by JMH is analyzed in Fig. 4, a high frequency of work performed through Cloud Storage 2 can be found. Therefore, Cloud Storage 2, which has insufficient storage space, can be considered as highly utilized storage by JMH so that analysis on its utilization frequency of internally stored data should be required. When the execution log of JMH is examined, data of work and service types performed through Cloud Storage 2 showed a concentration on pictures. However, regarding the data storage status of Cloud Storage 2, the largest percentage of file types is video so that current Cloud Storage 2 is deemed to be inefficient for data management in this case. Thus, 30 GB of video file, which has a low execution frequency among the internal data in Cloud Storage 2, is migrated to Cloud Storage 1, thereby securing more storage space for highly utilized Cloud Storage 2. In summary, the performance process for storage optimization is as follows:

1) Storage Optimization Recognition: This step recognizes the need of storage optimization, mostly through Storage Analyzer during the upload process. If Storage Analyzer recognizes a storage deficiency in the middle of the process of inspection for available storage of each Cloud Storage while uploading data of users, it delivers the information of the corresponding Cloud Storage to Access Pattern Analyzer and requests an execution history analysis.

2) Execution History Analysis: This step analyzes the access pattern of a user, which is performed by Access Pattern Analyzer. During this process, Access Pattern Analyzer identifies the user access pattern based on the recent execution history of a user. Here, the following considerations are taken into account. First, frequency of work performed by a user in the corresponding Cloud Storage is analyzed. For example, if a user does not access the Cloud Storage frequently, performing storage optimization may negatively degrade efficiency. If frequency of works performed by a user in the corresponding Cloud Storage is determined as high, Access Pattern Analyzer analyzes the data and service type performed in the corresponding Cloud Storage and delivers this to Storage Analyzer.

3) Storage State Analysis: This step is performed by Storage Analyzer. Data stored in the corresponding Cloud Storage are identified and their relationship with information delivered from Access Pattern Analyzer are analyzed. The critical elements in this process are to analyze storage distribution of stored data and utilization frequency, and to discover low-utilized data by matching information delivered from Access

Pattern Analyzer. If low-utilized data are discovered by the analysis, data migration process is performed.

4) Data Migration: This step migrates the low-utilized data in the corresponding Cloud Storage, which is performed by Storage Locator and Storage Analyzer. In this process, the decision to migrate cloud storage is performed according to the optimal storage determination method explained previously. Once cloud storage to be migrated is determined, low-utilized data are migrated, and the performance result is sent to the Directory Mapper.

In summary, the storage optimization process provides cloud storage selection and data management based on access patterns of users.

V. CONCLUSION

In this paper we proposed a service platform called Meta-Cloud by which mutually independent cloud storage services are integrated to form a virtual storage service environment. We described the overall system architecture and required components of the proposed Meta-Cloud as well as the service execution flow and the features for service quality improvement. The goal of Meta-Cloud is to provide a general interface to provide users with a virtual storage service environment. Thus, users do not need to be concerned about the storage status of application programs and data in the cloud storages but share data efficiently utilizing a number of cloud storages.

The Meta-Cloud platform proposed in this paper is in the early stages of development so that more research is needed to improve it. For example, development of its prototype is needed as well as development of a decision-making algorithm for various conditions and a definition of the interface for scalability of the Storage Layer.

ACKNOWLEDGMENT

This work (Grants No. 2013-B04, 2013-0554) was supported by the Content Convergence Software Research Center funded by the GRRC program of Gyeonggi Province, South Korea.

REFERENCES

- [1] K. Brett and B. Aazhang, "Cellular networks with an overlaid device to device network," *IEEE 42nd Asilomar Conference*, pp.1537-1541, 2008.
- [2] P. Jänis, CH. Yu, K. Doppler, CB. Ribeiro, C. Wijting, K. Hugl and V. Koivunen, "Device-to-Device Communication Underlaying Cellular Communications Systems," *IJCNIS*, pp.169-178, 2009.
- [3] (2013) The Bluetooth website. [Online]. Available: <http://www.bluetooth.com/>
- [4] K. Doppler, M. Rinne, C. Wijting, CB. Ribeiro, and K. Hugl. "Device-to-device communication as an underlay to LTE-advanced networks," *IEEE Communications Magazine*, p.42-49, 2009.
- [5] P. Mell, and T. Grance, "The NIST definition of cloud computing," *NIST special publication 800.145*, 2011.
- [6] MP. Papazoglou, "Service-oriented computing: Concepts, characteristics and directions," *Proc. WISE*, p.3-12, 2003.
- [7] W. Zhu, C. Luo, K. Wang and S. Li, "Multimedia cloud computing," *IEEE Signal Processing Magazine*, p.59-69, 2011.
- [8] D. Agrawal, S. Das and A. Abbadi, "Big data and cloud computing: current state and future opportunities," *The 14th International Conference on Extending Database Technology*, p. 530-533, 2011.



Dong-Joo Choi is a senior student at the department of Kyonggi University, Korea. He is pursuing his B.S degree. His current research interest includes mobile cloud computing, open mobile platforms, and mobile network services. After earning his B.S., he plans to start M.S. program at Kyonggi University