A Brokerage Approach for Secure Multi-Cloud Storage Resource Management

Muhammad Ihsan Haikal Sukmana[1(✉)], Kennedy Aondona Torkura[1], Sezi Dwi Sagarianti Prasetyo[2], Feng Cheng[1], and Christoph Meinel[1]

[1] Hasso Plattner Institute, University of Potsdam, Potsdam, Germany
{muhammad.sukmana,kennedy.torkura,feng.cheng,christoph.meinel}@hpi.de
[2] University of Potsdam, Potsdam, Germany
prasetyo@uni-potsdam.de

Abstract. Nowadays, more cloud customers are utilizing multiple cloud service providers (CSPs) to store their data in the cloud as it provides better data availability and service reliance than storing in the single CSP. However, there are several challenges faced by cloud customers to securely manage their cloud storage resources for cloud end-users (a user or a service) in the multi-cloud scenario, such as diverse APIs and service implementations in multiple CSP as CSP is not required to comply with cloud computing standards and multi-cloud resource management skill gap. In this paper, we present a unified multi-cloud storage resource management framework for managing cloud storage resources and their configurations for Object Storage and Identity and Access Management services following the cloud brokerage approach. We propose a unified cloud storage resource model continuing our previous work to tackle the various data and cloud access control models of cloud storage resources in multiple CSPs. Based on the unified model, we introduce a unified multi-cloud storage resource management platform to manage cloud storage resources and grant/revoke access for the cloud end-user developed for two popular public CSPs: Amazon Web Services and Google Cloud. The unified platform collects and processes information about the cloud storage resources that allows cloud customers to discover, create, delete, modify, evaluate, and monitor cloud storage resources across various CSPs.

Keywords: Multi-cloud storage · Cloud brokerage · Resource management · Access management · Object storage service · Identity and Access Management service · Cloud management platform

1 Introduction

Storage service is one of the most used cloud computing services as it provides cheaper data storage and better data availability and scalability compared to in-house data storage [13]. However, cloud storage services could still be susceptible

N. Park et al. (Eds.): SecureComm 2020, LNICST 336, pp. 102–119, 2020.
https://doi.org/10.1007/978-3-030-63095-9_6

to outage even though they guarantee up to 99.9% uptime. For example, in 2017, Amazon Web Services Simple Storage Service (AWS S3) went down for 4 h causing several web services to be unavailable and massive financial loss [15].

More cloud customers are using storage service from multiple cloud service providers (CSPs) to store their data in the cloud, or commonly known as multi-cloud storage approach [17,19]. The approach provides better data availability and service reliability compared to a single CSP usage as the data could still be accessed in case one or several CSPs are inaccessible due to outage [13,14].

Due to the cloud shared responsibility model [3], cloud customers subscribing to the storage service in a CSP are responsible for securely managing their cloud storage resources, i.e., the buckets and their stored data, the CSP credential, and the resource configurations. Cloud customers must be able to securely create, delete, and modify available cloud storage resources for the cloud end-users, e.g., users, applications, and services, ensure the resources are secure from unauthorized users and monitor resource activities across multiple CSPs.

However, there are several challenges faced by cloud customers to securely adopt and manage their storage resources across multiple CSPs. CSPs are not obligated to follow any cloud computing standards that affect each CSP to have different data models, service implementations, and APIs with other CSPs [22]. Therefore, cloud customers have to deal with the heterogeneity of the CSPs to manage their cloud resources on their own where the complexity is growing with the number of CSPs subscribed by the cloud customers [18]. Meanwhile, multi-cloud orchestration API and tools might not be sufficient to fulfill cloud customer's needs as it does not provide full CSP service functionalities.

In this paper, we present a unified multi-cloud storage resource management framework for securely managing cloud storage resources and its access for Object Storage and Identity and Access Management (IAM) services of various CSPs. Our work provides secure storage resource lifecycle management in a multi-cloud storage environment from the cloud customer's perspective for cloud end-users using the cloud brokerage approach [6].

Our contributions in this paper are as follow:

- We propose a unified cloud storage resource model built on top of the CSP's native API continuing our previous work of unified cloud access control model [21] to manage the information of cloud storage resources and its access for Amazon Web Services (AWS)[1] and Google Cloud (GC)[2].
- We develop a unified multi-cloud storage resource management platform that focuses on four resource management processes: resource discovery, resource orchestration, resource assessment, and resource monitoring.
- We introduce a unified cloud activity log format to normalize cloud activity log messages of different formats from various CSPs.

The structure of this paper is as follows: Sect. 2 presents several related works in the multi-cloud storage resource management area. Section 3 explains

[1] https://aws.amazon.com/.
[2] https://cloud.google.com/.

the overview of the multi-cloud storage approach and the challenges faced by cloud customers managing the storage resources in the multi-cloud scenario. In Sect. 4 we present our unified cloud storage resource model based on our previous work [21] to tackle various data and access control models of cloud storage resources from different CSPs. Section 5 introduces our unified multi-cloud cloud resource management platform based on our unified model that allows cloud customers to discover, create, delete, modify, monitor, and evaluate the cloud storage resources across multiple CSPs. Section 6 discusses how our unified platform solves the challenges of multi-cloud storage resource management. Finally, Sect. 7 summarizes our work and the future work of our platform.

2 Related Works

Although there have been several works regarding resource management in a multi-cloud environment, very few works are focusing on the multi-cloud storage resource management and its security area.

Hill and Humprey [7] presented a CSP vendor-agnostic cloud storage abstraction layer (CSAL) that allows an application to access Blob, Table, and Queue storage services in the multiple CSPs. It utilizes a single namespace across all storage services to maintain the metadata of each storage entity. Rafique et al. [17] introduced an adaptive middleware platform for (semi-)autonomous storage architecture management across multiple CSPs for three different scenarios: performance optimization, peak-load condition, and evolving pricing scheme. It continuously monitors the storage system's metrics that allow for identifying the changing condition of the system and optimizing the multi-cloud data placement strategy. Krotsiani and Spanoudakis [10] proposed a certification model for non-repudiation in the cloud storage services to ensure neither data owner nor CSP could deny the activities happening in the CSP. It uses a non-repudiation mechanism based on the fair multi-party non-repudiation scheme and continuous monitoring and assessment to detect the anomaly and suspicious behavior. [4] developed a multi-cloud storage broker API to provide portability and easier migration between different CSPs. It is based on a layered ontological framework to map and abstract common functionalities of object storage service.

Our work is different from the works above as we propose a unified storage resource management framework that would allow cloud customers to securely manage cloud storage resources and its access for cloud end-users in the multi-cloud environment continuing our previous work [21]. We propose a unified cloud storage resource model and a unified cloud activity log format for storing and processing the information of cloud storage resources of various formats in multiple CSPs into a single format. We implement a unified multi-cloud storage resource management platform following the cloud brokerage approach that allows for secure cloud storage resource lifecycle management across multiple CSPs.

3 Background

3.1 Multi-Cloud Storage Resource Management Overview

The usage of storage service from a single CSP is susceptible to vendor lock-in and service unavailability threats as the data could not be retrieved due to the CSP outage [16]. Cloud customers could utilize storage service from more than one CSP to resolve this issue known as the multi-cloud storage approach [17,19]. It provides better data interoperability and availability than utilizing storage service from a single CSP as the data could migrate between CSPs and still be retrieved in case one or several CSPs are unreachable [14,16]. The approach might utilize various data storage strategies by storing multiple objects across various CSPs, e.g., erasure coding, replication, or fragmentation [13,14].

When cloud customers utilize storage services from one or more CSPs, they are responsible to comply with the cloud shared responsibility model implemented by the CSPs [3]. Each CSP is responsible to operate and manage the underlying hardware components to provide the storage services and ensure that cloud customer's storage resources could not be accessed by other unauthorized cloud customers or known as cross-tenant data leakage threats [1,5,22].

Meanwhile, cloud customers are responsible to manage their storage resources across various CSPs from unauthorized users [21,22]. Cloud resource management is a process of managing and allocating available resources in the cloud for the requiring entity to fulfill its requirements and objectives [9,12]. It helps cloud customers to utilize cloud resources efficiently and securely while guaranteeing the Quality of Service for the entities. There are three entities involved in the cloud resource management process [9]:

- **Cloud Service Provider (CSP)**: The CSP manages its infrastructure to provide necessary services and its resources for its customers. It is responsible to fulfill the cloud customer's expected level of service based on the Service Level Agreement (SLA) agreed with cloud customers. We assume that the CSPs are trusted entities as they will execute the commands issued by the cloud customers using the CSP's native APIs and will not unauthorizedly access cloud customer's data.
- **Cloud customers** : Cloud customers subscribe to the CSP to use its services and resources. They are responsible to manage their cloud resources and fulfill the SLA agreed with the cloud end-user.
- **Cloud end-user**: Application, service, or a person that requires certain access to the cloud resources provided by the cloud customers to do its job.

The cloud resource management process for storage service requires cloud customers to orchestrate cloud storage resources, secure the resources from unauthorized users, and monitor its activities [22]. This includes the data uploaded to the cloud, the buckets where the data is stored, the CSP credential(s), and the resource configuration that determines who has what kind of access to the buckets and its stored data. They are also responsible to provide necessary access to the cloud storage resources for the cloud end-user following the agreed SLA between cloud customers and cloud end-users.

3.2 Security Challenges of Multi-Cloud Resource Management

The usage of the multi-cloud storage approach creates several challenges for cloud customers to securely manage their cloud storage resources across various CSPs. Each CSP utilizes various hardware and software resources to build its cloud environments without any obligation to follow any cloud computing standards available in the market [21,22]. This affects each CSP to implement its mechanisms and service implementations of the same cloud service to be different from other CSPs, such as API and data model.

Therefore, cloud customers have to deal with the heterogeneity of the CSPs to manage their cloud resources as several CSPs do not provide cloud interoperability functionality to communicate between services in multiple CSPs [13,22]. The resource management process in a multi-cloud environment requires processing the information coming from CSP's complex environment where it could be difficult to have accurate global information about the cloud resources [12]. They are expected to collect and process the information of the resources across different CSPs by themselves where the management complexity is growing with the number of CSPs subscribed by the cloud customers [18].

Cloud customers are also required to ensure that their cloud storage resources are correctly configured. The configurations of cloud storage resources could follow cloud security best practices and standards available in the market to ensure that it is secure from unauthorized users, such as the Center for Internet Security (CIS) Benchmark[3]. However, cloud customers might lack the knowledge and skill of multi-cloud management and cloud security [12]. They often require to use each CSP's management platform and API to manage the cloud storage resources in various CSPs, which could create limited visibility of cloud storage resources and its activities [3,26]. It could also make it difficult to enforce security and access control towards their cloud resources due to the heterogeneity of the security implementation from various CSPs and the loss of physical access control caused by outsourcing data storage to the cloud [2,22].

There are several multi-cloud APIs and tools available that provide cloud interoperability and multi-cloud orchestration that can be used to manage the resources in the multi-cloud scenario, such as jclouds[4] and Libclouds[5]. However, cloud customers still require to provide an abstraction layer and a unifying environment to achieve multi-cloud resource management while using the APIs and tools [26]. Also, these APIs and tools might not provide full CSP native functionality for all cloud services, e.g., jclouds does not provide user account management in IAM service or bucket storage configuration functionality.

4 Unified Cloud Storage Resource Model

We propose a unified cloud storage resource model continuing our previous work of unified cloud access control model [21] to solve the challenges of various data

[3] https://www.cisecurity.org/cis-benchmarks/.
[4] https://jclouds.apache.org/.
[5] https://libcloud.apache.org/.

models of storage resources from different CSPs due to no obligation of CSP to follow any cloud computing standard [22]. It combines the information of cloud storage resources and cloud access control models to solve the challenges of security in multi-cloud storage resource management explained in Sect. 3.

The unified cloud storage resource model helps cloud customers to manage the cloud storage resources across multiple CSPs, such as automated cloud storage resource creation. It also can be used to store the state of cloud storage resources that consists of various data and access control models from different CSPs in a single format. The cloud resource states in the unified format then could be analyzed for different use cases, e.g., discover the relationship between the resources and the entities that have access to it or check the compliance of the resources against cloud security standards and best practices.

We implement our proposed model to manage cloud storage resources available in AWS Simple Storage Service (S3)[6], AWS IAM[7], GC Storage[8], GC IAM[9], and GC Cloud Resource Manager (CRM)[10]. Our unified model could also be extended for Storage and IAM services from other CSPs. The unified cloud storage resource model consists of nine entities as can be seen in Fig. 1.

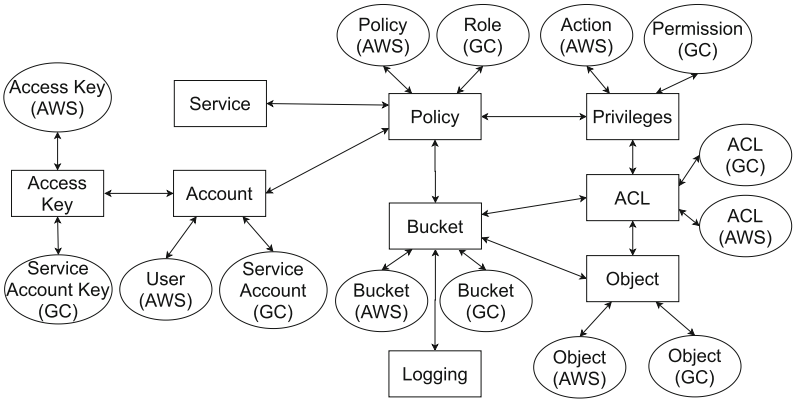

Fig. 1. Unified cloud storage resource model and its implementation on top of Object Storage and Identity and Access Management services in AWS and GC

– **Bucket**: Bucket is a logical abstraction of object storage where the objects are stored in the CSP. It represents Bucket both in AWS[11] and GC[12].

[6] https://aws.amazon.com/s3/.
[7] https://aws.amazon.com/iam/.
[8] https://cloud.google.com/storage/.
[9] https://cloud.google.com/iam/.
[10] https://cloud.google.com/resource-manager/.
[11] https://docs.aws.amazon.com/en_pv/AmazonS3/latest/dev/UsingBucket.html.
[12] https://cloud.google.com/storage/docs/json_api/v1/buckets.

- **Object**: Object is the logical abstraction of the file stored in the Bucket. It represents Object in AWS[13] and GC[14].
- **Account**: Account is the identity of an entity created in CSP's IAM service. It consists of User[15] in AWS and Service Account[16] in GC.
- **Service**: Service represents the identity of CSP service.
- **Privilege**: Privilege is the possible action/permission in the CSP's services. It consists of Action[17] in AWS and Permission in GC[18].
- **Policy**: Policy is a set of Privileges and its state (allow/deny) that regulates cloud-level access control between the entity and the cloud resource. Policy is represented as Policy[19] in AWS and Role[20] in GC. In general, there are two types of policy assignment:
  1. *IAM-level Policy*: Policy is attached to an IAM entity that allows or denies access to CSP services and its resources. In AWS, Policy can be assigned directly to User, Group, or Role. In GC, Role can be assigned to Service Account, Google account and group, G Suite domain, and cloud identity domain.
  2. *Resource-level Policy*: Policy is assigned to a resource (e.g., Bucket) and its CSP service that determines who is authorized to access the resource. In AWS, Policy can be assigned to Bucket by specifying the IAM entities or AWS service accessing it. In GC, a Role can be assigned to Service Account, Google account and group, G Suite domain, and cloud identity domain in regards to the Bucket.
- **Access Control List** (ACL): ACL is a list of access permission to buckets and/or its object that defines the entity and its type of access. It is a legacy access control mechanism that predates IAM-level access control. It represents ACL both in AWS[21] and GC[22].
- **Logging**: Logging is the logging configuration of a Bucket[23][24] where all activities of a Bucket are logged and delivered to the target Bucket.
- **Access Key**: Access Key is the credential of Account used for authentication and allowing programmatic calls to services in multiple CSPs. It contains the access key ID and secret key. The privileges of Access Key follow the Policy

[13] https://docs.aws.amazon.com/en_pv/AmazonS3/latest/dev/UsingObjects.html.
[14] https://cloud.google.com/storage/docs/json_api/v1/objects.
[15] https://docs.aws.amazon.com/en_pv/IAM/latest/UserGuide/id_users.html.
[16] https://cloud.google.com/iam/docs/service-accounts.
[17] https://docs.aws.amazon.com/en_pv/IAM/latest/UserGuide/reference_policies_actions-resources-contextkeys.html.
[18] https://cloud.google.com/storage/docs/access-control/using-iam-permissions.
[19] https://docs.aws.amazon.com/en_pv/IAM/latest/UserGuide/access_policies.html.
[20] https://cloud.google.com/iam/docs/understanding-roles.
[21] https://docs.aws.amazon.com/AmazonS3/latest/dev/acl-overview.html.
[22] https://cloud.google.com/storage/docs/access-control/lists.
[23] https://docs.aws.amazon.com/AmazonS3/latest/dev/ServerLogs.html.
[24] https://cloud.google.com/storage/docs/access-logs.

set in Account to ensure that it can only access its authorized resources. It represents Access Key[25] in AWS and Service Account Key[26] in GC.

5 Unified Multi-Cloud Storage Resource Management Platform

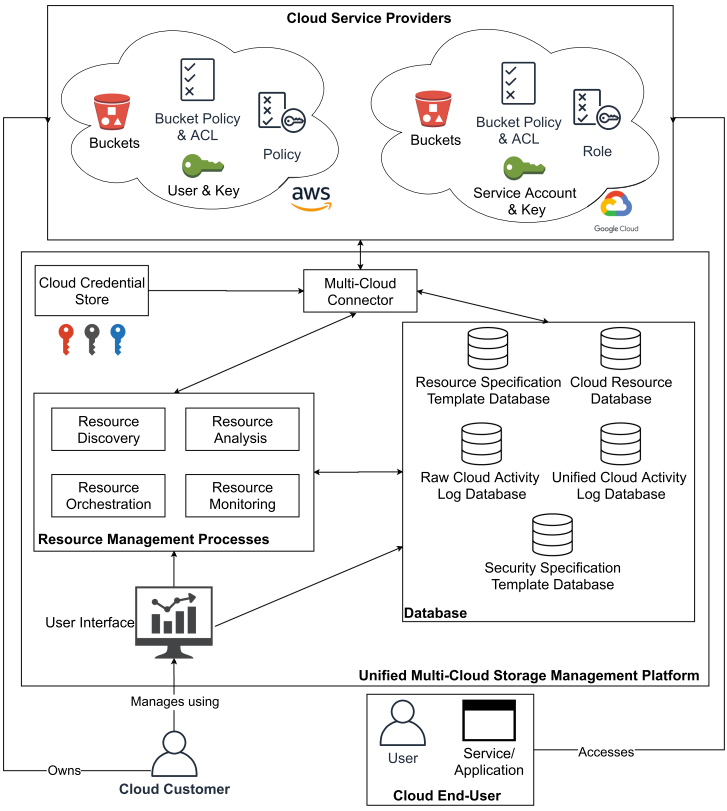

Fig. 2. Overview of unified multi-cloud storage management platform

We propose a unified multi-cloud storage resource management platform to provide cloud customers with holistic visibility and management capabilities for all cloud storage resources across multiple CSPs. It utilizes the unified cloud storage resource model explained in Sect. 4 to manage the information about cloud storage resources across multiple CSPs. Cloud customers only need to use the

[25] https://docs.aws.amazon.com/en_pv/IAM/latest/UserGuide/id_credentials_access-keys.html.

[26] https://cloud.google.com/iam/docs/creating-managing-service-account-keys.

unified platform to manage the storage resources in various CSPs instead of utilizing each CSP's management platform and API.

We chose the cloud brokerage approach [6,23] as the basis for our unified platform to manage the relationship between cloud customers, cloud storage resources in multiple CSPs, and cloud end-users. The platform provides centralized multi-cloud management as it collects, pre-processes, and stores the information on cloud storage resources with different data models in a unified format. It utilizes an abstraction layer built on top of Object Storage and IAM services APIs of various CSPs to support the multi-cloud storage resource management. This could simplify information asymmetry of cloud storage resources thus reducing the complexity of decisions taken by the cloud customers to manage the storage resources and their configurations access across multiple CSPs [6].

The unified multi-cloud storage resource management platform consists of the cloud credential store, multi-cloud-connector, and several databases as it focuses on four main resource management processes: resource discovery, resource orchestration, resource assessment, and resource monitoring. Figure 2 shows an overview of our unified multi-cloud storage management platform.

5.1 Multi-Cloud Connector

The multi-cloud connector is the gateway between our unified multi-cloud storage management platform with multiple CSPs. It provides an abstraction layer that is built on top of CSP's native APIs to ensure that the platform can access Object Storage and IAM services full native functionalities. We are utilizing the APIs of AWS S3, AWS IAM, GC Storage, GC IAM, and GC CRM services to access the cloud storage resources. All commands made by the unified platform are translated into CSP's native API commands by the connector. It also downloads the cloud activity logs generated by multiple CSPs that will be explained in Sect. 5.6.

5.2 Cloud Credential Storage

Cloud credential storage securely stores an Access Key for each CSP to allow the unified platform to access Storage and IAM services across various CSPs. The key is generated from the Account with adequate privileges to list, create, modify, and delete cloud storage resources where it can only be accessed via the platform. When the unified management platform issues a request to a CSP, the multi-cloud connector first requests the required Access Key to cloud credential storage before sending the request to the CSP.

5.3 Resource Discovery Process

Resource discovery is the process to detect and register all available created resources for each service in the CSP [11]. The unified multi-cloud storage management platform provides resource discovery by automatically gathering the

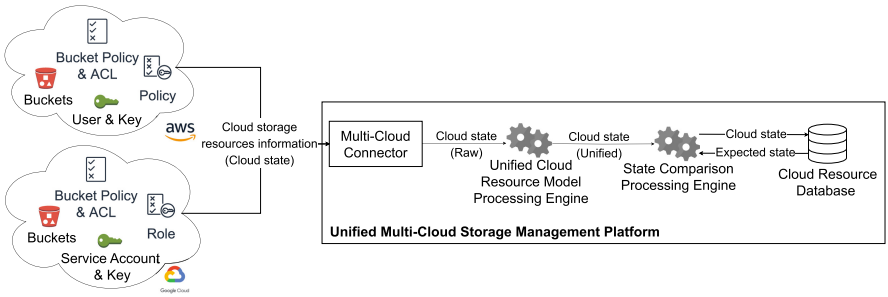

Fig. 3. Overview of resource discovery process

information of all cloud storage resources and their configurations across multiple CSPs in a single format. It runs periodically in the background to monitor any changes in the cloud resources. When it runs for the first time, the platform does not need to have prior knowledge of cloud storage resources owned by cloud customers. Figure 3 shows an overview of the resource discovery process.

Multi-cloud connector first sends a request to each CSP service to retrieve the information of all available cloud storage resources. Depending on the CSP's API capabilities, the information about the cloud resources and their configurations, e.g. name, type, Policy, and ACL, are then retrieved by the multi-cloud connector. Cloud storage resource information that could not be collected during the discovery process due to the limitation of the CSP's API could be added manually later by the cloud customers, e.g., the secret key of Access Key is only available once it is newly generated.

The cloud storage resource's raw information is then processed by the Unified Cloud Resource Model Processing engine to parse the information with different data models from various CSPs to our unified cloud storage resource model as explained in Sect. 4. The processed cloud storage information is then stored in the Cloud Resource database. An example of unified cloud storage information:


```
{
   "name":"exampleBucket",
   "type":"Bucket",
   "csp":"AWS",
   "creationDate":"2019-01-02T21:27:04.000+0000",
   "location":"eu-central-1:Frankfurt",
   "bucketConfiguration":{
      "logging":{
         "enabled":false
      },
      "accessors":[
         {
            "name":"TestUser",
            "effect":"Allow",
            "type":"ACL",
```

```
            "entity":"User Grantee",
            "privileges":[
               "s3:ListBucket",
               "s3:PutObject",
               "s3:DeleteObject"
            ]
         }
      ]
   },
   "deleted":false
}
```


We incorporate the state transition model into our resource discovery process to track the changes made into the cloud storage resource [25]. When the resource discovery process runs for the first time, the cloud storage resource information in a unified format stored in the Cloud Resource database is regarded as the **expected state**. After the initial resource discovery process, the information of storage resource is then regarded as the **cloud state**. These states are then compared using the State Comparison Processing engine. If the states are different, cloud customers could decide whether to store the cloud state in the Cloud Resource database as the expected state or retain the expected state by reversing any changes happening in the storage resources across multiple CSPs.

Using the information of cloud storage resources and their configurations stored in the unified format, cloud customers could then associate the cloud storage resources with the information of the cloud end-users. They could also maintain a consistent and accurate global state of cloud resources across multiple CSPs instead of manually list the created cloud resources and their configurations of each service in each CSP using its management dashboard or API.

5.4 Resource Orchestration Process

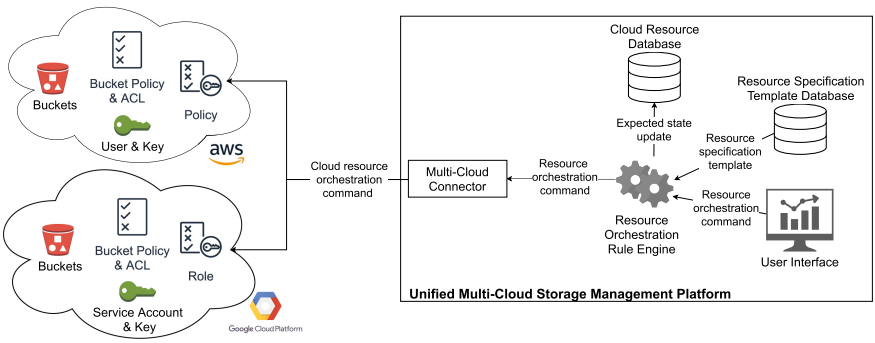

Fig. 4. Overview of resource orchestration process

Resource orchestration is the process of allocating the resources in the CSP to fulfill the requirements of cloud end-users. We follow the unified cloud storage resource model from Sect. 4 to help cloud customers create, delete, and modify cloud storage resources and their configurations for the cloud end-users. Figure 4 shows an overview of the resource orchestration process.

Cloud customers could create, delete, and modify the storage resource in one or multiple CSPs by providing necessary cloud storage resource specification using the user interface to generate resource orchestration command. They could also create a resource specification template for cloud resources and their configurations stored in the Resource Specification Template database. The template is used to automatically create and configure necessary cloud resources for the cloud end-users to reduce the possibility of misconfiguration due to human error.

Cloud customer's resource orchestration command and resource specification template are then processed by Resource Orchestration Rule engine to consolidate the resource orchestration command. It then updates the expected state with the information of created, deleted, or modified cloud storage resources in the Cloud Resource database. The resource orchestration command is then translated by the multi-cloud connector to the specific CSP's API commands.

Cloud customers should follow the least privilege principle, privilege separation concept, or cloud security best practices and standards while orchestrating cloud resources for the cloud end-users [21]. This is to ensure the cloud end-users only have limited access to the authorized cloud resources following their roles or responsibilities, thus avoiding insider threat or over-privileged access.

5.5 Resource Assessment Process

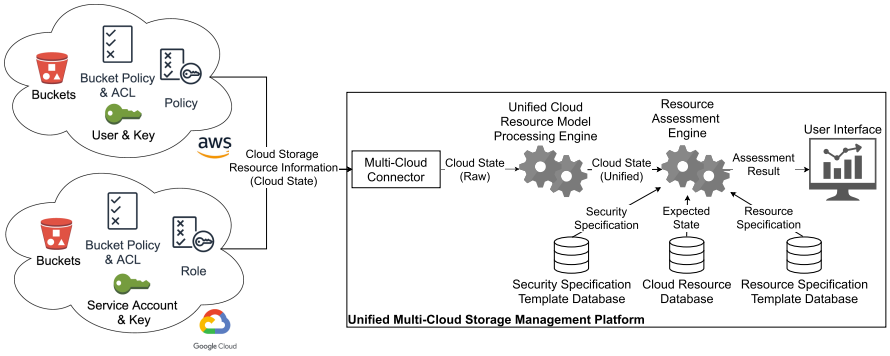

Fig. 5. Overview of resource assessment process

Resource assessment is the process of evaluating the cloud resources against the specifications set by cloud customers. It ensures that the resources are correctly

and securely configured that could only be accessed by its authorized cloud end-users [24]. Figure 5 shows an overview of the resource assessment process.

Our method for resource assessment is as follows: the raw information of cloud storage resources and their configurations, or cloud state, is first retrieved periodically and parsed with the Unified Cloud Resource Model Processing engine to follow our unified format. The Resource Assessment engine then compares the unified cloud state with the expected state stored in the Cloud Resource database to detect if there are any unauthorized modifications [25].

The Resource Assessment engine also evaluates the unified cloud state and the expected state against the security specifications and the resource specifications that are fetc.hed from the Security Specification Template and the Resource Specification databases, respectively. The specifications could be derived from cloud computing best security practices or recommendations, such as the Center for Internet Security's AWS benchmark[27]. It could also be derived following the cloud end-users' requirements to ensure the cloud end-users could only access its authorized cloud resources with limited actions.

Finally, the Resource Assessment engine will generate the assessment result for the cloud customers. If there are unauthorized modifications to the cloud state or the cloud storage resource configurations do not comply with the security and resource specifications, the assessment result will include the violations against the security and resource specifications and recommended actions to be taken to address the violations. Cloud customers could take necessary actions to improve the cloud storage resources' configurations to ensure that the resources are secure and can only be accessed by authorized cloud end-users.

5.6 Resource Monitoring Process

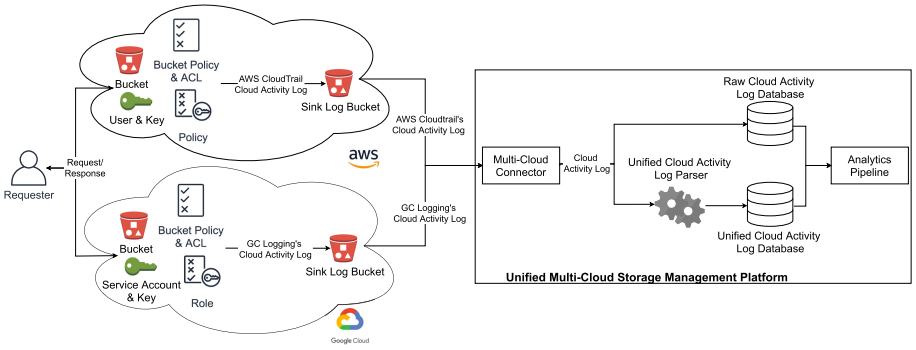

Fig. 6. Overview of resource monitoring process overview

Resource monitoring is a process of monitoring the usage and the activities of the resources in the CSP. As cloud customers outsource their files to the cloud,

[27] https://www.cisecurity.org/benchmark/amazon_web_services/.

they lose the full control of files as it could be accessed by anyone. Unauthorized users and authorized cloud end-users could directly interact with the Buckets where the files are stored using the CSP management dashboard, API request using CSP credential or Access Key, signed URL generated from the Access Key, or unauthenticated URL of Buckets and the files. The information provided by the local system owned by the cloud customers might not be enough to give full oversight of cloud activities, therefore they require a new source of information that informs the activities of cloud storage resources [2].

Therefore, we collect cloud activity logs generated by AWS CloudTrail[28] and Google Cloud Logging[29] to monitor the activities happening in cloud storage resources across multiple CSPs. The log contains the detailed information of all activities happening in the CSP's services, e.g., incoming API requests to the cloud resources and its responses in Object Storage service.

However, there are several challenges in processing cloud activity log from different CSPs. Although technically these services are already logging the activities happening in the CSP, the log messages can only be viewed and processed using the CSP's logging and monitoring service. It also requires the cloud customers to actively store or retrieve the log as it may be deleted after a certain period of time[30]. Each CSP has its log format structure and information quality [20], for example, AWS CloudTrail provides more information with better data structure's consistency compared to GC Logging. Cloud customers would be responsible to actively retrieve and process the cloud activity logs while dealing with different cloud logs from various CSPs to gain necessary information about the activities happening in the cloud storage resources.

The resource monitoring process follows the data warehouse method [8], which consists of extraction, transformation, and loading (ETL) steps, to transform semi-structured data of cloud activity log files provided in JSON format to structured data. We propose a unified cloud activity log format to normalize different log formats from various CSPs to a single format. We first select the necessary information needed from the available cloud activity log fields. We then normalize the value that are in different formats or could contain information for multiple log fields. Finally, we combine the information from cloud activity log files from multiple CSPs to give an overview of the events happening to the cloud storage resources in multiple CSPs [20]. Our proposed unified cloud activity log format can be seen in Table 1.

Our method for resource monitoring is as follows: Cloud activity log files are delivered into a specific sink Bucket that provides inexpensive and long-term storage for the log files. Depending on the CSP, the cloud activity log file could be delivered to the Bucket every 5 min[31] up to one hour[32]. Once the cloud log

[28] https://aws.amazon.com/cloudtrail/.
[29] https://cloud.google.com/logging, formerly Google Cloud Stackdriver.
[30] https://docs.aws.amazon.com/awscloudtrail/latest/userguide/view-cloudtrail-events.html.
[31] https://aws.amazon.com/cloudtrail/faqs.
[32] https://cloud.google.com/logging/docs/export/using_exported_logs.

Table 1. Unified cloud activity log format and the parsing from AWS CloudTrail and GC Logging

Unified Cloud Activity Log	AWS CloudTrail	GC Logging	Description
eventId	eventID	–	Event identifier
timestamp	eventTime	timestamp	Event timestamp
csp	"AWS"	"GC"	CSP type
service	eventSource	protoPayload.serviceName	CSP service name
resourceName	requestParameters	protoPayload.request Parameter or protoPayload. resourceName	Resource name
resourceType	requestParameters	protoPayload.request	Resource type
resourceLocation	awsRegion	resource.label.location	Resource location
method	eventName	protoPayload.methodName	Request method
ipAddress	sourceIPAddress	protoPayload.request Metadata.callerIP	Requester IP address
userAgent	userAgent	protoPayload.request Metadata.caller SuppliedAgent	Requester user agent
responseCode	errorCode	protoPayload.status.code	Response status code
responseMessage	errorMessage	protoPayload.status.message	Response message
requesterCredential	userIdentity	protoPayload.authentication Info.principalEmail	Requester identity

file has been delivered to the Bucket, the multi-cloud connector then downloads the log file to our resource management platform.

After the cloud activity log files have been downloaded, it is then stored into Raw Cloud Log Activity database while it is processed by the Unified Cloud Activity Log Parser to parse cloud activity log files into our unified log format and store it in Unified Cloud Log Activity database. Finally, the raw and unified cloud log messages are then pushed into the analytics pipeline for further processing. Figure 6 shows an overview of the resource monitoring process.

6 Discussion

Our unified multi-cloud storage resource management framework could solve the security challenges of managing cloud storage resources across multiple CSPs faced by the cloud customers as explained in Sect. 3.2.

The unified cloud storage resource model helps to normalize various data and cloud access control models of storage resources from different CSPs. We focused on developing our unified model on the storage and IAM services of AWS and GC as both CSPs employ quite a similar cloud access control model following role-based access control, which is useful for associating cloud storage resources with cloud end-users. Our proposed model differs from our previous work of unified cloud access control model [21] as it includes more cloud storage resources types and their configurations that could be utilized for various multi-cloud management strategies, e.g., cloud brokerage [6] or cloud federation [12].

Our unified multi-cloud storage resource management platform provides holistic visibility and secure multi-cloud storage resource management. Our abstraction layer for multiple CSPs implemented in our unified platform is built on top of CSP's native APIs to ensure that the unified platform can access the full functionality of the services provided by the CSP. This is different from the multi-cloud APIs where it provides abstraction by focusing only on the common functionalities and data structure of the CSP's APIs.

The unified platform allows cloud customers to automatically discover created cloud storage resources and orchestrate necessary cloud storage resources for cloud end-users to ensure that the cloud storage resources are not misconfigured due to human error. The cloud resources are also evaluated periodically against cloud computing's security best practices and standards and cloud customer's system requirements to make the resources are secure and accessible only for authorized cloud end-users.

We chose to monitor the cloud storage resources using the cloud activity log instead of the storage event log used in our previous work [20]. This is because the cloud activity log is not limited only to the Bucket and Object operations in the Object Storage service but also other services of the CSPs. Our unified cloud activity log format could be used to normalize cloud activity log files that have semi-structured and complex data in nested JSON format to be simplified and structured data that can be used for monitoring the activities in cloud storage resources. The proposed log format could also be used for different purposes, e.g., cloud forensic and security analytics.

Our unified platform is not as sophisticated as multi-cloud orchestration services available in the market where it provides Infrastructure-as-a-Code abstraction layer where cloud infrastructures could be defined using a human-readable configuration template, such as Terraform[33] or Chef[34]. However, our unified platform focuses on resource discovery, resource assessment, and resource monitoring processes that are not available in multi-cloud orchestration services.

7 Conclusion and Future Works

In the past few years, more cloud customers utilize Object Storage service from multiple CSPs to store their data to provide better data availability. However, cloud customers face several challenges of securely managing their cloud storage resources across different CSPs for the cloud end-users. In this paper, we propose a unified multi-cloud storage resource management framework that allows cloud customers to discover, create, delete, modify, evaluate, and monitor cloud storage resources in various CSPs. We introduce a unified cloud storage resource model that continues our previous model to tackle different data models of various CSPs to determine the state of cloud storage resources. We develop a unified multi-cloud storage resource management platform that collects, pre-processes, stores, and manages the information on cloud storage resources and their configurations

[33] https://www.terraform.io/.
[34] https://www.chef.io.

centrally across multiple CSPs. Our unified platform follows the cloud brokerage approach that will help cloud customers to manage cloud storage resources used by the cloud end-users. We also propose a unified cloud activity log format implemented in our platform to normalize cloud activity log messages of different formats from various CSPs.

We are currently researching various security analytics scenarios in a multi-cloud storage environment to ensure the cloud storage resources are secure, such as the correlation process using cloud activity log and storage event log. We are also extending our unified platform to support different resource types in other CSPs, e.g., virtual machine or container in Microsoft Azure and Openstack.

References

1. Amazon Web Services: Shared responsibility model. https://aws.amazon.com/compliance/shared-responsibility-model/ (2020). (Accessed 14 July 2020)
2. Bohli, J.M., Gruschka, N., Jensen, M., Iacono, L.L., Marnau, N.: Security and privacy-enhancing multicloud architectures. IEEE Trans. Dependable Secure Comput. **10**(4), 212–224 (2013)
3. Cloud Security Alliance: Top threats to cloud computing: The egregious 11 (2019). https://cloudsecurityalliance.org/download/artifacts/top-threats-to-cloud-computing-egregious-eleven/
4. Elango, D.M., Fowley, F., Pahl, C.: An ontology-based architecture for an adaptable cloud storage broker. In: Mann, Z.Á., Stolz, V. (eds.) ESOCC 2017. CCIS, vol. 824, pp. 86–101. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-79090-9_6
5. Factor, M., et al.: Secure logical isolation for multi-tenancy in cloud storage. In: 2013 IEEE 29th Symposium on Mass Storage Systems and Technologies (MSST), pp. 1–5. IEEE (2013)
6. Heilig, L., Lalla-Ruiz, E., Voß, S.: A cloud brokerage approach for solving the resource management problem in multi-cloud environments. Comput. Ind. Eng. **95**, 16–26 (2016)
7. Hill, Z., Humphrey, M.: Csal: a cloud storage abstraction layer to enable portable cloud applications. In: 2010 IEEE Second International Conference on Cloud Computing Technology and Science, pp. 504–511. IEEE (2010)
8. Hu, H., Wen, Y., Chua, T.S., Li, X.: Toward scalable systems for big data analytics: a technology tutorial. IEEE Access **2**, 652–687 (2014)
9. Jennings, B., Stadler, R.: Resource management in clouds: survey and research challenges. J. Netw. Syst. Manage. **23**(3), 567–619 (2015)
10. Krotsiani, M., Spanoudakis, G.: Continuous certification of non-repudiation in cloud storage services. In: 2014 IEEE 13th International Conference on Trust, Security and Privacy in Computing and Communications, pp. 921–928. IEEE (2014)
11. Lee, C.A.: Cloud federation management and beyond: requirements, relevant standards, and gaps. IEEE Cloud Comput. **3**(1), 42–49 (2016)
12. Liaqat, M., et al.: Federated cloud resource management: review and discussion. J. Netw. Comput. Appl. **77**, 87–105 (2017)
13. Mansouri, Y., Toosi, A.N., Buyya, R.: Data storage management in cloud environments: taxonomy, survey, and future directions. ACM Comput. Surv. (CSUR) **50**(6), 91 (2018)

14. Nachiappan, R., Javadi, B., Calheiros, R.N., Matawie, K.M.: Cloud storage reliability for big data applications: a state of the art survey. J. Netw. Comput. Appl. **97**, 35–47 (2017)
15. Newton, C.: How a typo took down s3, the backbone of the internet - the verge. https://www.theverge.com/2017/3/2/14792442/amazon-s3-outage-cause-typo-internet-server (2017). (Accessed on 7 August 2020)
16. Petcu, D.: Multi-cloud: expectations and current approaches. In: Proceedings of the 2013 international workshop on Multi-cloud applications and federated clouds, pp. 1–6 (2013)
17. Rafique, A., Van Landuyt, D., Reniers, V., Joosen, W.: Towards an adaptive middleware for efficient multi-cloud data storage. In: Proceedings of the 4th Workshop on CrossCloud Infrastructures & Platforms, pp. 1–6 (2017)
18. Raj, P., Raman, A.: Multi-cloud management: technologies, tools, and techniques. Software-Defined Cloud Centers. CCN, pp. 219–240. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-78637-7_10
19. Schnjakin, M., Korsch, D., Schoenberg, M., Meinel, C.: Implementation of a secure and reliable storage above the untrusted clouds. In: Computer Science & Education (ICCSE), 2013 8th International Conference on, pp. 347–353. IEEE (2013)
20. Sukmana, M.I., Torkura, K.A., Cheng, F., Meinel, C., Graupner, H.: Unified logging system for monitoring multiple cloud storage providers in cloud storage broker. In: 2018 International Conference on Information Networking (ICOIN), pp. 44–49. IEEE (2018)
21. Sukmana, M.I., Torkura, K.A., Graupner, H., Cheng, F., Meinel, C.: Unified cloud access control model for cloud storage broker. In: 2019 International Conference on Information Networking (ICOIN), pp. 60–65. IEEE (2019)
22. Takabi, H., Joshi, J.B., Ahn, G.J.: Security and privacy challenges in cloud computing environments. IEEE Security & Privacy **8**(6), 24–31 (2010)
23. Toosi, A.N., Calheiros, R.N., Buyya, R.: Interconnected cloud computing environments: challenges, taxonomy, and survey. ACM Comput. Surv. (CSUR) **47**(1), 1–47 (2014)
24. Torkura, K.A., Sukmana, M.I., Cheng, F., Meinel, C.: Slingshot-automated threat detection and incident response in multi cloud storage systems. In: 2019 IEEE 18th International Symposium on Network Computing and Applications (NCA), pp. 1–5. IEEE (2019)
25. Torkura, K.A., Sukmana, M.I., Strauss, T., Graupner, H., Cheng, F., Meinel, C.: Csbauditor: proactive security risk analysis for cloud storage broker systems. In: 2018 IEEE 17th International Symposium on Network Computing and Applications (NCA), pp. 1–10. IEEE (2018)
26. Varghese, B., Buyya, R.: Next generation cloud computing: new trends and research directions. Future Gener. Comput. Syst. **79**, 849–861 (2018)