IOSR Journal of Engineering (IOSRJEN) **www.iosrjen.org**
ISSN (e): 2250-3021, ISSN (p): 2278-8719
Vol. 08, Issue 8 (August. 2018), ||V (IV) || 50-56

Analysis Of Secured Storage And Secured Dataflow In Cloudiot (Storage And Dataflow Management In Cloudiot)

M. Manoranjitham,
Assistant Professor, Department Of Computer Science And Engineering,Apollo Engineering College, Chennai, India.
Corresponding Author: M. Manoranjitham,

Abstract: The integration of cloud and Internet of Things (IoT), named CloudIoT, has been considered as an enabler for many different applications.However the security issue is one main concern that some organizations hesitate to adopt such technologies while some just ignore the security issue while integrating the CloudIoT into their business. Therefore, given the cloud-resource providers and IoT devices, how to evaluate their security level becomes an important issue to promote the adoption of CloudIoT as well as reduce the business security risks. To solve this problem, considering the importance of the business data in CloudIoT, we develop an end-to-end security assessment framework based on software defined network (SDN) to evaluate the security level for the given CloudIoT offering. Specially, in order to simplify the network controls and focus on the analysis about the data flow through CloudIoT, we develop a three-layer framework by integrating SDN and CloudIoT, which consists of different indicators to describe its security features. Then, the interviews from industries are carried out to understand the importance of these features for the overall security. Furthermore, given the relevant evidences from the CloudIoT offering, the Google Brillo and Microsoft Azure IoT Suite, our framework can effectively evaluate the security level which can help the consumers in selecting the CloudIot.

Keywords:Storage management system; Dataflow management system; Secured Storage.

Date of Submission: 06-08-2018 Date of acceptance: 23-08-2018

I. INTRODUCTION

[1]The Internet of Things (IoT) has recently emerged as a novel networking paradigm to connect a large amount of smart objects for data sharing and exchanging, so that we can measure, communicate, and interact with the real physical world.On the other hand, cloud computing has been accepted as a cost-effective approach for providing high performance computing and virtually unlimited storage resource. Therefore, the integration of these two complementary technologies, the sensor-capability from IoT and the computing-capability from Cloud, has been accepted as a novel IT paradigm, named CloudIoT, for many different applications, including smart grid, smart cities, healthcare,video surveillance, environmental monitoring, etc. Actually, the CloudIoT is playing an important role for the current IT system, especially for the critical infrastructure. Considering the fact that information security has become increasing important for current IT environment while we can observe many cyber attacks these years, resulting in to a major lose for around million populations, the security of CloudIoT is no doubt an urgent issue for both industry and academic. Most of the industries and even government organizations are using cloud storage to store the employee , customer and citizen's personal data. In that case the security to the data is very important . While using the cloud the option given can be of three they are public cloud, private cloud and the hybrid cloud. Some organizations fail to notice the security issues of the Cloud storage or some volunteer to ignore or avoid the security issues. They should monitor and be worried about the data stored in cloud. So it is essential to promote the adoption of CloudIot and reduce the Business Security Risk. And to solve this problem considering the importance of business data in CloudIot we developed an end to end Security frame work based on the network SDN and CloudIot . This SDN evaluate the security levels for the given CloudIot offerings. Especially in order to simplify the network controls and focus on the analysis about dataflow through CloudIot. We developed a three layer frame work by integrating the SDN and CloudIot which consist of different indicators to describe its security features. Then, the interviews from industries are carried out to understand the importance of these features for the overall security. Furthermore, given the relevant evidences from the CloudIoT offering, the Google Brillo and Microsoft Azure IoT Suite, our framework can effectively evaluate the security level which can help the consumers in selecting the CloudIot.

II. EASE OF USE

A. Screening Tool

A screening tool can be deployed to scan computer file systems, server storage, and inspect outbound network traffic. The tool searches for the occurrences of plaintext sensitive data in the content of files or network traffic. It alerts users and administrators of the identified data exposure vulnerabilities.

For example, an organization's mail server can inspect the content of outbound email messages searching for sensitive data appearing in unencrypted messages. Data leak detection differs from the anti-virus (AV) scanning (e.g., scanning file systems for malware signatures) or the network intrusion detection systems (NIDS) (e.g., scanning traffic payload for malicious patterns). AV and NIDS typically employ automata-based string matching, which match static or regular patterns.

III. ANALYSIS OF DISTRIBUTED FILE SYSTEM

In Cloud storage for reliability of the stored data we go for distributed systems .The data are distributed and stored in two servers . if one server fails to retrieve then the data can be fetched from the chunk server.

B. DISTRIBUTED FILE SYSTEM

Cloud computing is the computing style providing IT related function with service form. Cloud computing is largely divided into 3 classes such as SaaS(Software as a Service), PaaS(Platform as a Service), IaaS(Infrastructure as a Service). In order to provide data safety stored in cloud, at first, safe storing place in IaaS class should be provided. The Service providing service place in cloud computing environment is called cloud storage service. Base technology to provide this cloud storage service is distributed file system. Let's look into GFS and HDFS which are mostly used in distributed file system.

GFS: GFS(Google File System) is the distributed file system made to provide cloud service in Google. GFS is consisted of client, master server and chunk server, and roles of each object are as follows. Client: this provides self interface similar to file system interface and communicate with master server and chunk server on behalf of application program[13].

Master server: this manages meta-data of file system such as name space, access control information, mapping information between file and chunk, chunk location information, etc. These meta-data are stored in the memory of master server and quickly inform the location of data to client. Also, they control overall system operation such as creating chunk copy, adjusting number of copies, returning unused store space, chunk server health check, etc.

Chunk server: chunk server manages chunk which is stored data unit and supports input and output of data requested by client. Chunk server regularly reports Heartbeat message to master server. Also, this detects data error using checksum and deletes error detected chunk. How GFS is operated can be fully supposed by component role of prior GFS. When storing file, client sends file information to be stored by own to master server and master server sends chunk server location and handle of actually storing file to client. Afterward, the client divides own data into chunk with fixed size. And then it sends divided chunk to chunk server. When reading file, client searches own data in master server and receives chunk server location where these data are stored. Afterward, it receives chunk through communication with chunk server and can have original data by summing these.

C. REQUIREMENTS OF SAFE SEARCH AND SHARING IN CLOUD ENVIRONMENT

The following requirements should be met for safe search and sharing to be secured under Cloud storage environment.

Confidentiality: Data transmitted between remote data server and client terminal should be identifiable only by proper persons.

Search speed: The client who has limited system resources should be able to quickly search documents including word files from documents stored in cloud storage systems.

Traffic efficiency: Communication volume should be small for the energy efficiency between client and server, and efficiency of network resources.

Calculation efficiency: Calculation efficiency should be provided for index generation and execution of search, and for sharing data with other users safely

D. WRITING SCENARIO

In suggested method considering cloud storage structure, encode index possible for sharing and search is stored in master server. User encodes keyword necessary at data search later to be able to search by oneself only and

sends this to master server. Master server sends chunk information for data storage to user and user divides data into chunks and stores in designated chunk server.

E. READING SCENARIO

User sends trapdoor which is able to search data without exposing keyword information to master server. Master server searches data having keyword by using trapdoor in encoded index. And then it sends chunk information having corresponding data to user. User acquires data by summing these after receiving each chunk from chunk server where is storing data.

F. SHARING SCENARIO

In order to share data with desired user and in order for shared user to freely share data with another user, re-encryption should be done for shared user to be able to search encoded index only. The user acquired index of sharing data can always search corresponding data by keyword and download them.

IV. ARCHITECTURE OF SECURED CLOUD STORAGE

At its core, the architecture consists of three components: a data processor (DP), that processes data before it is sent to the cloud; a data verifier (DV), that checks whether the data in the cloud has been tampered with; and a token generator (TG), that generates tokens that enable the cloud storage provider to retrieve segments of customer data; and a credential generator that implements an access control policy by issuing credentials to the various parties in the system (these credentials will enable the parties to decrypt encrypted files according to the policy).

A. CONSUMER ARCHITECTURE

Alice's data processor prepares the data before sending it to the cloud. Bob asks Alice for permission to search for a keyword. Alice's token and credential generators send a token for the keyword and a credential back to Bob. Bob sends the token to the cloud. The cloud uses the token to find the appropriate encrypted documents and returns them to Bob. At any point in time, Alice's data verifier can verify the integrity of the data.

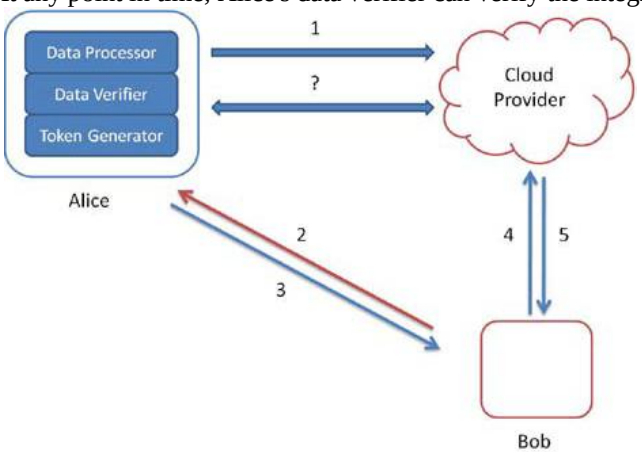

Fig 1. Consumer Architecture

B. ENTERPRISE ARCHITECTURE

Each MegaCorp and PartnerCorp employee receives a credential. MegaCorp employees send their data to the dedicated machine. the latter processes the data using the data processor before sending it to the cloud. the PartnerCorp employee sends a keyword to MegaCorp's dedicated machine. the dedicated machine returns a token. the PartnerCorp employee sends the token to the cloud. the cloud uses the token to find the appropriate encrypted documents and returns them to the employee. More precisely, in this case the dedicated machines only run data verifiers, token generators and credential generators while the data processing is distributed to each employee.

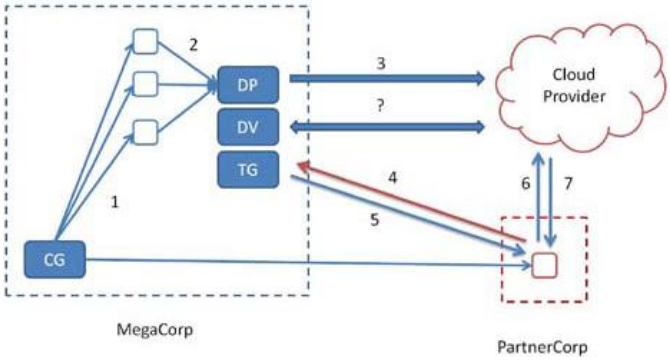

Fig 2. Enterprise Architecture

V. RELATED WORK

C. SECURE INDEX MANAGEMENT SCHEME ON CLOUD STORAGE ENVIRONMENT

Keeping pace with the increase of digital information in use, Cloud storage is in service, which can store one's data from distance through network and various devices and easy to access. Unlike the existing removable storage necessary in order to carry data, it is used many users because it has no limit of memory capacity and no need to carry storage medium. As many users save a great volume of date in Cloud storage, its reliability has become a focus of issue. To protect it from unethical managers and attackers, researches are being conducted on application of a variety of cryptography systems such as searchable encryption and proxy re-encryption to Cloud storage system. However, existing searchable encryption technology is inconvenient in the cloud storage environment in which the user uploads data in person, and those data are shared with others, whenever it is necessary to do, and those with whom data are shared change frequently. In this we propose a searchable re-encryption scheme by which user can share data with others safely by generating searchable encryption index, and re-encrypting it.

D. ON SCHEMA MATCHING WITH OPAQUE COULMN AND DATAFLOW VALUES

The schema matching problem at the most basic level refers to the problem of mapping schema elements (for example, columns in a relational database schema) in one information repository to corresponding elements in a second repository. While schema matching has always been a problematic and interesting aspect of information integration, the problem is exacerbated as the number of information sources to be integrated, and hence the number of integration problems that must be solved, grows. Such schema matching problems arise both in "classical" scenarios such as company mergers, and in "new" scenarios such as the integration of diverse sets of query able information sources over the web. Purely manual solutions to the schema matching problem are too labor intensive to be scalable; as a result, there has been a great deal of research into automated techniques that can speed this process by either automatically discovering good mappings, or by proposing likely matches that are then verified by some human expert. In this paper we present such an automated technique that is designed to be of assistance in the particularly difficult cases in which the column names and data values are "opaque," and/or cases in which the column names are opaque and the data values in multiple columns are drawn from the same domain. Our approach works by computing the "mutual information" between pairs of columns within each schema, and then using this statistical characterization of pairs of columns in one schema to propose matching pairs of columns in the other schema

E. SOFTWARE DEFINED NETWORK [SDN]

As mentioned SDN stands for Software Defined Networking. We have seen the networking where in routers, switches, firewalls are installed based on the networking requirement after initial network layout. There will be problem when the requirement gets changed at later point of time. It will be difficult to modify the existing networking components and only option will be to dump the existing one and replace the one with higher capacity.

Here software components of the networking devices can not be modified. This leads to the development of SDN.In SDN software components can be customized and can also be configured based on requirements. This can be done independent of any hardware device i.e. all the networking hardware devices can be defined by changing the software installation at deployment time. These helps in enhancing the data flow across the systems. The effect of SDN can be easily felt by networking service providers and big business enterprises but not by the end user. In contrast, the end user of networking system will feel enhanced services, security and seamless usage of service.The other major enabler of SDN evolution are as follows:
• Change in the usage of traffic from different kind of networking devices by the user.
• Demand of scalability of storage, computing and networking resources by large and dynamic enterprises.

• Massive parallel processing on the servers and big data flow between them. SDN moved distributed data centres to the cloud region and this is considered to be major motivator for SDN evolution.

F. THREE LAYERED ARCHITECTURE OF SDN

SDN architecture composed of Application layer, controller layer and data plane layer. The traditional network had control plane, data plane and management plane in a single networking device. In software defined networks control plane and management planes are separated and known as controller.

Controller communicate with data plane using Open flow protocol and with open flow API integrated into data plane layer. Application layer sits on top of controller.

Open flow protocol basically allows external control software i.e. controller to control the data path of a switch by using flow table. The major benefit of adopting the SDN is world's largest networks such as google, yahoo, verizon, microsoft, facebook, NTT Communication Deutsche Telekom already supports SDN basedarchitecture.

The other benefit is that centralized management of networkingdevices.Improvement in automating the network components.

SDN can be used for multi-tenancy application. Multi-tenancy is the network where in distributed data centers installed for different customers can communicate securely.

Here multiple tenants can share the same physical resources. Also each tenant is assigned with unique logical resources.In summary SDN delivers networks with flexibility, scalability and efficiency.

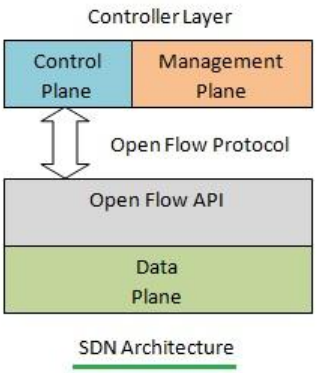

Fig3. SDN Architecture

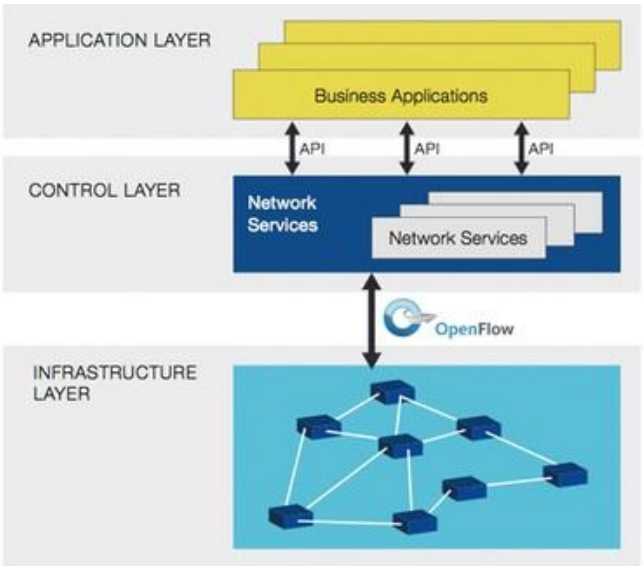

Fig4. SDN Three Layer Architecture

G. DETECTING AND RESOLVING UNSOUND WORKFLOW VIEWS FOR EFFICIENT PROVENANCE ANALYSIS

Technological advances have enabled the capture of massive amounts of data in many different domains, taking us a step closer to solving complex problems such as global climate change and uncovering the secrets hidden in genes. Workflow management systems are therefore increasingly used for managing and

analyzing this data, allowing users to specify complex, multi-step, "in-silicon" experiments or analyses. To ensure reproducibility and verifiability of results, many workflow systems are now providing support for provenance.The provenance of a data item is the sequence of steps used to produce the data, together with the intermediate data and parameters used as input to those steps. In general, it can be thought of as a graph which captures the causal dependencies between entities such as data and processes, and queries of provenance as calculating transitive closures of dependencies. As workflows become large and complex, the size of the provenance graph as well as the cost of answering transitive closure queries becomes problematic, and a number of techniques have recently been proposed for reducing the size of the provenance graph and complexity of calculating provenance information.

H. IMPLEMENTING THE CORE COMPONENT CRYPTOGRAPHICALLY SECURED SERVICES

1. Confidentiality Assurance

In a cryptographic storage service, the data is encrypted on-premise by the data processor(s). This way, customers can be assured that the confidentiality of their data is preserved irrespective of the actions of the cloud storage provider. This greatly reduces any legal exposure for both the customer and the provider.

2. Geographic restrictions

In a cryptographic storage service data is only stored in encrypted form so any law that pertains to the stored data has little to no effect on the customer. This reduces legal exposure for the customer and allows the cloud storage provider to make optimal use of its storage infrastructure, thereby reducing costs.

3. Subpoenas

In a cryptographic storage service, since data is stored in encrypted form and since the customer retains possession of all the keys, any request for the (unencrypted) data must be made directly to the customer.

VI. FAULT TOLERANT CLOUD BASED SECURED INFORMATION STORAGE

Cloud computing provides a promising opportunity forboth small and large organizations to transition from traditionaldata centers to cloud services, where the organizations can bemore concerned with their applications, services, and data ratherthan the underlying network infrastructures and their associatedcost. There are major concerns, however, with data security,reliability, and availability in the cloud. In this paper, we addressthese concerns by proposing a novel security mechanism forsecure and fault-tolerant cloud-based information storage. Wepresent a formal model of the security mechanism using coloredPetri nets (CPN). The model utilizes multiple cloud serviceproviders as a cloud cluster for information storage, and a servicedirectory for management of the cloud clusters including servicequery, key management, and cluster restoration. Our approachnot only supports maintaining the confidentiality of the storeddata, but also ensures that the failure or compromise of anindividual cloud provider in a cloud cluster will not result in a compromise of the overall data set.

VII. CONCLUSION

The secured storage is obtained by using the three layeredsoftware defined network integrated with Iot. And the cloud storage of Microsoft and google are been used for the experimental purpose and it resulted best to the users**.** Our framework can effectively evaluate the security level of the cloud offerings which can help the consumers in selecting the CloudIot.

REFERENCES

[1]. J. Baek, R. Safavi-Naini, and W. Susilo. Public key encryption with keyword search revisited. In International conference on Computational Science and Its Applications, pages 1249-1259. Springer-Verlag, 2008.

[2]. Belkin, N.J. and Croft, W.B. Information filtering andinformation retrieval: two sides of the same coin? Commun.ACM 35, 12 (1992), 29-38.

[3]. Liu, H. and Jacobsen, H.-A. A-TOPSS: a publish/subscribe system supporting approximate matching. Proceedings of the 28th international conference on Very Large Data Bases, VLDB Endowment (2002), 1107-1110.

[4]. Erhard Rahm, Philip A. Bernstein: A survey of approaches to automatic schema matching. VLDB Journal 10(4) (2001).

[5]. Jayant Madhavan, Philip A. Bernstein, Erhard Rahm:Generic Schema Matching with Cupid. VLDB 2001: 49-58.

[6]. Silvana Castano, Valeria De Antonellis, Sabrina De Capitani di Vimercati: Global Viewing of Heterogeneous Data Sources. TKDE 13(2): 277-297 (2001)

[7]. Q.Wang, C.Wang, J. Li, K. Ren, and W. Lou. Enabling public verifiability and data dynamics for storage security in cloud computing. In European Symposium on Research in Computer Security (ESORICS '09), volume 5789 of Lecture Notes in Computer Science, pages 355{370. Springer, 2009.
[8]. B. Zhang and F. Zhang, "An efficient public key encryption with conjunctive-subset keywords search", Journal of Network and Computer Applications, vol. 34, no. 1, (2011).
[9]. O. Biton, S. C. Boulakia, S. B. Davidson, and C. S. Hara, "Querying and managing provenance through user views in scientific workflows," in Proc. 24th Int. Conf. Data Eng., 2008, pp. 1072–1081. [Online]. Available: http://dx.doi.org/10.1109/ ICDE.2008.4497516.
[10]. S. Kamara and K. Lauter, "Cryptographic Cloud Storage. Proceedings of Workshops on Financial Cryptography and Data Security", (2010) January 25-28; Canary Islands, Spain.

M. Manoranjitham," Analysis Of Secured Storage And Secured Dataflow In Cloudiot (Storage And Dataflow Management In Cloudiot)." IOSR Journal of Engineering (IOSRJEN), vol. 08, no. 8, 2018, pp. 50-56.