

Data Access Control with Revocable Multiauthority Cloud Storage

P.Srilakshmi¹, Dr.A.V.R.Mayuri², K.Asha Rani³

¹Assistant Professor, G .Pulla Reddy Engineering College (Autonomous),
Department of Computer Science & Engineering, Nandyal Road, Kurnool
Srilakshmi.reddy99@gmail.com

²Assistant Professor, G. Pulla Reddy Engineering College (Autonomous),
Department of Computer Science & Engineering, Nandyal Road, Kurnool
mayuriavr@gmail.com

³Assistant Professor, G. Pulla Reddy Engineering College (Autonomous),
Department of Computer Science & Engineering, Nandyal Road, Kurnool
Asha.ashreddy@gmail.com

Abstract: Data access control is an effective way to ensure the data security in the cloud. Due to data outsourcing and untrusted cloud servers, the data access control becomes a challenging issue in cloud storage systems. Cipher text-Policy Attribute based Encryption (CP-ABE) is regarded as one of the most suitable technologies for data access control in cloud storage, because it gives data owners more direct control on access policies. However, it is difficult to directly apply existing CP-ABE schemes to data access control for cloud storage systems because of the attribute revocation problem. In this paper, an expressive, efficient and revocable data access control scheme for multi-authority cloud storage systems, where there are multiple authorities co-exist and each authority is able to issue attributes independently. Precisely, a revocable multi-authority CP-ABE scheme, and apply it as the underlying techniques to design the data access control scheme. A attribute revocation method can efficiently achieve both forward security and backward security.

Keywords: Cipher text-Policy Attribute based Encryption (CP-ABE), Certificate Authority, Attribute Authority, Data Consumers, Data Owners

I.Introduction

Cloud storage is an important service of cloud computing, which offers services for data owners to host their data in the cloud. This new paradigm of data hosting and data access services introduces a great challenge to data access control. Because the cloud server cannot be fully trusted by data owners, they can no longer rely on servers to do access control. Cipher text Policy Attribute-based Encryption (CP-ABE) is regarded as one of the most suitable technologies for data access control in cloud storage systems, because it gives the data owner more direct control on access policies. The authority can be the registration office in a university, the human resource department in a company, etc. The data owner defines the access policies and encrypts data according to the policies. Each user will be issued a secret key reflecting its attributes. A user can decrypt the data only when its attributes satisfy the access policies. There are two types of CP-ABE systems: single-authority CP-ABE where all attributes are managed by a single authority, and multi-authority CP-ABE where attributes are from different domains and managed by different authorities. Multi-

authority CP-ABE is more appropriate for data access control of cloud storage systems, as users may hold attributes issued by multiple authorities and data owners may also share the data using access policy defined over attributes from different authorities. For example, in an E-health system, data owners may share the data using the access policy “Doctor AND Researcher”, where the attribute “Doctor” is issued by a medical organization and the attribute “Researcher” is issued by the administrators of a clinical trial. However, it is difficult to directly apply these multi-authority CP-ABE schemes to multi-authority cloud storage systems because of the attribute revocation problem. In multi-authority cloud storage systems, users’ attributes can be changed dynamically. A user may be entitled some new attributes or revoked some current attributes. And his permission of data access should be changed accordingly. However, existing attribute revocation methods either rely on a trusted server or lack of efficiency, they are not suitable for dealing with the attribute revocation problem in data access control in multi-authority cloud storage systems. In this we first propose a revocable multi authority CP-ABE scheme, where an

efficient and secure revocation method is proposed to solve the attribute revocation problem in the system. A attribute revocation method is efficient in the sense that it incurs less communication cost and computation cost, and is secure in the sense that it can achieve both backward security (The revoked user cannot decrypt any new cipher text that requires the revoked attribute to decrypt) and forward security (The newly joined user can also decrypt the previously published ciphertexts, if it has sufficient attributes). Our scheme does not require the server to be fully trusted, because the key update is enforced by each attribute authority not the server. Even if the server is not semi-trusted in some scenarios, our scheme can still guarantee the backward security. Then, we apply our proposed revocable multi-authority CP-ABE scheme as the underlying techniques to construct the expressive and secure data access control scheme for multi-authority cloud storage systems with the help of following authorities

1. Certificate Authority:

The CA is a global trusted certificate authority in the system. It sets up the system and accepts the registration of all the users and AAs in the system. For each legal user in the system, the CA assigns a global unique user identity to it and also generates a global public key for this user. However, the CA is not involved in any attribute management and the creation of secret keys that are associated with attributes. For example, the CA can be the Social Security Administration, an independent agency of the United States government. Each user will be issued a Social Security Number (SSN) as its global identity.

2. Attribute Authorities:

Every AA is an independent attribute authority that is responsible for entitling and revoking user's attributes according to their role or identity in its domain. In our scheme, every attribute is associated with a single AA, but each AA can manage an arbitrary number of attributes. Every AA has full control over the structure and semantics of its attributes. Each AA is responsible for generating a public attribute key for each attribute it manages and a secret key for each user reflecting his/her attributes.

3. Data Consumers:

Each user has a global identity in the system. A user may be entitled a set of attributes which may come from multiple attribute authorities. The user will receive a secret key associated with its attributes entitled by the corresponding attribute authorities.

4. Data Owners:

Each owner first divides the data into several components according to the logic granularities and encrypts each data component with different content keys by using symmetric encryption techniques. Then, the owner defines the access policies over attributes from multiple attribute authorities and encrypts the content keys under the policies.

5. Cloud Server:

Then, the owner sends the encrypted data to the cloud server together with the cipher texts. They do not rely on the server to do data access control. But, the access control happens inside the cryptography. That is only when the user's attributes satisfy the access policy defined in the cipher text; the user is able to decrypt the cipher text. Thus, users with different attributes can decrypt different number of content keys and thus obtain different granularities of information from the same data.

Compared to the conference version of this work, we have the following improvements:

1. We modify the framework of the scheme and make it more practical to cloud storage systems, in which data owners are not involved in the key generation. Specifically, a user's secret key is not related to the owner's key, such that each user only needs to hold one secret key from each authority instead of multiple secret keys associated to multiple owners.

2. We greatly improve the efficiency of the attribute revocation method. Specifically, in our new attribute revocation method, only the cipher texts that associated with the revoked attribute need to be updated, all the cipher texts that associated with any attribute from the authority (corresponding to the revoked attribute) should be updated. Moreover, in our new attribute revocation method, both the key and the cipher text can be updated by using the same update key, instead of requiring the owner to generate an update information for each cipher text, such that owners

are not required to store each random number generated during the encryption.

3. We also highly improve the expressiveness of our access control scheme, where we remove the limitation that each attribute can only appear at most once in a cipher text.

II. SYSTEM ARCHITECTURE

What is cloud computing?

Cloud computing is the use of computing resources (hardware and software) that are delivered as a service over a network (typically the Internet). The name comes from the common use of a cloud-shaped symbol as an abstraction for the complex infrastructure it contains in system diagrams. Cloud computing entrusts remote services with a user's data, software and computation. Cloud computing consists of hardware and software resources made available on the Internet as managed third-party services. These services typically provide access to advanced software applications and high-end networks of server computers.

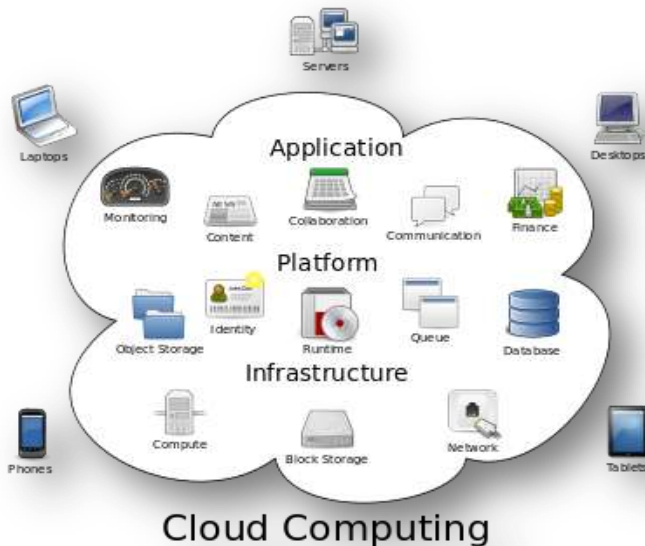


Fig 1: Structure of cloud computing

How Cloud Computing Works?

The goal of cloud computing is to apply traditional super computing, or high-performance computing power, normally used by military and research facilities, to perform tens of trillions of computations per second, in consumer-oriented applications such as financial

portfolios, to deliver personalized information, to provide data storage or to power large, immersive computer games. The cloud computing uses networks of large groups of servers typically running low-cost consumer PC technology with specialized connections to spread data-processing chores across them. This shared IT infrastructure contains large pools of systems that are linked together. Often, virtualization techniques are used to maximize the power of cloud computing.

Characteristics and Services Models:

The salient characteristics of cloud computing based on the definitions provided by the National Institute of Standards and Terminology (NIST) are outlined below:

- **On-demand self-service:** A consumer can unilaterally provision computing capabilities, such as server time and network storage, as needed automatically without requiring human interaction with each service's provider.
- **Broad network access:** Capabilities are available over the network and accessed through standard mechanisms that promote use by heterogeneous thin or thick client platforms (e.g., mobile phones, laptops, and PDAs).
- **Resource pooling:** The provider's computing resources are pooled to serve multiple consumers using a multi-tenant model, with different physical and virtual resources dynamically assigned and reassigned according to consumer demand. There is a sense of location-independence in that the customer generally has no control or knowledge over the exact location of the provided resources but may be able to specify location at a higher level of abstraction (e.g., country, state, or data center). Examples of resources include storage, processing, memory, network bandwidth, and virtual machines.
- **Rapid elasticity:** Capabilities can be rapidly and elastically provisioned, in some cases automatically, to quickly scale out and rapidly released to quickly scale in. To the consumer, the capabilities available for provisioning often appear to be unlimited and can be purchased in any quantity at any time.
- **Measured service:** Cloud systems automatically control and optimize resource use by leveraging a metering capability at some level of abstraction appropriate to the

type of service (e.g., storage, processing, bandwidth, and active user accounts). Resource usage can be managed, controlled, and reported providing transparency for both the provider and consumer of the utilized service.



Fig 2: Characteristics of cloud computing

III Challenges and Our Contributions

The main challenge of our construction is to formulate a reasonable security model and provide formal security proofs when combining CP-ABE with proxy re-encryption. Our contribution can be summarized as follows.

Firstly, we provide the definition for attribute revocation in CP-ABE with honest-but-curious servers, and formulate the security model to reflect possible attacks. Secondly, the proposed scheme enables the authority to revoke any attribute of users at any time while placing a minimal load on him. Thirdly, the proposed scheme is provably secure under the Decisional Bilinear Diffie-Hellman (DBDH) assumption. Last but not least, our method is applicable to the KP-ABE counterpart in which the authority is able to revoke any partial access privilege of users. To the best of our knowledge, this paper is among the first formally addressing the issue of user/attribute revocation in ABE although it focuses on a practical setting.

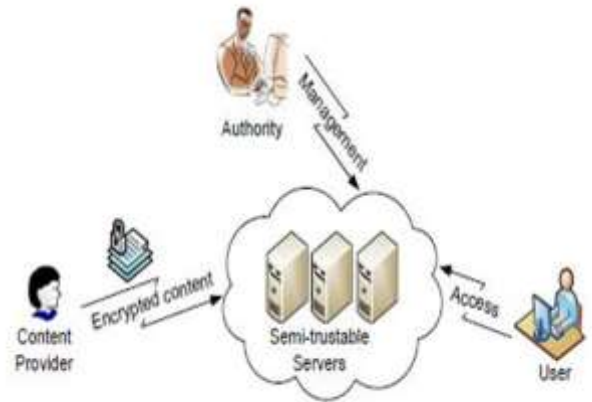


Fig3. An example application scenario of data sharing

3.1. Attribute Based Encryption (ABE) and its two flavors:

In a conventional public key cryptosystem, there are two distinct keys: a public key and a private key. Bob uses Alice's public key to encrypt a message to Alice. Alice uses her private key to decrypt the message. How does Bob know the public key of Alice? Alice may have given the public key using a secure communication channel. This works only if there is already some trust/familiarity between both Bob and Alice. What if Bob and Alice do not know each other? This method fails. In an unknown open system, we need a trusted third party (TTP) to uniquely bind public keys to users or another entity such as an organization. This is where we need a PKI (Public Key Infrastructure). A PKI has one more trusted entities called Certification Authorities (CAs). For example, VeriSign is a CA. CA issues Alice a certificate (which contains the public key of Alice) signed by the CA's public key after verifying Alice's credentials. Bob can now retrieve Alice's certificate and verify it is authentic by checking the signature on it. Certificates may need to be revoked later due to various reasons. For example, if Alice's private key is stolen, she will have to ask the CA to revoke its certificate. How does Bob know if a certificate is revoked? The CA maintains a revocation list which allows Bob to verify if a given certificate is revoked or not. The traditional PKI is somewhat cumbersome as one needs to retrieve certificates, check revocation list, and then encrypt. Is there a better way of doing public key cryptography? Some smart researchers came up with the

idea of using user identities (for example, your email address) as public keys. Such systems are called IBE (identity based encryption).

The idea is as follows:

There is a TTP called KGS (Key Generation Server). Given an identity, KGS generates a private key and the identity acts as the public key. For example, Alice's identity is her email address `alice@example.com`. Alice uses this identity to obtain a private key from the KGS. Now Bob encrypts a message using Alice's email. Only Alice can decrypt the message since the identity, the public key, `alice@example.com` belongs to Alice and only she can obtain the private from the KGS. Notice that there is a huge trust placed on the KGS. The security of the whole system relies on the security of the KGS and how well the KGS authenticates users before issuing private keys.

The idea of IBE was further improved to support much better systems. The concept of attribute-based encryption (ABE) has been introduced by Sahai and Waters. ABE can be considered as a generalization of identity based encryption (IBE), where, as mentioned earlier, the encryption is based on some identity. Thus, ABE is more expressive than IBE. In an ABE system, the plaintext is encrypted with a set of attributes. The KGS, which possesses the master key, issues different private keys to users after authenticating the attributes they possess.

Thus, these private keys are associated with the set of attributes each user possesses. In its basic form, a user can decrypt a cipher text if and only if there is a match between the attributes of the cipher-text and the user's key. For example, Alice has the attributes "role = doc" and "age > 18". Now Bob encrypts a message using the attributes ("role = student" AND "age > 18"). Alice can decrypt the message as she satisfies both attributes. Bob encrypts another message using the attributes ("role = professor" OR "role = staff"). Alice cannot decrypt the message as she does not satisfy the policy. (The workings of the actual ABE schemes are a little different from the above examples, but they give the essential idea behind the schemes.)

The initial ABE system is limited only to threshold policies where there should be at least k out of n attributes common between the attributes used to encrypt the plaintext and the attributes users possess. Pirretti gave an implementation of such a threshold ABE system using a variant of the Sahai-Waters Large Universe construction. For example, Bob encrypts a message for any 3 attributes out of the 5 attributes $\{a_1, a_2, a_3, a_4, a_5\}$. Alice has the attributes $\{a_1, a_2, a_4, a_5\}$ and Eve has $\{a_1, a_2\}$. While Alice can decrypt Bob's message, Eve cannot as she does not satisfy the threshold policy.

Since the initial threshold scheme, a few variants have been introduced to provide more expressive ABE systems. Out of them there are two important variants.

1. Key Policy ABE (KP-ABE)
2. Cipher-text Policy ABE (CP-ABE)

Goyal introduced the idea of KP-ABE systems and Bethencourt introduced the idea of CP-ABE systems.

Let's try to understand the idea behind these two variants using diagrams.

3.1.1. Key Policy ABE (KP-ABE)

As shown in the fig3.1.1, in KP-ABE, Bob encrypts a message using a set of attributes. It defines an access structure, which is a threshold tree of the policy that Bob wants to enforce. Alice and Tim tries to decrypt the message. The attributes Alice has satisfy the access structure and hence she can derive the key and decrypt the document. The attributes Tim has do not satisfy the access structure and therefore cannot derive the key to decrypt the message. The key idea here is that the key is associated with the policy using an access structure.

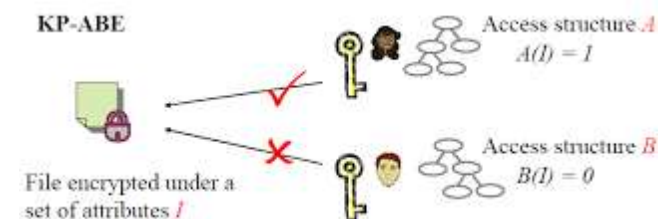


Fig3.1.1. Key Policy ABE

3.1.2. Cipher-text Policy ABE (CP-ABE)

As shown in the fig3.1.2, CP-ABE reverses the role of encryption and key derivation. The encryption is associate with an access structure which is constructed using the policy. KGS simply issues private keys for the attributes users have. If users (rather their attributes) satisfy the owner defined access structure, they can decrypt it. The second variant is more closer to encryption found in open systems as the cipher-text is associated the policy.

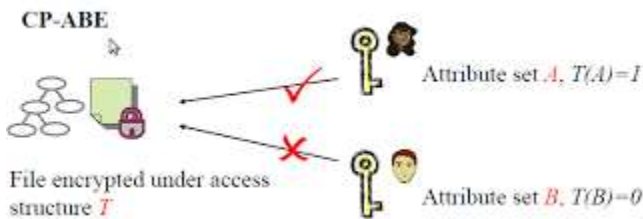


Fig3.5.2. Cipher-text Policy ABE

IV DriveHQ

About DriveHQ

Based in Silicon Valley, DriveHQ is the first cloud IT solution provider before the term —cloud| was made popular. DriveHQ has offered a broad array of cloud services for over 10 years, quite a few years before competitors like Dropbox, Box, Google Drive, Egnyte, Carbonite, Sky Drive, iCloud and Amazon S3 started offering cloud services. DriveHQ has been steadily growing its business and increasing its customer base. It has over 2million registered users. We are a technology and service oriented company. For the last 10 years, we have developed more and better technologies than any of our competitors.

Many people just assumed that companies like Google, Dropbox or Box are innovative. In reality it is not true, e.g.: Google did not invent search; it did not invent smart phone; it bought Android through acquisition, and Android is really based on Linux and Java (vs. Apple and Microsoft who developed their own smart phone operating systems). In the cloud storage industry, Google Drive is many years late and it does not bring anything new. In fact, it simply copied features of other cloud services. Neither Dropbox nor Box pioneered cloud storage industry. Their success is closely linked with their marketing efforts and VC funding. DriveHQ pioneered the cloud storage and IT industry, our

success is linked with our superior technologies and services, and our loyal customers.

We have already compared DriveHQ cloud IT service with some other major cloud storage service providers. There are many cloud service providers; it is not possible to compare with all of them. However, many cloud services are very similar, so we don't need to compare with all of them. This document tries to give some general idea about how to compare DriveHQ service with other cloud storage services.

The differences among cloud storage services

Cloud storage services mainly can be categorized as follows:

- **Raw Cloud Storage Service:** Usually offered as part of IaaS (Infrastructure as a Service)service. It lacks advanced features.
- **Cloud Folder Sync Service:** Usually include web browser based online storage service and a folder sync client to sync a local folder with a remote folder.
- **Cloud Data Backup Service:** Automatically backup data on a local PC to the cloud using a backup client program.
- **Cloud File Sharing and Collaboration Service:** Share files/folders online, collaborate with others.
- **Cloud Hosting Service:** Cloud-based FTP server hosting, email server hosting; web hosting, file hosting, etc.
- **Cloud Drive Mapping Service:** Map cloud storage as a network drive. Different cloud service providers may offer different category services; even within the same category, the actual service can be still very different.

What to look for when choosing a Cloud Service Provider?

With so many available options, it is hard to choose the best service for you. Below are a few principles in choosing the best provider:

- **Understand the differences of each cloud service:** DriveHQ has a very detailed feature comparison chart. Other websites may also offer such comparisons. Trust the one that provides the detailed analysis.
- **Understand your current and future needs:** You might be looking for one feature only. But ask yourself if you need any other features in the future, and also think if you

will use other features if it is already included for no extra charge.

- **Understand how many user accounts you need:**

Number of user licenses is often the most important pricing factors, for example, Egnyte and Box charge \$180/user/year, Dropbox charges \$125/user/year and Google Drive charges \$50-\$100/user/year.

- **Keep in mind about scalability:** You might be lured to some easy / simple service. However, being easy and simple also means having fewer services/features. When you have more employees or your storage increases, such service may no longer be suitable feature-wise and cost-wise.

- **The service provider's track record:** Always choose a business that has been proven. In DriveHQ's case, ten years in business, many years of profitability, and strong customer reputation speak for itself. On the other hand, most other cloud service providers have not fully proven themselves. Most of them relying on huge marketing budget and have not achieved profitability.

The weaknesses of other cloud-based folder synchronization service

There are many companies mimicking Dropbox service. In fact, when you analyze services from Box, Egnyte and Google Drive, the folder sync part is almost identical. They all create a special folder on your computer and sync the folder with the cloud storage. While we have not fully reviewed other services, but even Google Drive is copying Dropbox.

- **The weaknesses of other cloud backup service**

Cloud based backup service is often more secure and reliable than local backup service. Your data can survive major disasters like fire, flood, earthquake, etc. However, any backup-only service is inadequate for businesses / enterprises. Most businesses and enterprises need to access, share and collaborate files remotely.

- **The weaknesses of other cloud-based file sharing and collaboration service**

Most other cloud service providers have relied on browser based file sharing and collaboration service. However,

requiring all users to use browsers for file management, sharing and collaboration not only require re-training of many employees, but also decreases employee productivity as web browser-based service is inherently less efficient than native applications like DriveHQ ,FileManager or WebDAV drive mapping.

- **The weaknesses of traditional hosting services**

Traditional hosting services use conventional FTP servers, email servers and web servers. In some cases, these hosting services are more versatile. However, DriveHQ's hosting service is easier than traditional hosting services. You just need to sign up a DriveHQ account and you will have access to these hosting services immediately. Moreover, such hosting services are seamlessly integrated with DriveHQ's cloud IT service, making it extremely easy to manage files, folders, user accounts and access permissions.

- **The weaknesses of raw cloud storage services**

Amazon offers raw cloud storage service with its S3 service. It is designed to be part of the internet's infrastructure service (IaaS). Because of this, it lacks advanced features. If a business does not have strong technical expertise, such IaaS service requires a lot of development and integration work and should be avoided. If a business needs huge amount of storage space, then DriveHQ's storage price can match or beat Amazon's S3.

- **The weaknesses of other cloud-based WebDAV drive mapping service**

WebDAV drive mapping clearly seems to be the most straight-forward cloud storage service. You can map your cloud storage as a network drive. It works just like a local drive. Some other cloud service providers may also offer WebDAV drive mapping service. However, regular WebDAV drive mapping service is inefficient and unreliable. That's why most cloud service providers have copied Dropbox's folder sync feature instead of pushing for WebDAV.

IV SERVICE MODELS & BENEFITS

4.1 Services Models:

Cloud Computing comprises three different service models, namely Infrastructure-as-a-Service (IaaS), Platform-as-a-Service (PaaS), and Software-as-a-Service (SaaS). The three service models or layer are completed by an end user layer that encapsulates the end user perspective on cloud services. The model is shown in figure below. If a cloud user accesses services on the infrastructure layer, for instance, she can run her own applications on the resources of a cloud infrastructure and remain responsible for the support, maintenance, and security of these applications herself. If she accesses a service on the application layer, these tasks are normally taken care of by the cloud service provider.

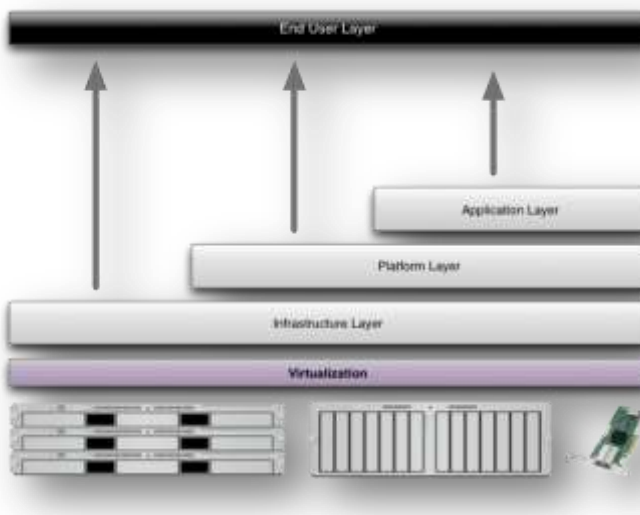


Fig 4 :Structure of service models

4.2 Benefits of cloud computing:

1. **Achieve economies of scale** – increase volume output or productivity with fewer people. Your cost per unit, project or product plummets.
2. **Reduce spending on technology infrastructure.** Maintain easy access to your information with minimal upfront spending. Pay as you go (weekly, quarterly or yearly), based on demand.
3. **Globalize your workforce on the cheap.** People worldwide can access the cloud, provided they have an Internet connection.
4. **Streamline processes.** Get more work done in less time with less people.

5. **Reduce capital costs.** There's no need to spend big money on hardware, software or licensing fees.
6. **Improve accessibility.** You have access anytime, anywhere, making your life so much easier!
7. **Monitor projects more effectively.** Stay within budget and ahead of completion cycle times.
8. **Less personnel training is needed.** It takes fewer people to do more work on a cloud, with a minimal learning curve on hardware and software issues.
9. **Minimize licensing new software.** Stretch and grow without the need to buy expensive software licenses or programs.
10. **Improve flexibility.** You can change direction without serious "people" or "financial" issues at stake.

Advantages:

1. **Price:** Pay for only the resources used.
2. **Security:** Cloud instances are isolated in the network from other instances for improved security.
3. **Performance:** Instances can be added instantly for improved performance. Clients have access to the total resources of the Cloud's core hardware.
4. **Scalability:** Auto-deploy cloud instances when needed.
5. **Uptime:** Uses multiple servers for maximum redundancies. In case of server failure, instances can be automatically created on another server.
6. **Control:** Able to login from any location. Server snapshot and a software library lets you deploy custom instances.
7. **Traffic:** Deals with spike in traffic with quick deployment of additional instances to handle the load.

V CONCLUSION

This paper introduces revocable multi-authority CP-ABE scheme that can support efficient attribute revocation. Then, a secure data access control scheme for multi-authority cloud storage systems is constructed. It also proved that the scheme was secure in the random oracle model. The revocable multi-authority CP-ABE is a promising technique, which can be applied in any remote storage systems and online social networks etc. This scheme does not require the server to be fully trusted, because the key update is enforced by each attribute

authority not the server. Even if the server is not semi-trusted in some scenarios, the introduced scheme can still guarantee the backward security. Then, we apply revocable multi-authority CP-ABE scheme as the underlying techniques to construct the expressive and secure data access control scheme for multi-authority cloud storage systems.

References

- [1] B. Waters, Cipher text-Policy Attribute-Based Encryption: An Expressive, Efficient, and Provably Secure Realization,“ in Proc. 4th Int’l Conf. Practice and Theory in Public Key Cryptography (PKC’11), 2011, pp. 53-70
- [2] S. Jahid, P. Mittal, and N. Borisov, Easier: Encryption-Based Access Control in Social Networks with Efficient Revocation,“ in Proc. 6th ACM Symp. Information, Computer and Comm. Security (ASIACCS’11), 2011, pp. 411-415.
- [3] A.B. Lewko and B. Waters, New Proof Methods for Attribute- Based Encryption: Achieving Full Security through Selective Techniques,“ in Proc. 32st Ann. Int’l Cryptology Conf.: Advances in Cryptology - CRYPTO’12, 2012, pp. 180-198.
- [4] K. Yang and X. Jia, Attribute-Based Access Control for Multi-Authority Systems in Cloud Storage,“ in Proc. 32th IEEE Int’l Conf. Distributed Computing Systems (ICDCS’12), 2012, pp. 1-10.
- [5] M. Li, S. Yu, Y. Zheng, K. Ren, and W. Lou, Scalable and Secure Sharing of Personal Health Records in Cloud Computing Using Attribute-Based Encryption,“ IEEE Trans. Parallel Distributed Systems, vol. 24, no. 1, pp. 131-143, Jan. 2013.