An Effective and Fine Grained Big Data access Control Scheme with Protection Policy

Prema C.

Assistant Professor,
Department of CSE,
R. R. Institute of Technology, Bengaluru,
Karnataka, India

Arjun

UG Student,
R.R Institute of Technology,
Visvesvaraya Technological University,
Bangalore, India

ABSTRACT

In order to control the access of the huge amount of big data becomes a very challenging issue, especially when big-data are stored in the cloud. Cipher text- policy attribute-based encryption (CP-ABE) is a promising encryption technique that enables end-users to encrypt their data under the access policies defined over some attributes of data consumers and only allows data consumers whose attributes satisfy the access policies to decrypt the data. In CP-ABE, the access policy is attached to the cipher text in plaintext form, which may also leak some private information about end-users. Existing methods only partially hide the attribute values in the access policies, while the attribute names are still unprotected. In this paper, we propose an efficient and fine-grained big data access control scheme with privacy-preserving policy. Specifically, we hide the whole attribute (rather than only its values) in the access policies. To assist data decryption, we also design a novel attribute bloom filter to evaluate whether an attribute is in the access policy and locate the exact position in the access policy, if it is in the access policy. Security analysis and performance evaluation show that our scheme can preserve the privacy from any linear secret-sharing schemes access policy without employing much over head.

Keywords: CP-ABE, Big data, Decryption, Secret-sharing.

INTRODUCTION:

In the era of big data, a huge amount of data can be generated quickly from various sources (e.g., smart phones, sensors, machines, social networks, etc.). Toward these big data, conventional computer systems are not competent to store and process these data. Due to the flexible and elastic computing resources, cloud computing is a natural fit for storing for execution of processing pipelines among heterogeneous event processing engines as a workflow. In traditional workflow, tasks are executed once or several times at some control flows like iterations. In contrast to it, streaming workflows that are constantly responding to environmental conditions based on stream inputs allow tasks in the workflow to be invoked multiple times in , which involves movement of huge amount of data between execution unquestionable trust in the cloud provider, in some cases corroborated by reports of external auditors. While providers may offer security enhancements such as protection of data at rest, end-users have limited or no control over such mechanisms. There is a clear need for usable and cost-effective cloud platform security mechanisms suitable for organizations that rely on cloud infrastructure. One such mechanism is platform integrity verification for compute hosts that support the virtualized cloud infrastructure. Several large cloud vendors have signaled practical implementations of this mechanism, primarily to protect the cloud infrastructure from insider threats and advanced persistent threats. We see two major improvement vectors regarding these implementations. First, details of such proprietary solutions are not disclosed and can thus not be implemented and improved by other cloud platforms. Second, to the best of our knowledge, none of the solutions provides cloud tenants a proof regarding the integrity and processing big data. With cloud computing, end-users

store their data into the cloud, and rely on the cloud server to share their data to other users (data consumers). In order to only share end-users' data to authorized users, it is necessary to design access control mechanisms according to the requirements of end-users.

RELATEDWORK:

In Computer science, cloud computing describes a type of outsourcing of computer services, similar to the way in which electricity supply is outsourced. Users can simply use it. They do not need to worry where the electricity is from, how it is made, or transported. Every month, they pay for what they consumed. The idea behind cloud computing is similar: The user can simply use storage, computing power, or specially crafted development environments, without having to worry how these work internally. Cloud computing is usually Internet-based computing. The cloud is a metaphor for the Internet based on how the internet is described in computer network diagrams; which means it is an abstraction hiding the complex infrastructure of the internet. It is a style of computing in which IT-related capabilities are provided "as a service", allowing users to access technology-enabled services from the Internet ("in the cloud")without knowledge of, or control over the technologies behind these servers. Cloud computing is a general concept that utilizes software as a service (SaaS), such as Web 2.0 and other technology trends, all of which depend on the Internet for satisfying users' needs. For example, Google Apps provides common business applications online that are accessed from a web browser, while the software and data are stored on the Internet servers. Cloud computing is often confused with grid computing (a form of distributed computing whereby a "super and virtual computer" is composed of a cluster of networked, loosely-coupled computers, working together to perform very large tasks), utility computing (the packaging of computing resources, such as computation and storage are provided as a measured service that have to be paid similar to a traditional public utility such as electricity) and autonomic computing (computer systems capable of self-management).Many cloud computing deployments are powered by grids, have autonomic characteristics and are billed like utilities, but cloud computing can be seen as a natural next step from the grid-utility model. Some successful cloud architectures have little or no centralized infrastructure or billing systems at all including peer-to-peer networks like Bit Torrent and Skype. The majority of cloud computing infrastructure currently consists of reliable services delivered through data centers that are built on computer and storage virtualization technologies.

PROPOSED SYSTEM:

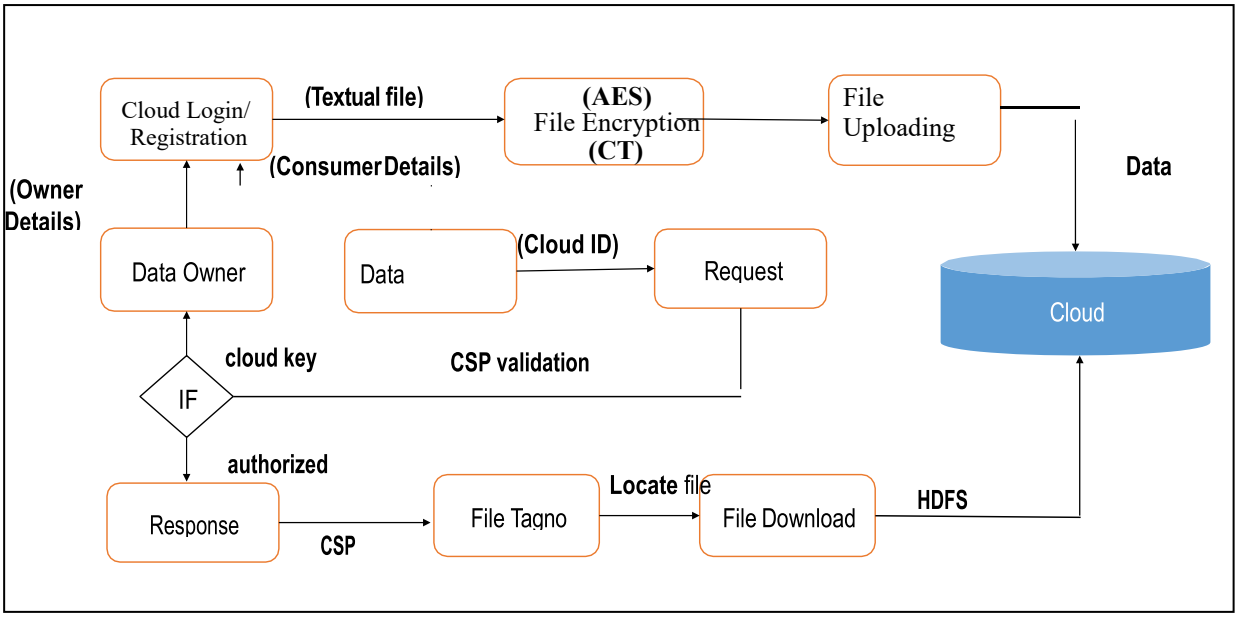
Cipher-text policy attribute-based encryption (CP-ABE) is a promising encryption technique that enables end-users to encrypt their data under the access policies defined over some attributes of data consumers and only allows data consumers whose attributes satisfy the access policies to decrypt the data. In CP-ABE, the access policy is attached to the cipher-text in plaintext form, which may also leak some private information about end-users. Cloud computing is a general term for anything that involves delivering hosted services over the Internet. It is an emerging computing technology that uses the internet and central remote servers to maintain data. This system is very helpful for different users so that they can easily use the system without any external support to software and hardware. They can also access their personal files at any computer on internet. This technology allows for much more efficient computing by centralizing storage, memory, processing and bandwidth. According to the different types of services offered, cloud computing can be considered as of three layers. we propose an efficient and fine-grained big data access control scheme with privacy-preserving policy. Specifically, we hide the whole attribute (rather than only its values) in the access policies. To assist data decryption, we also design a novel attribute bloom filter to evaluate whether an attribute is in the access policy and locate the exact position in the access policy if it is in the access policy.

EXISTING SYSTEM:

When the attributes are hidden, not only the unauthorized users but also the authorized users cannot know which attributes are involved in the access policy, which makes the decryption a challenging problem. Due to this reason, existing methods do not hide or anonymize the attributes. Toward this problem, some works have been proposed to hide the access policy. In two constructions are proposed to partially hide the access policy. Instead, they only hide the values of each attribute by using wildcards , hidden vector encryption ,and inner product encryption proposed a "Trusted Cloud Compute Plat-form" (TCCP) to ensure VMs are running on a trusted hardware and software stack on a remote and initially untrusted host. To enable this, a trusted coordinator stores the list of attested hosts that run a "trusted virtual machine monitor" which can securely run the client's VM. Trusted hosts

maintain in memory an individual trusted key used for identification each time a client launches a VM. This paper presents a good initial set of ideas for trusted VM launch and migration, in particular the use of a trusted coordinator. We further build a novel attribute bloom filter (ABF) to locate the attributes to the anonymous access policy, which can save a lot of storage overhead and computation cost especially for large attribute universe.

SYSTEM DESIGN:



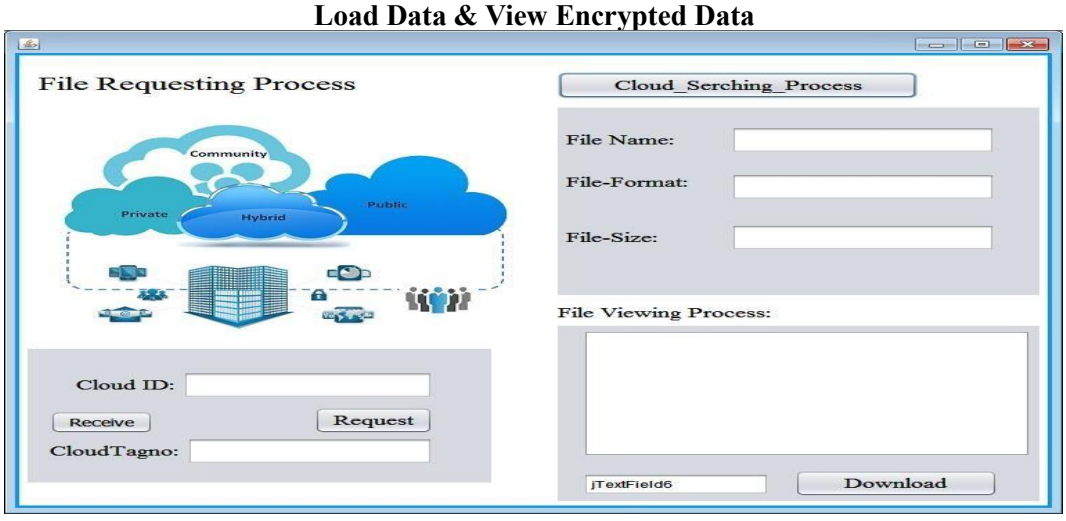
System Architecture (overall design)

AES Algorithm: AES is an iterated symmetric block cipher, which means that:
a) AES works by repeating the same defined steps multiple times. b) AES is a secret key encryption algorithm. c) AES operates on a fixed number of bytes
AES as well as most encryption algorithms is reversible. This means that almost the same steps are performed to complete both encryption and decryption in reverse order. The AES algorithm operates on bytes, which makes it simpler to implement and explain. This key is expanded into individual sub keys, a sub keys for each operation round. As mentioned before AES is an iterated block cipher. All that means is that the same operations are performed many times on a fixed number of bytes.

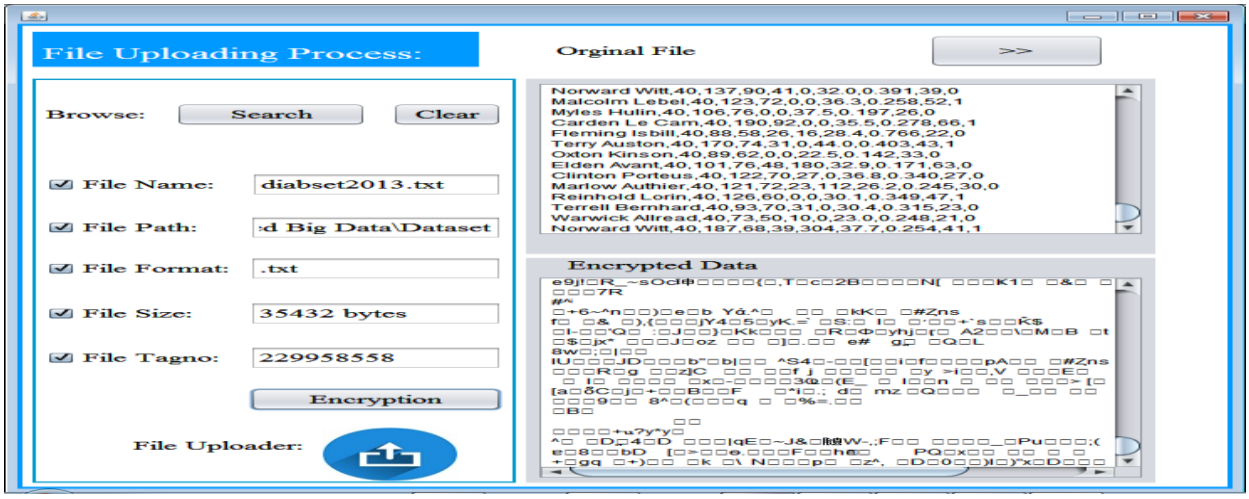
RESULTS:

We have demonstrated that our scheme is selectively secure against chosen plaintext attacks. Moreover, we have implemented the ABF by using Murmur Hash and the access control scheme to show that our scheme can preserve the privacy from any LSSS access policy without employing much overhead.

Load Data & View Encrypted Data



Data Consumer Request



CONCLUSION AND FUTURE WORK:

From a tenant point of view, the cloud security model does not yet hold against threat models developed for the traditional model where the hosts are operated and used by the same organization. However, there is a steady progress towards strengthening the IaaS security model. In this work we presented a framework for trusted infrastructure cloud deployment, with two focus points: VM deploy mention trusted compute hosts and domain-based protection of stored data. We described in detail the design, implementation and security evaluation of protocols for trusted VM launch and domain-based storage protection. In our future work, we will focus on how to deal with the offline attribute guessing attack that check the guessing "attribute strings" by continually querying the ABF.

REFERENCES:

B. Grobauer, T. Walloschek, and E. Stocker, "Understanding cloud computing vulnerabilities, "IEEE Security and Privacy, Vol.9, No. 2, 2011, pp. 50-57.

D. Boru, D. Kliazovich, F. Granelli, P. Bouvry, and A. Y. Zomaya,"Energy-efficient data replication in cloud computing datacenters," In IEEE Globecom Workshops, 2013, pp. 446-451.

K. Bilal, M. Manzano, S. U. Khan, E. Calle, K. Li, and A.Zomaya, "On the characterization of the structural robustness of data center networks,"IEEE Transactions on Cloud Computing, Vol. 1, No. 1, 2013, pp. 64-77.

K. Bilal, S. U. Khan, L. Zhang, H. Li, K. Hayat, S. A. Madani,N. Min-Allah, L. Wang, D. Chen, M. Iqbal, C. Z. Xu, and A. Y.Zomaya, "Quantitative comparisons of the state of the art datacenter architectures," Concurrency and Computation: Practice and Experience, Vol. 25, No. 12, 2013, pp. 1771-1783.

M. Hogan, F. Liu, A.Sokol, and J. Tong, "NIST cloud computing standards roadmap," NIST Special Publication, July 2011.

W. A. Jansen, "Cloud hooks: Security and privacy issues in cloud computing, "In 44th Hawaii IEEE International Conference on System Sciences (HICSS), 2011, pp. 1-10.

W. K. Hale, "Frequency assignment: Theory and applications, "Proceedings of the IEEE, Vol. 68, No. 12, 1980, pp. 1497-1514.[7] K. Hashizume, D. G. Rosado, E. Fernandez-Medina, and E. B. Fernandez, "An analysis of security issues for cloud computing," Journal of Internet Services and Applications, Vol. 4, No. 1,2013, pp. 1-13.

Y. Deswarte, L. Blain, and J-C. Fabre, "Intrusion tolerance in distributed computing systems," In Proceedings of IEEE Computer Society Symposium on Research in Security and Privacy, Oakland CA, pp. 110-121, 1991.
