Poster: Assured deletion on Public Clouds

Kopo M. Ramokapane
Security Lancaster Research Centre
Lancaster University
United Kingdom
k.ramokapane@lancaster.ac.uk

Awais Rashid
Security Lancaster Research Centre
Lancaster University
United Kingdom
marash@comp.lancs.ac.uk

Jose M. Such
Security Lancaster Research Centre
Lancaster University
United Kingdom
j.such@lancaster.ac.uk

I. Introduction

With the advent of cloud computing, concepts such as Storage as a Service (StaaS) have emerged and become popular especially on public Clouds. This concept involves having a huge amount of storage space and offering it to many different customers. The advantage for customers is that they are not involved in any management of the underlying infrastructure. Other attracting features of Cloud storage include elasticity, pay-as-you-go and data redundancy. Data is replicated over different physical locations to ensure accessibility even in cases of natural disasters. However, while there are benefits, cloud storage also presents some novel challenges. Inadvertent exposure of sensitive data is still a major concern for potential cloud customers. While research [1],[2] has looked into the issue of side channel attacks and various other types of attacks exploiting virtualization, security issues arising from insecure deletion have not been considered to date in this context.

Contractual and service level agreements play a vital role in terms of how services will be offered but this requires tenants to trust their providers without technical assurances. It is important for cloud providers to technically verify to tenants that their data has been disposed of securely. Secure deletion means removal or destruction of data from disks such that it can no longer be recovered or considered to have any useful meaning. Issues of insecure deletion are well-understood in non-cloud contexts [3]; nonetheless, the nature of Cloud computing brings additional challenges with regards to guarantees on secure deletion. A number of cloud computing features such as multi-tenancy, virtualization, elasticity and data backup pose various challenges with regards to providing deletion assurances to cloud tenants.

The following section discusses how the above cloud features make secure deletion guarantees a challenge for Cloud Service Providers (CSPs).

II. Challenges and vulnerabilities

A. Virtualization

Unlike in traditional computing, public cloud tenants are offered virtualized storage that may be scattered across the physical infrastructure. Storage virtualization hides the physical details of the actual storage device from tenants. It also introduces additional multiple layers as shown in Fig. 1 which, may cache data as it passes through. Data lineage in a virtualized environment is hard to trace since data flow is not linear. It is therefore difficult for CSPs to guarantee deletion from these logical layers and addresses.

B. Multi-tenancy

Cloud Computing is a multi-tenant environment; tenants share resources from the same physical device. When a normal delete operation is executed or a tenant is de-provisioned the previously held space is made available for reuse by other tenants. These pose a threat to data that may have been left by the previous owner. A malicious new cloud user may not write anything on their allocated space but rather inspect the storage for any data remnants. This could lead to data leakage. It is also not practically easy to securely delete data from a physical storage device while still in use by other tenants.

C. On-Demand Elasticity

Cloud Computing also offers on-demand elasticity. When a tenant needs more storage space or processing power there is likelihood that a particular virtual instance would be moved to another physical location in the cloud where the needed resources can be supplied. This process is called live migration. Most live migrations are likely to happen randomly; a tenant's virtual service might migrate to the next convenient host that is available without any systematic control. Providing assured deletion in such environments is a difficult task since data remnants may be scattered all over the infrastructure.

D. Backup and Availability

Data snapshots are replicated all over the cloud to provide fault-tolerance and high availability. These snapshots are stored permanently and usually have a long retention period. Unfortunately there are times when deleting such snapshots is what the tenants desire. Cloud providers have no central point of control to guarantee deletion to tenants, i.e. all backup copies have been expunged from their cloud within a required time. Some cloud providers use subcontracts or offline storage media for backup purposes. While this could be of benefit, assuring deletion in such layouts and arrangements is a challenge.

III. Our contribution

In this poster, we present a model for assured data deletion (ADD) that provides probabilistic proof that tenant's data has been completely deleted from CSP's infrastructure. Our model is fitting in that it allows the CSP to generate a proof verifying that data has been completely deleted from its storage infrastructure; which is in contrast to existing approaches whose delete guarantees are dependent on trust and SLAs between service providers and clients.

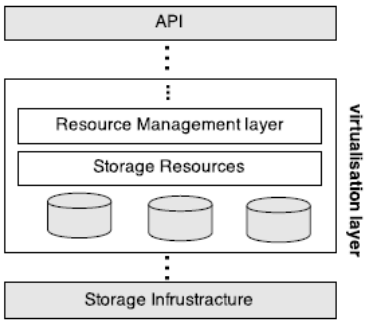

Fig. 1. A high level architecture of many layers that result from virtualization.

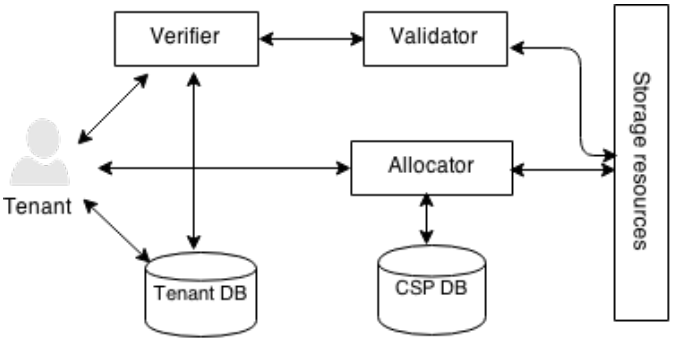

Fig. 2. Conceptional Assured Data Deletion(ADD) Architecture

Design and Assumptions: Fig. 2 depicts our model, which includes a verifier, allocator, a validator and two database storages. A verifier challenges the CSP about its deletion claims while the validator's purpose is to provide results to the verifier. The allocator joins location metadata to tenant's data blocks. It also sends signed computations back to the tenant. The tenant database stores signed metadata from the CSP.

We make standard assumptions that the verifier, validator and the allocator are all trustworthy and that the scheme is protected from any type of disabling or tempering. All communication between components is assumed to be secure.

Preliminaries: In this discussion, we use *T, P, D, L*, D_m and L_m to denote tenant, service provider, data block, Location, data block metadata and location address metadata respectively.

Storing data: Before *D* is uploaded to the cloud storage, D_m of *D* is stored in tenant database T_d. Upon receiving the *D*, *P* will attach L_m to *D*. L_m simply contains the information about the location address where *D* is stored in the cloud. After storing the *D*, *P* would compute a function of D_m and L_m and send response R_s[*stored*]=[D_m,L_m] back to *T* which will be stored in T_d. If *L* changes then *P* makes a new computation of D_m and L_m and the new output is sent back to *T* to update its records.

Data deletion: For simplicity, we describe the protocol when *T* wants to delete part of its data from the storage. It should be noted that this does not include deleting data from backup media. Suppose tenant *T* wants to delete *D* stored at location *L* from the cloud, *T* will send delete request R_q[*delete*]=[D_m,L_m] to *P*. *P* will attempt deletion operation and send response R_s[*deleted*]=[D_m,L_m,*true*] to *T* if the operation completed successfully. *T* will store this information into T_d.

Proof of deletion: To verify *P*'s assertion, *T* will challenge *P* with a random question relating to the disputed data block. A sample challenge might be "Compute D_m and L_m for address *L*" or "return D_m of *D* stored at *L*". The validator would make the computation and send a response R_s[*output*]=[D_m,L_m] back to *T*. The verifier would also make the same computation using information from T_d and produce output T_o=[D_m,L_m] and then both output results are compared. Claim of deletion from the infrastructure is valid only when both results are not the same. If *L* is empty or holds a different data block the computation should return a different computation therefore proving that previous data held there has been removed.

IV. Conclusion

While cloud computing is continuing to revolutionize the computing world, providing guarantees for outsourced data is continuing to be a challenge. Most of such verifications rely heavily on SLAs and contracts but contractual agreements cannot technically give evidence that an expected operation has been performed accordingly. One of these challenges is providing assured deletion to cloud tenants.We have highlighted and discussed cloud computing features that make secure deletion guarantees in public cloud storage difficult. We have also proposed a solution which attempts to provide deletion assurances to tenants.

Although we believe that our architecture represents a significant step forward in providing deletion assurances to tenants, there are some limitations which our designed architecture cannot handle. Firstly, our architecture has only been designed to provide assurances to tenants when they only want some part of their data deleted from the cloud infrastructure. Secondly, it does not provide deletion assurances for data hosted at subcontract sites and offline storages. These limitation are not necessarily fundamental, but they require additional research to overcome.

V. Future Work

Our future work includes evaluating our architecture using two different criteria: generality and performance overhead. To evaluate the generality of our architecture, we will assess whether the deletion verification works successfully. Performance overhead observations will include evaluating (i) deletion time: time taken by the system to delete tenant's data and notify the tenant, (ii) upload time: time taken to upload data and save metadata in both databases and lastly (iii) Verification time; time taken to verify deletion to tenant.

References

[1] T. Ristenpart, E. Tromer, H. Shacham, and S. Savage, "Hey, You, Get Off of My Cloud: Exploring Information Leakage in Third-Party Compute Clouds," *Ccs'09: Proceedings of the 16th Acm Conference on Computer and Communications Security*, pp. 199-212, 2009

[2] Y. Tang, P. P. C. Lee, J. C. S. Lui, and R. Perlman, "FADE: Secure Overlay Cloud Storage with File Assured Deletion," *Security and Privacy in Communication Networks*, vol. 50, pp. 380-397, 2010.

[3] J. Reardon, D. Basin, and S. Capkun, "SoK: Secure Data Deletion," *2013 Ieee Symposium on Security and Privacy (Sp)*, pp. 301-315, 2013