Chapter 9
Securing Cloud Storage

Jacques Jorda
Institut de Recherche en Informatique de Toulouse, Université Paul Sabatier, France

Abdelaziz M’zoughi
Institut de Recherche en Informatique de Toulouse, Université Paul Sabatier, France

ABSTRACT

Data storage appears as a central component of the problematic associated with the move of processes and resources in the cloud. Whether it is a simple storage externalization for backup purposes, use of hosted software services or virtualization in a third-party provider of the company computing infrastructure, data security is crucial. This security declines according to three axes: data availability, integrity and confidentiality. Numerous techniques targeting these three issues exist, but none presents the combined guarantees that would allow a practical implementation. The authors’ solution relies on the integration of these techniques to a virtualization middleware. Quality of service definition allows specifying the nature of the security to implement with a seamless access.

INTRODUCTION

Cloud storage can be used in many ways. First, it can serve as a data backup, such as Apple’s iCloud service, or Dell Datasafe online storage service. Even Amazon’s Simple Storage Service (also known as Amazon S3) can be used in that way. It can also be used to store used by cloud computing services: both services accessible via internet (Software as a Service (SaaS), such as those provided in the Microsoft Azure offer) and server virtualization require cloud storage to store data or virtual filesystems disks.

This cloud storage has all the advantages of cloud computing: only what is used has to be paid, without heavy initial investment or hidden fees. It also has the same pitfalls: data security being critical, the confidence in the operator abilities to ensure this security must be complete.

In the next section, we will present the different possible security vulnerabilities relative to cloud data storage, and the way to tackle them. These vulnerabilities are classical (data leakage due to an external user) or specific to cloud computing
Securing Cloud Storage

issue (data leakage due to a cloud administrator). We will see how to ensure data security in this context. Classical encryption techniques can be used in conjunction with data distribution services linked to the storage resource nature found in these environments. Track access methods can complete these security techniques to detect unauthorized accesses, including by the host infrastructure managers. We will conclude this section by showing why no global solution is available at this time.

In the following section, we will introduce the principle of storage virtualization. We will present the architecture of our storage virtualization middleware, and show how it can handle the problematic of cloud storage security. This virtualization can be implemented at the host infrastructure to provide virtual storage resources to the applications or to user virtual machines disks; it can be extended to user virtual machines to provide secure and seamless accessible backup storage volumes.

BACKGROUND

When data are moved in the cloud, the storage security is critical. Because data are no longer managed by their owner, the latter must be ensured that security is maintained. This security is defined by three main parameters: confidentiality, integrity and availability. These three parameters are used to characterize threats likely to affect the data.

Data availability is crucial. Storage provider must ensure that the data will be available independently of what could happen, by committing fast turnaround time subjected to a penalty. Providers such as Amazon with its Simple Storage Service are likely to be unavailable for several hours. When it happens, the user computing infrastructure becomes partly ineffective (when only few services are hosted by Saas solutions) or completely ineffective (when the all user server infrastructure is virtualized and hosted by the provider), leading to significant losses. To always ensure data availability, the provider must thus implement not only redundancy solutions but also backups.

Data integrity issue relies on the guarantee sought by the user that all his externalized data are effectively present and non-altered in its provider infrastructure. Data integrity damages may results from malicious third party attacks, hosting infrastructure vulnerabilities, or conscious choice of the hosting infrastructure to delete non-accessed data to maximize its costs. This issue begins to be well studied in the literature with, for example, the use of POR (Proof Of Retreivability) and PDP (Proof of Data Possession) tools. These techniques allow detecting data integrity damage without requiring local data copy storage by the user (the aim being exactly to externalize the storage).

Data confidentiality remains one of the main concerns and the major barrier to the development of cloud services. This confidentiality is vulnerable to conventional threats (injection attacks, cross-site scripting…) but also to specific cloud computing threats (hypervisor flaws, management of the security perimeter within a company, confidence in the provider). At this level, differences can be separated according to the user infrastructure externalization degree in the provider infrastructure:

- The user uses the application server hosting (database, web services, mail servers…), these applications being accessed by the user with secured connexions. Usually, it involves a minimal user infrastructure, because some applications can not be placed in the cloud and desktop application data remain within the own user infrastructure – and therefore require, at least, both authentication and file servers. Here, data protection is complex because the application is managed by the provider.
- The user virtualizes its servers (and when necessary the desktops). Here, the all user server infrastructure is moved to the cloud and all his data are hosted on dedicated vir-