Cryptographic Cloud Storage

Seny Kamara and Kristin Lauter

Microsoft Research
{senyk,klauter}@microsoft.com

Abstract. We consider the problem of building a secure cloud storage service on top of a public cloud infrastructure where the service provider is not completely trusted by the customer. We describe, at a high level, several architectures that combine recent and non-standard cryptographic primitives in order to achieve our goal. We survey the benefits such an architecture would provide to both customers and service providers and give an overview of recent advances in cryptography motivated specifically by cloud storage.

1 Introduction

Advances in networking technology and an increase in the need for computing resources have prompted many organizations to outsource their storage and computing needs. This new economic and computing model is commonly referred to as cloud computing and includes various types of services such as: infrastructure as a service (IaaS), where a customer makes use of a service provider's computing, storage or networking infrastructure; platform as a service (PaaS), where a customer leverages the provider's resources to run custom applications; and finally software as a service (SaaS), where customers use software that is run on the providers infrastructure.

Cloud infrastructures can be roughly categorized as either private or public. In a private cloud, the infrastructure is managed and owned by the customer and located on-premise (i.e., in the customers region of control). In particular, this means that access to customer data is under its control and is only granted to parties it trusts. In a public cloud the infrastructure is owned and managed by a cloud service provider and is located off-premise (i.e., in the service provider's region of control). This means that customer data is outside its control and could potentially be granted to untrusted parties.

Storage services based on public clouds such as Microsoft's Azure storage service and Amazon's S3 provide customers with scalable and dynamic storage. By moving their data to the cloud customers can avoid the costs of building and maintaining a private storage infrastructure, opting instead to pay a service provider as a function of its needs. For most customers, this provides several benefits including availability (i.e., being able to access data from anywhere) and reliability (i.e., not having to worry about backups) at a relatively low cost.

While the benefits of using a public cloud infrastructure are clear, it introduces significant security and privacy risks. In fact, it seems that the biggest hurdle

R. Sion et al. (Eds.): FC 2010 Workshops, LNCS 6054, pp. 136–149, 2010.

to the adoption of cloud storage (and cloud computing in general) is concern over the confidentiality and integrity of data. While, so far, consumers have been willing to trade privacy for the convenience of software services (e.g., for web-based email, calendars, pictures etc), this is not the case for enterprises and government organizations. This reluctance can be attributed to several factors that range from a desire to protect mission-critical data to regulatory obligations to preserve the confidentiality and integrity of data. The latter can occur when the customer is responsible for keeping personally identifiable information (PII), or medical and financial records. So while cloud storage has enormous promise, unless the issues of confidentiality and integrity are addressed many potential customers will be reluctant to make the move.

To address the concerns outlined above and increase the adoption of cloud storage, we argue for designing a *virtual private storage service* based on recently developed cryptographic techniques. Such a service should aim to achieve the best of both worlds by providing the security of a private cloud and the functionality and cost savings of a public cloud. More precisely, such a service should provide (at least):

- confidentiality: the cloud storage provider does not learn any information about customer data
- integrity: any unauthorized modification of customer data by the cloud storage provider can be detected by the customer

while retaining the main benefits of a public storage service:

- availability: customer data is accessible from any machine and at all times
- reliability: customer data is reliably backed up
- efficient retrieval: data retrieval times are comparable to a public cloud storage service
- data sharing: customers can share their data with trusted parties.

An important aspect of a cryptographic storage service is that the security properties described above are achieved based on strong cryptographic guarantees as opposed to legal, physical and access control mechanisms. We believe this has several important benefits which we discuss further in Section 3.

This article is organized as follows. In Section 2 we describe, at a high level, a possible architecture for a cryptographic storage service. We consider both consumer and enterprise scenarios. We stress that this design is not intended to be a formal specification (indeed many important business and engineering questions would need to be addressed) but is only meant to serve as an *illustration* of how some of the new and non-standard cryptographic techniques that have been developed recently could be combined to achieve our goals. In Section 3 we give an overview of the benefits of a cryptographic storage service, e.g., reducing the legal exposure of both customers and cloud providers, and achieving regulatory compliance. In Section 4 we describe in more detail the relevant cryptographic techniques, including searchable encryption, proofs of storage and attribute-based encryption. Finally, in Section 5, we mention some cloud services that could be built on top of a cryptographic storage service such as secure back-ups, archival, health record systems, secure data exchange and e-discovery.