TREDISEC: Trust-aware REliable and Distributed Information SEcurity in the Cloud

TREDISEC Consortium

Cloud computing services are increasingly being adopted by individuals and companies thanks to their various advantages such as high storage and computation capacities, reliability and low maintenance costs. Yet, data security and user privacy remain the major concern for cloud customers since by moving their data and their computing tasks into the cloud they inherently lend the control to cloud service providers. Therefore, customers nowadays call for end-to-end security solutions in order to retain full control over their data.

Implementing existing end-to-end security solutions unfortunately cancels out the advantages of the cloud technology such as cost effective storage. For example, cloud storage providers constantly look for techniques aimed to maximize space savings. One of the most popular techniques that has been adopted by many major providers to minimize redundant data is data deduplication. Unfortunately, deduplication and encryption are two conflicting technologies. Two identical data segments become indistinguishable after being encrypted. This means that if data are encrypted by users in a standard way, the cloud storage provider cannot apply deduplication.

In TREDISEC, we aim at designing new security primitives that not only ensure data protection and user privacy but also maintain the cost effectiveness of cloud systems. With this goal, we will first identify the functional requirements that are crucial to the cloud business and explore non-functional requirements such as storage efficiency and multi-tenancy. We will further analyze the conflicts between these requirements and security needs in order to develop new solutions that address these shortcomings and enhance security. The main challenges resulting from the combination of security, functional and non functional requirements, and which TREDISEC aims at resolving thanks to the newly designed primitives, are the following:

- **Data confidentiality with data reduction**: As already mentioned in the introduction, storage efficiency functions such as deduplication or compression become ineffective when data is encrypted. While there exist few solutions based on convergent encryption, these either do not achieve acceptable levels of security or rely on the existence of a fully trusted third party. TREDISEC's new primitives will provide stronger data confidentiality guarantees while benefiting from the various advantages of data reduction techniques in the cloud.
- **Secure data processing with multi-tenancy**: In order not to cancel out the performance advantages of the cloud, there is a strong need for privacy preserving data processing solutions. Classical encryption solutions prevent the cloud from operating over encrypted data. Among data processing primitives, word search is one of the most fundamental operation. Existing privacy preserving word search solutions are not yet directly applicable in real industrial strength use cases. TREDISEC will extend current solutions with advanced features such as the ability to delegate

search operations to authorized third parties and consider the multi-tenant environment whereby a large number of tenants outsource their data and computation to the cloud.

- **Verifiability with data reduction and multi-tenancy**: Since data storage and processing operations are performed remotely by potentially malicious clouds, end-users should receive some guarantees on both the storage and processing of data. Existing storage integrity solutions rely on Proofs of Retrievability (PoR) which provide the end-user with the assurance that a data segment is actually stored in the remote storage; these solutions still induce high computational costs and cannot be combined with data reduction techniques. TREDISEC will enable the verification of data retrievability while data is deduplicated or compressed. Furthermore, data reduction techniques usually imply the storage of a single copy of data for several users; in this case, users having already outsourced data should prove their actual ownership. Existing Proof of Ownership (PoW) solutions are unfortunately not yet mature enough in terms of both performance and security. TREDISEC will design efficient and secure PoW schemes where deduplication takes place among multiple tenants and over encrypted data. Additionally, TREDISEC will also investigate existing processing verifiability solutions such as probabilistically checkable proofs in order to provide end-users with some cryptographic tools that efficiently verify the correctness of some dedicated functions.
- **Distributed enforcement of access control policies for multi-tenancy settings**: The security of a multi-tenant system require reliable access control polices and enforcement mechanisms. Current Attribute Based Access Control (ABAC) models fall short in the multi-tenancy settings since users' attributes can be distributed across different trust domains. TREDISEC will extend current ABAC models to govern access control in multi-tenant cloud storage systems. Furthermore, current cloud platforms are agnostic to the concept of shared ownership. TREDISEC will design new cryptographic primitives to enforce distributed usage of data while preventing malicious tenants from combining their credentials and escalating their access rights. TREDISEC will also investigate the problem of secure data deletion: end-users will have have cryptographic guarantees on the timely deletion of their data and the back-up copies.

The ultimate goal of TREDISEC is to converge to a unified framework where these primitives are integrated and where all objectives are satisfied to the highest extent possible. With this goal, we will explore different architectural models while following the end-to-end security principle as closely as allowed by functional and non-functional requirements. The resulting TREDISEC framework will be evaluated across several different use cases.