EPFL

Thèse n° 7391

Integrity and Metadata Protection in Data Retrieval

Présentée le 26 novembre 2021

Faculté informatique et communications
Laboratoire de systèmes décentralisés et distribués
Programme doctoral en informatique et communications

pour l'obtention du grade de Docteur ès Sciences

par

Kirill NIKITIN

Acceptée sur proposition du jury

Prof. J.-P. Hubaux, président du jury
Prof. B. A. Ford, directeur de thèse
Prof. J. Cappos, rapporteur
Prof. S. Capkun, rapporteur
Prof. K. Argyraki, rapporteuse

École
polytechnique
fédérale
de Lausanne

2021

Abstract

Secure retrieval of data requires integrity, confidentially, transparency, and metadata-privacy of the process. Existing protection mechanisms, however, provide only partially these properties: encryption schemes still expose cleartext metadata, protocols for private information retrieval neglect data integrity, and data-distribution architectures forego transparency. In this dissertation, by designing new cryptographic primitives and security architectures that provide a more comprehensive protection, we improve on the current security and privacy practices in data retrieval. First, we propose a new format for encrypted data; it protects both content and all encryption metadata, such as the application, the intended recipients, and the algorithms used. The format comes with a cryptographically-agile encoding scheme that facilitates efficient decryption of such ciphertexts without cleartext markers. Second, to address the lack of integrity in privacy-preserving data-retrieval protocols, we introduce the concept of single-server verifiable private information retrieval. In contrast to existing solutions where, in some deployment scenarios, a malicious server can violate client privacy by selectively tampering with the data, our approach ensures that an honest client either correctly obtains the data from the system's server or detects server misbehavior and aborts. Finally, we present a software-update framework that reinforces software-distribution processes. Building on the concepts of decentralization and verifiability, our framework eliminates single points of failure, enforces transparency, and ensures integrity and authenticity of software releases. By implementing and experimentally evaluating our primitives and framework, we demonstrate that better protection is practical and incurs only a modest additional cost.

Keywords: privacy, integrity, security, metadata protection, private information retrieval, verifiable, software updates, transparency.