Open Learning Forum

ProvChain: A Blockchain-based Data Provenance Architecture in Cloud Environment with Enhanced Privacy and Availability

Xueping Liang; Sachin Shetty; Deepak Tosh; Charles Kamhoua; Kevin Kwiat; and Laurent Njilla
Tennessee State University; Old Dominion University; Norfolk State University; and Air Force Research Laboratory
Date Published: 5/2017

Executive Summary written by GBBC

Introduction

Cloud computing has grown increasingly popular because of its data storage capabilities. However, the nature of cloud computing makes it uniquely vulnerable. To address security concerns, cloud systems must undergo frequent audits; these audits are only useful if administrators are working with accurate and tamper-free records. Therefore, cloud data provenance, the metadata generated by changes to cloud data, is essential to the security of a cloud computing system. Unfortunately, provenance data is also susceptible to tampering and could reveal sensitive underlying data.

ProvChain

Blockchain seems a perfect technology to create a distributed and highly secure provenance database. In the paper, the researchers present ProvChain, "a blockchain based data provenance architecture to provide assurance of data operations in a cloud storage operation, while enhancing privacy and availability at the same time." ProvChain monitors user operations and collects provenance data; time-stamped user operations are then published to the blockchain. A blockchain receipt is created for each provenance entry so the auditor can validate operations during an audit. User IDs are hashed to preserve privacy; auditors cannot identify the true owners of data, this can only be done by the service provider.

Conclusion

The researchers tested ProvChain using a provenance-enabled ownCloud application. Provenance service "brings an average of 6.49% of total overhead against original ownCloud application in terms of response time," which is a relatively small price to pay for vastly improved security. They found that provenance retrieval results in an average time cost of 0.838 seconds for file creation, 0.676 seconds for file changes, and 0.79 seconds for file sharing. Overall, ProvChain does not result in a significant overhead increase, though service providers will have to determine an appropriate way of compensating blockchain miners. Ultimately, ProvChain and other blockchain solutions for cloud computing may prove essential for institutions that prioritize security.