JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Improving Reliability in Cloud Data Storage and Sharing

Mr. Mule BalaEswar Reddy (M.C.A). Rajeev Gandhi Memorial college Of Engineering and Technology, Nandyal

Mr. O.Sampath Kumar MTech (Ph.D). Rajeev Gandhi Memorial college Of Engineering and Technology, Nandyal

Abstract

Each business in the group creates a valued digital signature. a certificate on the ciphertext of this digital data is then issued using the raw data that was received by the data owner. The data owner uploads his or her data to the Inter-Planetary File System network according to the data storing format, and after that, the access address of the data as well as the related certificate are saved on the blockchain ledger. Before submitting a data sharing request to the data owner, everyone on the system could use the data sharing schema to confirm the validity of shared data. A smart contract is used to manage the data sharing process, and escrow is required of all parties to promote trustworthiness. The security qualities comprising confidentiality, integrity, privacy, non-repudiation, and anonymity are guaranteed by the data storage and sharing protocols.

1. INTRODUCTION

1.1 Introduction

Worldwide data development has been exponential, and reliable data are among the most useful resources for both individuals and businesses. By 2025, it is expected that the total amount of data generated and stored worldwide will reach 175 zettabytes. Additionally, it is predicted that 5 billion people would be using data every day by the year 2025. Due to the high need for valuable data storage and sharing, there are also significant data security risks associated with these procedures. Currently, centralized and decentralized systems are the two basic types of data storage and sharing.Organizations can store data on their data center system for the centralized design. These systems can only be scaled up so far, and their operational expenses are substantial. Utilizing cloud storage services can lower costs, allow for flexible system growth, and make IoT systems more suited. The use of encryption techniques and access control mechanisms is suggested to safeguard the security and privacy of data storing and sharing. SECUREDL was proposed by [12] as a means of safeguarding private information kept in databases. The centralized architecture, however, has two drawbacks, namely: (1) availability, when the centralized systems crash due to system overload, denial-of-service or distributed denial-of-service (DoS/DDoS) attacks, or system errors, (2) data security, where stored data may be accessed, modified, or removed without authorization by system administrators or attackers who have breached the system.Because of its attributes including anonymity, transparency, decentralization, and auditability, blockchain (BC) technology serves as the main component of the systems used for the majority of decentralized architecture solutions. The shared data on the BC network cannot

currently be verified for accuracy and dependability using current technologies. Meaningful data (MD) is specifically defined as data that has been validated and certified by a reputable organization (RO). For instance, MD, a respectable medical institution with highly qualified doctors, publishes a diagnostic result from an electronic medical record. In the sphere of education, an MD lecture is one that has been evaluated and approved by a professional board of a respected university. In addition, MD must be stored on the system in a secure manner.

2. Literature Survey

• **Cloud Databases for Internet-of-Things Data:**, The so-called Internet of Things (IoT) was made possible by the advancement of pervasive computing, radio frequency identification (RFID) technology, and sensor networks. The core tenet of the Internet of Things is that interconnected "smart" things will create a worldwide infrastructure for information and communication that will enable the development of value-added services. A vast, dynamic, and expanding network of networks, encompassing billions of items like sensors, actuators, RFID tags, computers, mobile devices, and more, is how the Internet of Things is envisioned. A significant amount of data enters the network as a result of these devices producing and exchanging data at the same time. The major objective of the studies was to evaluate how well various cloud databases for IoT data performed in comparison to one another. As a result, the database servers were set up on a cloud-based Amazon EC2 instance.Four open-source databases—MySQL, MongoDB, CouchDB, and Redis—were used for the studies. We selected the databases from ones that were both well-liked and typical of their category (e.g., both SQL and NoSQL). Scalar sensor data and multimedia data were the focus of the examination for two distinct benchmarks. The aforementioned data kinds are anticipated to span a wide range of scenarios in terms of quantity, data formats, and applications given the diversity of Internet-of-Things data.

[2] Survey of Real-time Processing Technologies of IoT Data Streams: Considerable attention has been paid to the Internet of Things (IoT) technology, which links numerous physical objects to the Internet. By 2022, 50 billion devices would be connected to the Internet, Cisco said in a white paper, generating 14.4 trillion dollars in income. By 2020, 28 billion IoT devices will be installed, and the market will generate 700 billion dollars in annual revenue, according to IDC. In Gartner's hype cycle for 2014 and 2015, IoT was regarded as the most anticipated technology. According to METI in Japan, IoT will support a data-driven society in which the digital data gathered by IoT will have value added and be advantageous to society. IoT technology will have a significant impact on markets as well as For each iteration of stochastic gradient descent, a small number of samples are chosen at random as opposed to the entire data set. The number of samples from a dataset that are utilized to calculate the gradient for each iteration is referred to as the "batch" in the Gradient Descent algorithm. The batch is assumed to be the entire dataset in normal Gradient Descent optimization techniques like Batch Gradient Descent. The issue emerges when our dataset gets large, despite the fact that using the entire dataset is highly helpful for reaching the minima in a less noisy and less random manner. If you utilize a standard Gradient Descent optimization technique and your dataset has a million samples, you will need to employ every single one of them for the optimization.

3. OVERVIEW OF THESYSTEM

3.1 Existing System

The current system allows for data sharing and data production, but the solutions lack the ability to validate the correctness and dependability of shared data on the BC network. Data that has been validated and accepted by a reputable organization (RO) (MD) is considered meaningful data. The need for data storage and sharing is essential for approved digital data, which has the following drawbacks:

3.1.1 Disadvantages of Existing System

- Less feature compatibility
- Low accuracy.

3.2 Proposed System

Our suggested plan for gathering, storing, and disseminating information. In the data generating structure for a collection of respectable businesses that offer the same kind of service, we employ a group authentication mechanism. One of the organizations in the group uses raw data given by a data owner to create valuable online information, and based on the encrypted message of this digital information, concerns a certificate..

3.3 Methodology

In this project work, I used five modules and each module has own functions, such as:

1. Data Owner Module
2. Data Provider
3. Group Manager
4. Data user

1. DataOwner:

Login: Dataowner has to login with valid details which are used in his / her Registration
Register: Each and every Dataowner has to Register.
Upload files: uploads his files into the cloud.
View files : views all the files which are uploaded by the him/her and sends request if there is any requests.
Secured files: He can view his/her files are secured or not.
Logout :Finally logout from the system.

2. Data Provider:

Login: data provider has to login with valid details which are used in his / her Registration Register: Each and every data provider has to Register and management has to accept request.
View Data Owners Request: Views data owners requests and secure those files.
View Secured Files: data provider can view all files which are secured. User Requests: views user requests for files and sends keys to the users. Logout: Finally, logout from the system.

3. Group Manager:

Login: Manager will login with default details

Add providers: He adds the data providers and views all the data providers who are added by him

Logout: Finally, logout from the system.

4. Data User:

Register: User registers with details

Login: Data user will login with his credentials which are added in his/her registration.

View Files: He views all the files which are uploaded by the data provider.

Send Request: sends the request for file

My Files: views all the files which are accepted

Check Files: he needs hash codes to view the file content

Download file: if hash codes are valid the data will be decrypted and can download as txt file

Logout: Finally, logout from the system.

ALGORITHM 1: Greedy Heuristic

Data: Given a set of gateways with data items, a set of Mini-Clouds with a known capacity and a set of communication links among them

Result: Time taken to transfer all data items from the gateways and to replicate all data items

1 Initialization;

2 d let job J be composed of data item Dj, sourceSourcej, destination Destj and link Lj reflecting a transfer event;

3 d Let S be the set of all possible JOBs at any point intime.d For D data items, M mini- Clouds and Llinks/gateways, S will have D M L JOBs initially. repeat

4 Order S based on×ordering heuristic; foreach

J∈OBj S do

5 if Destj has no space for Dj then

6 Delete JOBj;

7 end

8 if Lj is not busy then

9 Schedule JOBj;

10 Delete all JOBk from S such that Dk = Dj and Sourcek = Sourcej;

11 Mark Lj busy;

12 end

13 end

14 In parallel, when any JOBj completes Mark Lj free;

15 delete JOBj;

16 If (Dj needs to be further replicated) Add newJOBk to S, such that Sourcek = Destj, Destk Sourcek and Destk gateway;

17 until S /= Ø;

4 Architecture

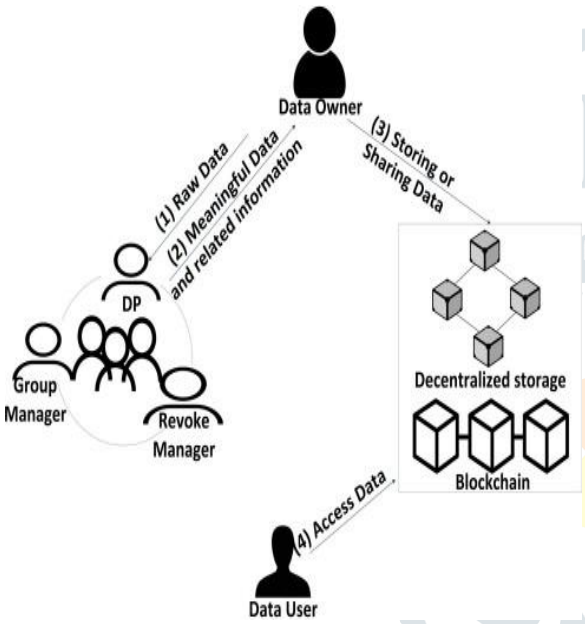
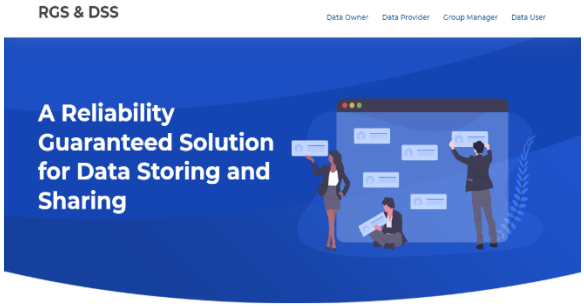


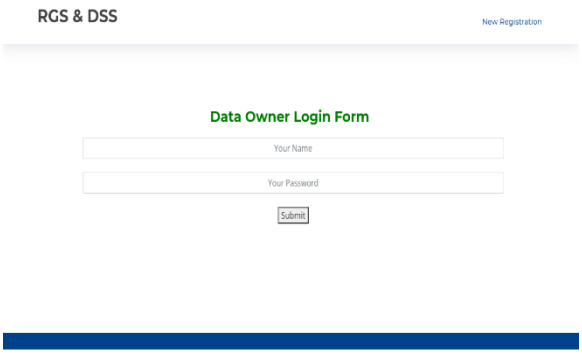
Fig 1: Frame work of proposed method

5 RESULTS SCREEN SHOTS

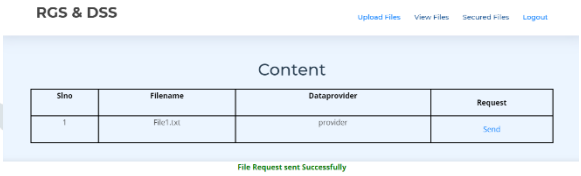
Home Page:



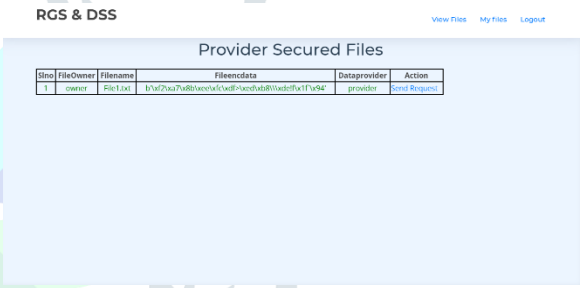
Owner Login:



Send Request:



Provider Secure Files:



7. CONCLUSION

✓ The three systems we suggest in this project are data producing, data storage, and data sharing. In the data production scheme, where RO is treated as a DP, a group manager organizes DPs that offer similar services into a group. A certificate on EMD is then issued by DP after it generates MD from RD that was sent by DO. We give the anonymity of DP and the privacy of DO, which have not been met by the existing solutions, in addition to the confidentiality and integrity of the stored data in the data storage scheme. Before submitting a request to DO for data sharing, everyone on the system can check the validity of provided data. Keep in mind that everyone can only confirm the accuracy of but cannot access the shared data's contents. The solutions currently available are unable to satisfy this property. Additionally, there are no intermediaries involved in the data sharing process because it is conducted directly between DO and DU.

The security analysis's findings demonstrate that the suggested methods satisfy the security requirements for secrecy, integrity, privacy, non-repudiation, and anonymity.

Future Enhancement

- ✓ In our upcoming work, we'll apply the suggested method to particular applications like the Internet of Things and electronic health records. The plans will then be assessed and improved. Additionally, there are no intermediaries involved in the data sharing process because it is conducted directly between DO and DU.

8. References

Serverless data pipeline approaches for IoT data in fog and cloud computing
Future Generation Computer Systems (IF 7.5) Pub Date: 2021-12-22 , DOI: 10.1016/j.future.2021.12.012
Shivananda R. Poojara, Chinmaya Kumar Dehury, Pelle Jakovits, Satish Narayana Srirama

Novel Fingerprinting Technique for Data Storing and Sharing through Clouds
Sensors (IF 3.9) Pub Date: 2021-11-17 , DOI: 10.3390/s21227647
Mehvish Fatima, Muhammad Wasif Nisar, Junaid Rashid, Jungeun Kim, Muhammad Kamran, Amir Hussain

Secure sharing of big digital twin data for smart manufacturing based on blockchain
Journal of Manufacturing Systems (IF 12.1) Pub Date: 2021-09-27 , DOI: 10.1016/j.jmsy.2021.09.014
Weidong Shen, Tianliang Hu, Chengrui Zhang, Songhua Ma

Automated Privacy Preferences for Smart Home Data Sharing Using Personal Data Stores
IEEE Security & Privacy (IF 1.9) Pub Date: 2021-09-06 , DOI: 10.1109/msec.2021.3106056
Yashothara Shanmugarasa, Hye-young Paik, Salil S. Kanhere, Liming Zhu

alidation of Architecture Effectiveness for the Continuous Monitoring of File Integrity Stored in the Cloud Using Blockchain and Smart Contracts
Sensors (IF 3.9) Pub Date: 2021-06-29 , DOI: 10.3390/s21134440
Alexandre Pinheiro, Edna Dias Canedo, Robson de Oliveira Albuquerque, Rafael Timóteo de Sousa Júnior

Data Trust Framework Using Blockchain Technology and Adaptive Transaction Validation
IEEE Access (IF 3.9) Pub Date: 2021-06-21 , DOI: 10.1109/access.2021.3091327
Sara Rouhani, Ralph Deters

Reliable Vehicle Data Storage Using Blockchain and IPFS
Electronics (IF 2.9) Pub Date: 2021-05-11 , DOI: 10.3390/electronics10101130
Hyoeun Ye, Sejin Park

DropStore: A Secure Backup System Using Multi-Cloud and Fog Computing
IEEE Access (IF 3.9) Pub Date: 2021-05-10 , DOI: 10.1109/access.2021.3078887
Reda Maher, Omar A. Nasr

Privacy Protected Blockchain Based Architecture and Implementation for Sharing of Students' Credentials
Information Processing & Management (IF 8.6) Pub Date: 2021-02-03 , DOI: 10.1016/j.ipm.2021.102512
Raaj Anand Mishra, Anshuman Kalla, An Braeken, Madhusanka Liyanage

Multi-Access Edge Computing Architecture, Data Security and Privacy: A Review
IEEE Access (IF 3.9) Pub Date: 2021-01-21 , DOI: 10.1109/access.2021.3053233
Belal Ali, Mark A. Gregory, Shuo Li

A Privacy-Preserving and Untraceable Group Data Sharing Scheme in Cloud Computing
IEEE Transactions on Dependable and Secure Computing (IF 7.3) Pub Date: 2021-01-11 , DOI: 10.1109/tdsc.2021.3050517
Jian Shen, Huijie Yang, Pandi Vijayakumar, Neeraj Kumar

On Blockchain-Enhanced Secure Data Storage and Sharing in Vehicular Edge Computing Networks
Applied Sciences (IF 2.838) Pub Date: 2021-01-04 , DOI: 10.3390/app11010414
Muhammad Firdaus, Kyung-Hyune Rhee

Data Trading Certification Based on Consortium Blockchain and Smart Contracts
IEEE Access (IF 3.9) Pub Date: 2021-01-01 , DOI: 10.1109/access.2020.3047398
Wei Xiong, Li Xiong.