Blockchain-based Database to Ensure Data Integrity in Cloud Computing Environments

Dr.B. Ravishankar,
Professor, IEM Department, B.M.S.
College of Engineering, Basavangudi,
Bengaluru 56004

Prateek Kulkarni,
Student, IEM Department, B.M.S.
College of Engineering, Basavangudi,
Bengaluru 56004

Vishnudas M V,
Student, IEM Department, B.M.S.
College of Engineering, Basavangudi,
Bengaluru 56004

***Abstract*—Data is nowadays an invaluable resource; indeed, it guides all business decisions in most of the computer-aided human activities. Threats to data integrity are thus of paramount relevance, as tampering with data may maliciously affect crucial business decisions. This issue is especially true in cloud computing environments, where data owners cannot control fundamental data aspects, like the physical storage of data and the control of its accesses. Blockchain has recently emerged as a fascinating technology which, among others, provides compelling properties about data integrity. Using the Blockchain to face data integrity threats seems to be a natural choice, but its current limitations of low throughput, high latency, and weak stability hinder the practical feasibility of any Blockchain-based solutions. In this paper, by focusing on a case study from the European SUNFISH project, which concerns the design of a secure by-design cloud federation platform for the public sector, we precisely delineate the actual data integrity needs of cloud computing environments and the research questions to be tackled to adopt Blockchain-based databases. First, we detail the open research questions and the difficulties inherent in addressing them. Then, we outline a preliminary design of an effective Blockchain-based database for cloud computing environments.**

***Keywords*—Block chain, Data Integrity, Cloud Based platform, Business Decision.**

I. INTRODUCTION

Data in these days is a key plus. It is strategic to drive any business call in numerous fields, starting from fagot and insurance, to health, education and public administration. As computer-aided human activities are relying additional and additional on knowledge, trusting knowledge has so become crucial. At constant time, the essential role of knowledge has created it an awfully appealing target for cyber-attacks, that aim at undermining the elemental independent agency properties that knowledge ought to manifest so as to be trustworthy.

Cyber-attacks against independent agency properties cause completely different damage on information trust according to the undermined property. Specifically, impairments accessibility prevents information to be regained only for temporary amount of your schedule. Compromising discloses the personal information and cannot be reverted back, but the original information square measure is still on the market. Instead, tampering with information probity could be extremely damaging attack that continuously has vital problems to data trust. Indeed, change of state with information will go unseen and drive operations maliciously, by deleting specific entries (i.e., to get rid of inconvenient traces) or by neutering specific sections of data (i.e., to alter information consumers' behavior). In 2015, Kaspersky laboratory has seen an enormous cyber-attack targeting over than one hundred monetary institutes worldwide that siphoned off cash from account balances for associate calculable worth of around $1 billion otherwise from confidentiality and availableness, once integrity is compromised there's no thanks to restore the first knowledge, it is lost forever. Therefore, as integrity attacks square measure delicate to be detected and extremely active, in this paper we have a tendency to concentrate on knowledge integrity instead of confidentiality or availableness. Data integrity problems square measure exacerbated in cloud computing environments, as knowledge homeowners hardly management wherever their knowledge square measure hold on, UN agency will really access them, and within which method. Nevertheless, a lot of and a lot of non-public and public organizations square measure outsourcing their knowledge, because it relieves the burden of maintenance price moreover because the overhead of storing knowledge locally" [10]. Therefore, guaranteeing knowledge integrity properties in cloud computing environments has become Associate in Nursing pressing got to address. Data integrity is usually assured by victimization, on one hand, cryptologic tools (i.e., digests, asymmetric keys) and, on the opposite hand, applicable knowledge replication ways. The cryptographic tools square measure so wont to sign single items of knowledge, so any formation attack is often promptly detected via cryptologic signature validations. Indeed, Associate in Nursing attack to effective would need the violation of the key keys, so to update knowledge signatures and circumvent the cryptologic integrity checks. These attacks square measure difficult to hold out, however once realized they are much undetectable. Therefore, it's extremely advocated to use applicable knowledge replication ways to make sure anyhow knowledge integrity. Copying and passing on the information over a group, critically alters the violation of data integrity: associate degree assaulter ought to compromise, while not being detected, all the copied information. This replication approach is wide adopted in observe, like, e.g., within the context of cloud computing environments, wherever there are a lot of distributed storage resources. However, although replication sure will increase the weightage for a roaring attack during a cloud setting, cloud providers themselves will conspire with attackers for simply violating information integrity. To impede these collusion attacks and to avoid trust on the guarantees claimed by cloud providers, we have a tendency to advocate associate degree innovative making use of the Blockchain technology to style and implement a dispensary and secure Blockchain-based info for cloud computing environments.

Despite the time-related obstacle, the intrinsic replication and distribution alternatives throughout this paper, we've got a

978-93-5406-901-7 © 2020 IEEE

bent to tend that once shed light-weight, by introducing several open analyses. Then, we have got to gift sensible analysis directions that lead due to active Blockchain-based on the one hand, we've got a bent to tend to elaborate on the issues we propose degree innovative Blockchain-based information that permits levelling study integrity we've to first focus in Section a pair of on a specific category of eventualities associated with SUNFISH, degree project concerning secure-by-design cloud federation, and to the information threats its case studies promptly. In Section three we've got a bent to tend to elucidate the Blockchain technology, its information integrity properties, and current limitations, whereas we've got time-dependent integrity, lack of performance and absence of stability. Bent to tend to stipulate in Section four the analysis inquiries vary to know active Blockchain-based databases. In Section five we tend to gift our answer and endeavor such queries and its application to the SUNFISH cloud federation case study.

II. ALGORITHM FOR INTEGRATED MD5 BASED INVERSE FUNCTION ENCRYPTION

```
Step1. Get string S & append padding Bits.
Step2. Append Length L.
Step3. Initialize MD Buffer MDB.
Step4. Process Message in 16-Word Blocks WB.
Step5. Take Shift,i,n as integer & Take str, str1, str2 as string.
Step6. Set str=WB, str1=””.
Step7. Set str to Lower case.
Step8. Set n=length.
Step9. Get array in ch1 [] from str to Char Array().
Step10. Take ch3, ch4 as character.
Step11. Shift=shf
Step12. Set i=0.
Step13. Repeat until i<n.
    a. if(Character.isLetter(ch1[i]))
        i. ch3=(char)(((int)ch1[i]*shift-97)%26+97)
    b. str1=str1+ch3
    c. otherwise if(ch1[i]==' ')
        i. str1=str1+ch1 [i];
```

The table and diagrams refers to the result after implementing the above algorithm.

TABLE I. INDICATES THE COMPARISON OF TRADITIONAL AND PROPOSED WORK

Security	Traditional	Proposed
Session level security	Yes	Yes
MD5 security	No	Yes
Multiplicative inverse	No	Yes
IP Tracing	Yes	Yes
Standard encryption	Yes	Yes
Authentication	Yes	Yes
Application Layer Security	Yes	Yes

TABLE II. PROBABILITY OF ATTACK IN TRADITIONAL AND PROPOSED TECHNIQUE

Work	Probability of attacks
Traditional Blockchain	0.015
Proposed Blockchain	0.01

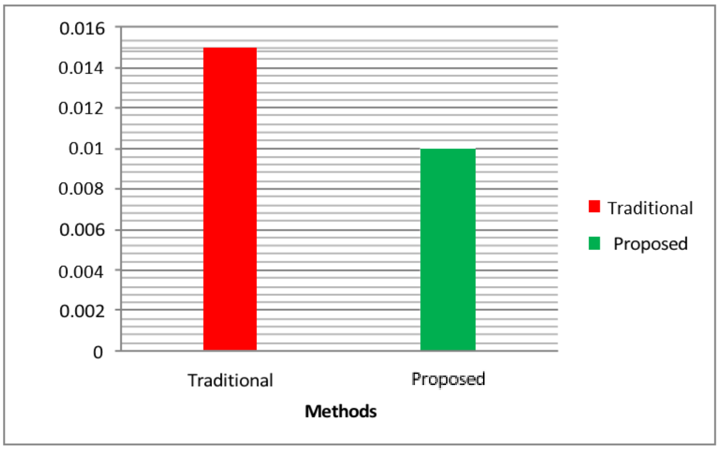


FIGURE1: COMPARISON OF PROBABILITY OF ATTACK IN TRADITIONAL & PROPOSED BLOCKCHAIN

III. CASE STUDY: EUROPEAN SUNFISH PROJECT

Nowadays, associate imperative would like public and personal firms to prompt and support and cooperate among their already deployed (private) cloud systems. Indeed, it's advocated that completely different cloud systems federate themselves into goal-oriented federations. Besides the multiple technical problems to handle, the creation and management of cloud federations got to face intimidating security problems, in the main associated with the non-disclosure of sensitive information and also the social control of integrity guarantees. to beat these security difficulties, the EU SUNFISH project aims at proposing a distributed, democratic cloud federation platform that will guarantee by-designing the safety of the managed information.

The SUNFISH proposal is Federation-as-a-Service (FaaS) [9], a replacement and innovative service that enables the secure creation and management of cloud knowledge and services. FaaS options advanced knowledge security services and innovative style principles resulting in a distributed and democratic cloud federation governance. For the sake of presentation, we tend to don't discuss the data security services, whereas we tend to extensively address the role of knowledge integrity in federation governance.

The intrinsic goal of cloud federations is sharing services among members by making regulated, secured inter-cloud interactions. the foundations governing these interactions, thus the service usage, square measure outlined in specific contracts. as an example, a member providing service could need that solely specific customers will use it which the service outputs got to be covert for privacy reasons. thanks to the high sensitivity of the information managed by cloud federations (e.g., personal and medical knowledge just in case of the general public sector), FaaS should give high assurances about the compliance of the member contracts. Indeed, besides the runtime implementation of the contracts, FaaS has got to guarantee the integrity of contracts, specifically that they cannot tamper with which all concerned members and should remember their existence. in addition, to ensure non-repudiable evidence of contract social control, all the inter-cloud interactions have to be monitored and therefore the logs hold on with robust integrity guarantees

Most of all, to foster a large adoption of cloud federations, FaaS advocates the absence of centralized governance. As a matter of truth, among federation members there cannot be a designed frontrunner (i.e., there's no primus repose pares), rather federation members type a network of peers. to the present aim, FaaS seeks to determine localized, democratic federation governance, therefore it should consider AN opportunely outlined, distributed information guaranteeing sturdy integrity guarantees. The novel style resolution for FaaS advocated by the SUNFISH project is based on the exploitation of a blockchain. To properly address the feasibleness of such an answer, significant threats to knowledge integrity ought to be known.

IV. DATA INTEGRITY THREATS AND ITS SOLUTION APPROACH

The Blockchain could be a quite novel technology that has appeared on the market within recent years, foremost used as a public ledger for the Bitcoin cryptocurrency. It primarily consists of consecutive enchained blocks containing records, that are replicated on the nodes of a p2p network. These records witness transactions that occurred between pseudonyms. Transactions might feature a cryptocurrency like, e.g., the Bitcoin, or different kinds of assets. the gathering of transactions and their intromission in chain blocks are applied in a very suburbanized fashion by distinguished nodes of the network, i.e. miners. Miners apply opportune block construction ways, i.e., the mining method, to realize agreement among all the miners on newly generated blocks. Bitcoin is an Associate in Nursing example of permission less Blockchain, i.e., there's no restriction for a node to become a miner. If instead there's Associate in Nursing authentication and authorization layer for miners, then the Blockchain is permissioned. The original mining method, still used for Bitcoin and Ethereum Blockchain, relies on the proof of labor (PoW). It consists of a very procedure intensive hashing task that's regulated according to the alleged Blockchain issue that regulates the typical time spent by miners to accomplish such a task and make a replacement block. Once a laborer achieves the creation of a replacement block, it broadcasts that block to all or any the opposite miners. They think about such a block because the latest of the chain and begin mining new blocks to be appended. For the sake of simplicity, we can say that when a laborer has created a replacement block, it becomes a part of the chain (if multiple miners concurrently add a block, a transient fork is formed that is sometimes quickly resolved as a result of by design miners forever think about the longest chain)

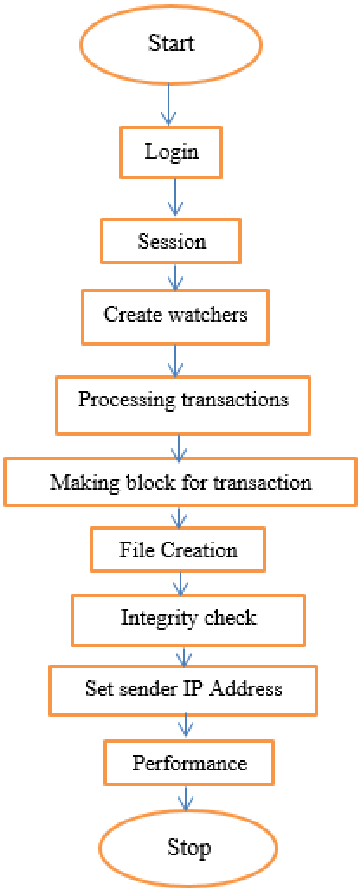
Now the solution approach,

The threat required for the associate aggressor to tamper the information (Threat T1) has been pre-previously mentioned throughout this section. At the smallest amount, all the miners of the first-layer Blockchain ought to be compromised, e.g. stealing their keys. at intervals the setting of FaaS, this could need offensive multiple distributed cloud suppliers at an equivalent time. though this unlikely state of affairs happens, the anchoring with the second-layer Blockchain ensures that alone the most recent set of operations on the knowledge are going to be subverted; all the others square measure tested by immutable, irreversible items of proof. The agreement formula featured by the first-layer Blockchain takes by choice into consideration all the miners, thence the member clouds. Therefore, there can't be any info operation completed whereas not all the members being tuned in to it (Threat T2). Collusion attacks (Threat T3) square measure instead corresponding to compromising first-layer Blockchain miners. All the private keys of these minor's square measure required to sign the messages needed to end AN info operation, thus even at intervals the case of 1 honest member attacked by a coalition fashioned by all the others, such honest member would possibly successfully react. Namely, it would forestall a malicious info operation to end by not deed its message within the agreement protocol. If the coalition attack was instead aimed to alter associated information already hold on at intervals the first-layer Blockchain, this implies that such information has been previously united on by all the members, thus associate honest member would possibly then prove it owns the intact version by showing the messages antecedently signed and sent by the opposite members.

V. EFFECTIVE BLOCKCHAIN-BASED DATABASE

In this section, we tend to tackle the analysis queries, which are simply raised and we define our proposal for more practical Blockchain-based information. This proposal is so meant to be a part of the SUNFISH project for guaranteeing high information integrity guarantees and, at an identical time, for being compliant with the performance and stability necessities required during a cloud computing setting. Our Blockchain-based information aims at providing a replicated information whose integrity is testified via adequate evidence hold on an innovative designed Blockchain system. Namely, we devise a two-layer Blockchain that, via the first-layer, ensures adequate performance and, via principled exploitation of the second-layer, ensures robust integrity guarantees. More specifically, the first-layer employs a light-weight distributed agreement protocol that assures low latency and high output. This layer aims at quickly and dependably storing pieces of evidence of every operation administrated on distributed information. However, this layer provides weak data integrity guarantees because of the dearth of a prisoner. Thus, the second layer is meant as a PoW-based Blockchain that stores shreds of evidence of (half of) the information operations logged by the first-layer. These pieces of evidence area unit hold on with robust information integrity guarantees however with poor performance. Indeed, the principled interaction between the 2 layers' permits getting associated overall performance improvement and effective assurances on information integrity.

The member clouds operate on the knowledge issue of the operations through the Information Inter-face. The operations area unit logged via applicable proof by the first-layer Blockchain, then they are dead on the distributed unit replicas. tons of space finally, the first-layer Blockchain is permissioned, and choices one jack on each member cloud. The miners, by hopping on a public/private key mix to sign messages, come back through agreement employing the therefore known as mining rotation agreement mechanism. Namely, it divides the time into rounds and, for every spherical, elects a jack as a pacesetter. The leader is then to blame for receiving new operations, linguistic communication them with its key, and broadcasting them to the alternative miners. Once all miners have signed the operations, they're going to become a district of the Blockchain: all the miners add these operations to their native ledger and apply them to their native duplicate. The interaction with the second-layer PoW-based Blockchain is completed via a Blockchain anchoring technique. The anchoring technique could also be an everyday operation that permits linking part of the first-layer Blockchain with (a block of) the second-layer Blockchain. especially, at positive intervals of it slow, a witnessed event containing the hash of the first-layer Blockchain up to the present operation is shipped to the second-layer Blockchain and, consequently, keep as the changeless, irreversible event. These hashes act as forensics proof for proving and validate the integrity of the data keep at intervals the first-layer Blockchain.



The above algorithm represents the Sample Propose Model for stable Blockchain based data-base

VI. CONCLUSION

In this paper we answer the we made the case study of European Sunfish project to realize Blockchain based database for cloud computing environment Our main contribution is the proposal of a high-level solution which answers these questions, and lies the foundations for the design of a Blockchain-based database able to provide the desired guarantees on data integrity, performance, and stability, later on we discussed about the proposal of stable Blockchain based database system. The approach surely permits to achieve integrity among then distributed replicas and to simplify the thread addressing. However, its availability can be critically affected by violating only a single miner.

VII. REFERENCES

[1] Giuseppe Ateniese, Randal Burns, Reza Curtmola, Joseph Herring, Lea Kissner, Zachary Peterson, and Dawn Song. Provable data possession at untrusted stores. In Proceedings of the 14th ACM conference on Computer and communications security, pages 598{609. ACM, 2007.

[2] Joseph Bonneau, Andrew Miller, Jeremy Clark, Arvind Narayanan, Joshua A Kroll, and EdwardW Felten. Sok: Research perspectives and challenges for bitcoin and cryptocurrencies. In 2015 IEEE Symposium on Security and Privacy, pages 104{121. IEEE, 2015.

[3] George Danezis and Sarah Meiklejohn. Centrally Banked Cryptocurrencies. In 23rd Annual Network and Distributed System Security Symposium, NDSS, 2016.

[4] ENISA. Security Framework for Governmental Clouds, 2015. Available at https://www.enisa.europa.eu/activities/Resilience-and-CIIP/cloud-computing/governmental-cloud-security/security-framework-for-govenmental-clouds.

[5] Ittay Eyal, Adem Efe Gencer, Emin Gun Sirer, and Robbert Van Renesse. Bitcoin-NG: A scalable Blockchain protocol. In 13th USENIX Symposium on Networked Systems Design and Implementation (NSDI 16), pages 45{59, 2016.

[6] Juan Garay, Aggelos Kiayias, and Nikos Leonardos. The Bitcoin Backbone Protocol: Analysis and Applications, pages 281{310. Springer Berlin Heidelberg, Berlin, Heidelberg, 2015.

[7] Trent McConaghy, Rodolphe Marques, Andreas Muller, Dimitri De Jonghe, Troy McConaghy, Greg McMullen, Ryan Henderson, Sylvain Bellemare, and Alberto Granzotto. BigchainDB: A Scalable Blockchain Database (DRAFT). 2016.

[8] Satoshi Nakamoto. Bitcoin: A peer-to-peer electronic cash system, 2008. Available at https://bitcoin.org/bitcoin.pdf.

[9] Francesco Paolo Schiavo, Vladimiro Sassone, Luca Nicoletti, and Andrea Margheri. FaaS:Federation-as-a-Service, 2016. Technical Report. Available https://arxiv.org/abs/1612.03937.

[10] Mehdi Sookhak, Abdullah Gani, Hamid Talebian, Adnan Akhunzada, Samee U. Khan, Rajkumar Buyya, and Albert Y. Zomaya. Remote Data Auditing in Cloud Computing Environments: A Survey, Taxonomy, and Open Issues. ACM Comput. Surv., 47(4):65:1{65:34, May 2015.

[11] Manisha,Dr.Jasvindra kaur, Improving data integrity using Blockchain technology volume 10 issue 1 pp .315-320 jan2018-jun2018.

[12] Edoardo Gaetani1, Leonardo Aniello1, Roberto Baldoni1, Federico Lombardi1, Andrea Margheri2, and Vladimiro Sassone2 Blockchain-based Database to Ensure Data Integrity in Cloud Computing Environments.