University of Piraeus

School of Information and Communication Technologies

Department of Digital Systems

Postgraduate Program of Studies

MSc Digital Systems Security

Master's thesis:

Hybrid Cloud Computing User Data Banks and Privacy Marketplace for SSI Systems (TI)

Supervisor Professor: Christos Xenakis

Name-Surname	Email	Student ID.
Masmanidis Vasilis	v.masmanidis@ssl-unipi.gr	mte2021

Contents

Abstract:

Customer credentials privacy has always been serious issues in the centralized platform such as cloud storage etc. credentials verification and dealing in these specific centralized ways could lead to the privacy breaches. Self-sovereign identity system using blockchain smart contract could ensure user privacy and better data control to do the basic operations. Casper has been discussed to address the privacy issues while using SSI and blockchain systems. The credentials and identity details are stored in the digital wallet through encryption which are hard to bypass by the cybercriminals to sabotage privacy. Casper employs Zero-Knowledge Proof mechanisms to verify the identity information from the credential proofs. In user data bank control over data is what needed the most which could by the implementation of SSI while using blockchain.

Introduction

Cloud computing provides the liberty of resource sharing, secure user data management and customer services making it unique in terms of commercial usage to get scalable business. Cloud computing deals in reducing cost, increase flexibility, and improve service innovation. But it is certain that everything comes at certain cost, so enjoying the benefits and perks of cloud computing it would also have some negative aspects such as interoperability, service level monitoring, compliance, security and privacy. In cloud computing the data is shared, dynamically scalable environment and resource pools, and considered as often virtualized. Cloud computing is a model that enable convenient access to the shared resources of computing including servers, storage operations and application via internet without interacting with the service providers. (Song, 2021)There should be Five basic characters of cloud computing enlisted by NIST, including on demand self-service, measured services, resource pooling, broad network access and elasticity. Encryption and decryption of private transaction that often happen in banks are need of an hour which should need to be done privately by ensuring and maintaining security and privacy. This could possibly be provided by the cloud computing models to the banking transactions. Idea of homomorphic encryption is best suited for banking data and users that helps in data management and privacy to highest level. (Song, 2021)

To get many services online in today modern age, cloud computing is second to none choice not only in economically viable but also considered as a seductive trend in marketing these days. The only things that keep the companies reluctant by not adopting to cloud technology is the security of the sensitive information while implementing the cloud technologies in their organization. most organization are opting for homomorphic encryption technique by encrypting banking data of the user. The data would be encrypted with “logical and” used for keyword search into an encrypted database. (Wenjuan Li, 2020)The lack of cloud service usage in different sectors, corporate sectors and in banks compel researchers and analyzers to do deep research about the methods that would enable such organization to use cloud computing technology in order to carry out their daily tasks, operations and customer care services in order to keep themselves head-to-head with modern day innovations. encouraging them to migrate and to outsource their computations to the cloud, and keep up with their core competing organizations. But the issue is shifting to the cloud technology will lead to lose control over the customer bank data such as account number, deposits and loan etc. (Tebaa, 2015)

There are many reasons that demands that data should be migrated to the private or public cloud, or to the hybrid cloud due to agility, scalability and cost saving feature. The only drawback that cloud computing implementation in organization holds is security, which is a risk to critical business and financial information of the users while implementing cloud in the banking sectors which has to be tackled by the management through data encryption and data encryption key. As the name indicates hybrid cloud is the combination of both private and public cloud which could turn to any of the above as per demand. It eliminates the isolation so that It can be identified as a private or public cloud. As the demand of the organization increases hybrid cloud working as a private cloud could immediately transform to public cloud due to capacity increases. In hybrid cloud, we can structure our applications as per the obligation of issues contingent on their criticality of ethics. The issue of the banking in the cloud computing environment is that the data of the owner can be accessed by the government of USA without consent of user and cloud service providers. (Tebaa, 2015)

To eliminate this, issue some cloud service providers are installing physical key management server in the data center. Other security providers are also encouraged to use their key management services. Banking sector is considered as massive data generators, speed of processing data and

storing the confidential data. In recent times the hikes in the prices of computers have put banks in serious trouble to buy a greater number of computers, to quench their working capabilities. Banks are in serious thoughts of cutting their IT costs but not compromising their data security and integrity. So, to maintain the data integrity and security while keeping the IT cost low as possible traditional way of IT implementation isn't the solution nor it is desired. So, in this case cloud computing is the best suited technology for business data operations, and been implemented in several industries and in financial sectors. Hybrid model considered to be more secure and reliable with the blockchain applications of maintaining security of the user's data in banks while doing transactions, storage and processing of data without losing integrity, and security of data is what we will discuss in details in later sections of this report. (Wenjuan Li, 2020)

Literature review

Data privacy

Data privacy is the appropriate use of data to the organization for specific purpose. Data being collected by the organizations from customers should not be disclosed to anyone apart from customer itself and should meet the business requirements. Every country especially Australia and USA impose penalty on discloser of the information and data to the customer if not provided sufficiently. Mostly in banking sector and financial institution the data that has been collected is to ensure the identity of the customer through personal identifiable information.

Data Security

Data security is dealing in three domains including confidentiality, integrity and availability of data to the customers. It shows how easily data is available to the customers and accessed. Data when retrieved should be accurate and reliable. Data security is the planning of data and keeping the information safe and sound by eliminating the unnecessary data which are found redundant. Data privacy and data security both are required for preventing the data from being accessed or getting exposed to the third party without permission and also driving results from data that has been collected. (Amal Ghorbel, 23 January 2017)

Banking sectors feels reluctant in moving their customer's data into cloud due to security and privacy issues. Somehow researchers have come up with certain solutions that would be implemented in the cloud technology to improvise the data security and privacy in the hybrid cloud technology. Different surveys have been conducted to study and analyze data leakage prevention

(DLP) solutions to dedicate or avoid the leakage of the private information which the data is in use, in the transient phase or in stored form. Privacy issue of data is basically due to location and legal aspects. Customers of banks are avoiding the cloud computing in banking sector due to they would lose the control over the data. Losing control means giving up on transparency, data loss and leakage of the information. Another major that has been detected is the data flow among different cloud models such as dynamic nature of cloud such as duplication of data, data flow across the border and preservation. Compliance and user preferences in data management also but hinders in cloud computing implementation in the banking sectors due to emerging privacy issues in the organizations. (Amal Ghorbel, 23 January 2017)

Personally identifiable information (PII)

Information that can be identified as an individual with certainty. Personal identifiable information has two categories.

> *Key attributes* each of these identify as an individual identity such as name, number, national identity card number, email and password. When anonymization technique is used these attributes are eliminated.
>
> *Quasi-identifier* this sub category dealing with the code, data of birth and home address. These attributes are then used for correlating with the other datasets and then identifying individuals.

Sensitive information

There are different kind of sensitive information attributes which needed to be discussed below.

Demography shows the characteristics of the individual such as gender, nation, job position etc.

Finance represents the financial related data that are usually required for the account creation such as credit card numbers, account balance and financial transactions.

Hardware Identity represents the data subject's hardware identifiers such as computer IP addresses, Mac address and hostname etc. (Amal Ghorbel, 23 January 2017) (Tebaa, 2015)

While *memberships* represent the information of the personal related to religious affiliation, political or to the community.

Interest and habits represent the data that are related to the interest of the people such as web browsers history, history of data being used, and shopping activities.

Health related information deals with the medical record, images, diagnostics and diseases which are sensitive information.

Intellectual production represents the data that are related to the idea or innovation before validation.

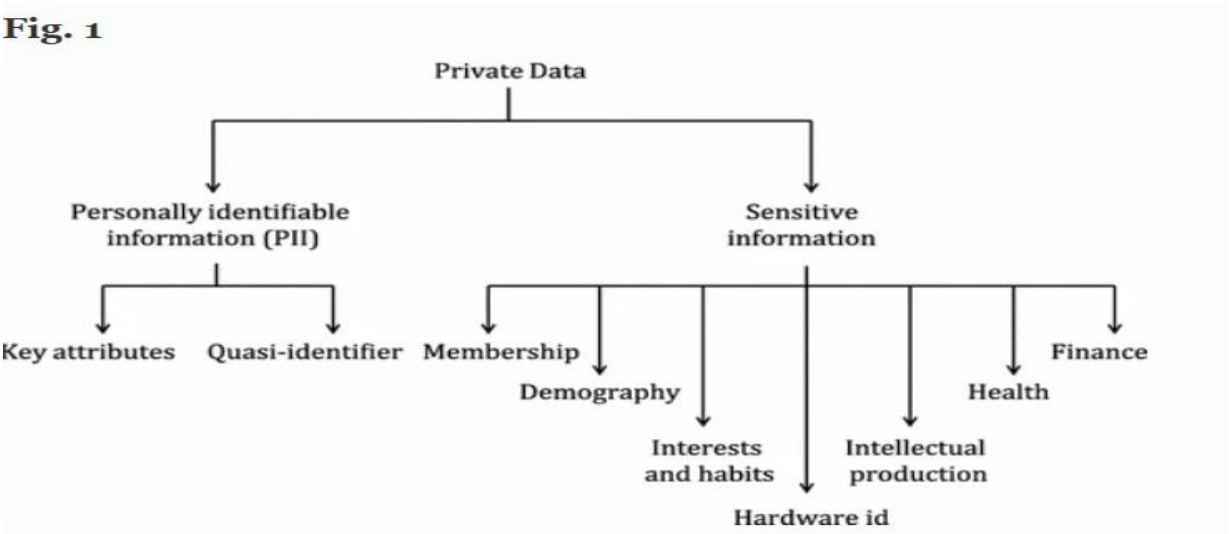


Figure 1. Types of private data

Cloud privacy issues

There are many advantages that are obtainable by the cloud while providing the solution to the businesses across the world like scalability, flexibility and agility, but as one has to know that everything comes with certain prices which is by compromising on privacy of the information that are being stored on the cloud. The users are led to move the data to cloud service providers. Outsourcing the data to external vendors or third party aggravates problems like who has given the access to the sensitive data. How many copies of the data would be stored at the cloud, storage location of data, all laws about security and privacy are followed and if violated who should take action, and if requested how to ensure that the required data would be deleted. Involved actors in the privacy case would be discussed in details below. (Tebaa, 2015)

Cloud service provider CSP

The cloud service providers are considered as the main entities in data hosting of the users. It comprises of internal employees and hardware devices, such as servers, web services interfaces, distributions, and load balancing algorithms. CSP can perform everything that a data and to play with it like replications of data, indexation, data mining and segmentation of the data. It also does storage of the data and duplication as well. CSP can intrude into the VM of the users while reading the data, and could disclosed it anyone, these actions grant threats to the privacy of the information. Intrusion to such information could be beneficial for some advertisement companies and could bring more revenue to share the confidential information about the users to the companies like choice of shopping etc. which is definitely against the well of the consumer. (Amal Ghorbel, 23 January 2017)

Cloud service broker CSB

This service assists in deploying applications in the cloud that could help to manage customer relationship management, management of cloud storage, management of documents and accounting. In some cases, this application needs to access to the private information of the users which later put the user in serious trouble of hiding his confidential information from the world. This application is highly doubted about sabotaging the private information, modify data and sending data to the third party. This notorious act could be considered as malicious threat to the confidential information.

Data owner

This entity is considerably the most powerful amongst all. It has all the authority to control and govern over the data it has. Data owner would decide if he is willing to give data to the cloud, private cloud, hybrid cloud or to the public cloud. Data owner would decide if he is willingly accepting the offer of using the service that are using cloud hosting. Sometimes the data owners are many in numbers instead of single guy, especially when the data is produced in group form or data is collected from different individuals. (Amal Ghorbel, 23 January 2017)

Cloud-based service administrator

Cloud services are provided by the organizations. The owner of the cloud services is called cloud service administrator. its key responsibility is to improve the services that they are offering. For that purpose, they always seek access to private database to gain information which often results

in disclosing of information of the customer private data. Sometimes the Cloud services hire administrator from outside organization which put serious threat to the privacy of the consumer. (Amal Ghorbel, 23 January 2017)

Cloud-based service third party

Third party actors are considered as trustworthy, but it can happen and in worst case it can be a serious threat to the private information and can threaten the privacy of the information. They can sell confidential information to the rival organization for good amount of money and that's where they are considered to be the worst threat to the privacy of the organization. (Amal Ghorbel, 23 January 2017)

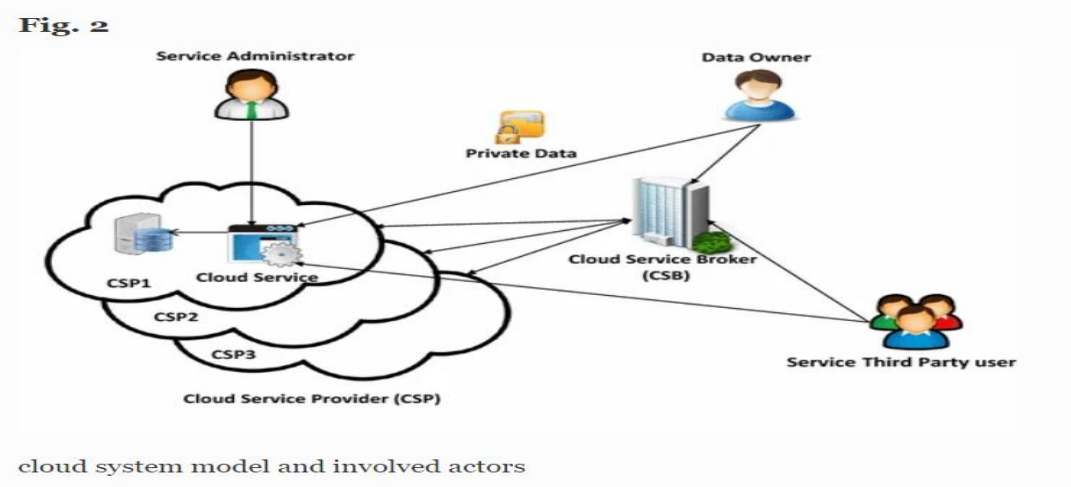


Figure 2. cloud system model and involved actors

Lack of user control in cloud-based services

In public cloud, remote mechanizes are used to store private data. Lack of control generally means lack of clearness of data processing, unknowing number of copies of data, and storage location information. These three has been serious issues since day one, and it becomes really hard to understand who is violating the privacy rules. For example, public cloud like Azure cloud has stored a data of user residing in Pakistan while his data is stored in USA, so the user is unaware of the privacy act of USA which might get exposed due to laws in USA. (Zhou & Zhang, 2013) Moreover, the restrained private data can be vulnerable to illegal practices. In this illegitimate data handling issues arises in which authorized actors performs unauthorized actions to the data that they keep. Illegitimate Data Dissemination (IDD) are applied sometimes in which plaintext data is

send to the unauthorized third parties. CSP most of the time generate income from this secondary data by exposing it to the third party which they sell on business organization for advertisement purpose. Another major concern is the leakage of the data that put hinders in adopting cloud services for many users and organizations.

Dynamic nature of the cloud environment

Dynamic nature of cloud could bring more risks for data privacy. Due to dynamic algorithms data can flow anywhere and has no specific place where it could store. So, user has no awareness about the location of data storage. The case become critical when the data flow in the plaintext form. Obviously, data could move to the different countries and every country has different security and privacy laws e.g. USA-Patriot Act put threats to the discloser of confidential information and data become susceptible to being disclosed. In some countries the user preferences and violation of legislation happen in storing data. some countries bother to share confidential information to other countries. Critical data sometimes need special protection and shouldn't be stored in other country. User usually feels insecure on storing data in the foreign country cloud service provider location. Some users impose the restriction that his data should be stored in the city in which he resides. That shows that how much information is critical to the user and sensitive. Data replications is another unresolved myth in cloud computing where data is copied to many servers to make it available but it gives the redundancy of the data which leads towards replications. Hence due to global view of the transaction of data across countries some data are restricted to be not processed due to rules and regulation. One cannot keep the track of the copies across the cloud servers. And it also dims the possibility of deletion of all data copies once requested by the client. (Zhou & Zhang, 2013)

Compliance with laws and user's preferences

Privacy compliance is another issue of the public cloud. The privacy issues that have been mentioned could be eradicated by the introducing policies for privacy and data security which would be done under compliance and enforcing them. The success of the compliance is totally depending upon the precision of the policies defined and the enforcement capabilities of those policies. However, data owners aren't aware of the practices of the compliances that data actors can performs to their data. For example, data owner who is shifting his data to the cloud has no knowledge of privacy policies and how they work, and they don't have right to design privacy

policy which could benefit them in future regarding privacy. Privacy policies are complex and hard to understand and also enforcement of several aspects of policy aren't implemented fully like dynamic data management, lack of user control etc. (Amal Ghorbel, 23 January 2017)

Accountability

Accountability has several definitions with respect to several subjects but according to data we can define the accountability as a "accountability is the responsibility of others information, data, and to protect them in term of privacy and security, taking responsibility of not misusing information, and held accountable if any misuse of the information happened". Auditing in cloud computing is closely related to the accountability and hence emphasizing on editing would improve the accountability of the cloud to ensure user that his data is safe, secure and privacy wont damage easily. Auditing deals with the data policy, who is collecting and processing data, information about the actions that would perform on data, and keep the data track alive. Mostly these days' data is duplicated and sent across the border which rises many issues in cloud computing. The auditing and monitoring of the data would be done in the intelligent way through data management of the cloud computing. (Amal Ghorbel, 23 January 2017)

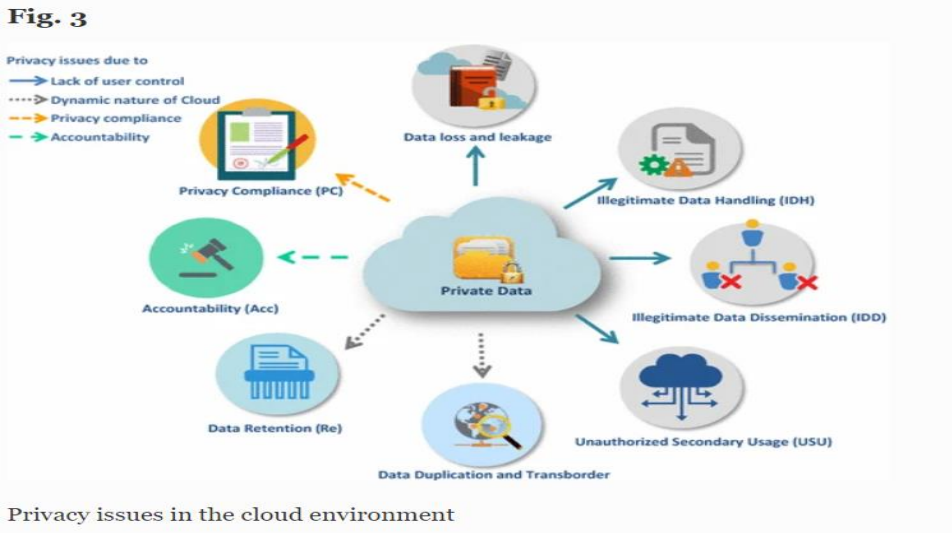


Figure 3. privacy in the cloud environment

Trust based blockchain approaches in cloud computing

Data in banking sector and transaction need a trust-based approach to maintain its privacy and security for credible interaction in the cloud computing environments. We will discuss different trust based blockchain system which will help users in the banking sectors to have trust in the

cloud services and would agree upon moving data to the cloud data centers. Three categories including blockchain enhanced trust interaction framework, basic blockchain trust framework and blockchain updated cloud data management would be discussed in details below. Basic trust based blockchain network has further two subcategories including identity authentication and access control and behavior management evaluation. While blockchain enhanced interaction framework consists of four subcategories 1) blockchain as a service, 2) blockchain cloud-based transaction, 3) blockchain resource allocation 4) and virtualization of cloud. While the last cloud data management has three sub research areas including data storage, data provenance, and data access model. We will discuss each and every one below in detail. (Jiyi Wu, 2021)

Identity authentication and access control

Identity management is the fundamental unit of the trust-based cloud service system. Customers and cloud service providers are considered as authenticated legitimate nodes in cloud services. Identity management system consists of third party which leads towards security threats and issues in privacy. Handling over the data to the third party rises the concern of security and also have chances of single point failure. it also has excessive authority of the certification center. To negotiate with such issue distributed system introduced identity federation which overcome security issue across multiple domains but it brings up complexity of the system design and operations. N. Alexopoulos et al first used the blockchain (open distributed ledger) system to develop an authentication for trust management systems. This model helps users to mitigate and tackle trust and privacy related issues and could give them response to them in an efficient way. By implementing trust-related information in an encrypted blockchain framework could prevent five major attacks in term of security breach. (Jiyi Wu, 2021)

K. Bendiab et al suggested a blockchain based identity management system. Service venders of cloud are enabled to manage their trust behavior and relationships with customers in the distributed models in a dynamic way through decentralization. The core idea that would be used in data banks to have inter-domain trust based blockchain system between the user and the cloud venders. This model can cover the basic of three characteristics of the user identity and privacy that are user credibility, authentication and satisfaction. This will surely boost the confidence of the user while choosing the idea of moving the data to the cloud. This research has shown the limitation of federation management system in trust management and also designed a cross-domain

authentication process by taking the dual role of CSP. To increase the trust among the users and cloud venders a novel security model has been developed called the blockchain Authentication and Trust Module to increase the credibility and validity of authentication in sensors network. (Jiyi Wu, 2021)

Furthermore, it has stated the introduction of the human-like knowledge-based trust system to enhance security in the distributed systems. The decentralized identity management system consists of two parts, including authentication and behavior management. The framework developed by researcher in which personal information of the user is linked with specific public address and reputation was represented by the tokens. The innovation has been introduced in the blockchain implementation task in which they introduced the inducement tasks the members can earn Rpcoin by revealing the wicked users, and fluctuation factor which can increase the activity of a system nodes and changes in their credibility. Through this way the possibility of increasing the security in the identity management system of the cloud services. (Jiyi Wu, 2021)

Behavior management and evaluation

To understand the credibility behavior of the entities behavior management system is best suited system for predicting and assessing. Behavior based model basically dealing with four types of servicing including identity management, authorization, authentication and charging. Identity management and authentication are handled by the key pairs of public and private while authorization is achieved by the smart contract, charging was realized through the payment gateways according to the services. User need fair in the transaction which would be provided by the smart contracts through service management and tenant management which ensure the fairness of the transactions to a certain degree. But this behavior system has the limitation that only licensed distributed ledger could participate which has legal credentials. (Jiyi Wu, 2021)This model focused on improving the credibility of the data by implementing processes like data reliability assessment, information source rating (1 or -1), miner selection (capability proof), blocks generation and verification, distributed consensus, and reputation calculation. (Jiyi Wu, 2021)

Blockchain-enhanced trust interaction framework and mechanisms

Blockchain-based cloud service framework (Blockchain-as-a-service)

To detect the violation in the service level agreements because it is not reliable and executed automatically, a role is added with the name of “witness” to ensure the credibility. The Nash

equilibrium theory was also introduced to assist cloud provider and users to negotiate and reduce the gas consumption. In this proposed model the nodes were gaining profit by monitoring the transactions in clouds. The nodes of the blockchain force the entities to complete the money obligation after all parties agree on transactions. In this model the cloud system would consist of two types of contracts one is smart contract and another one is SLA contract. (Jiyi Wu, 2021)In this blockchain proposed model the provider and user both first discuss the service fee, service duration, compensation and witnesses to be co-employed, etc., and then pool smart contract is executed through random selection of witnesses. This model is still considered as the theoretical model and its existence in the real world is still a difficult to prove.

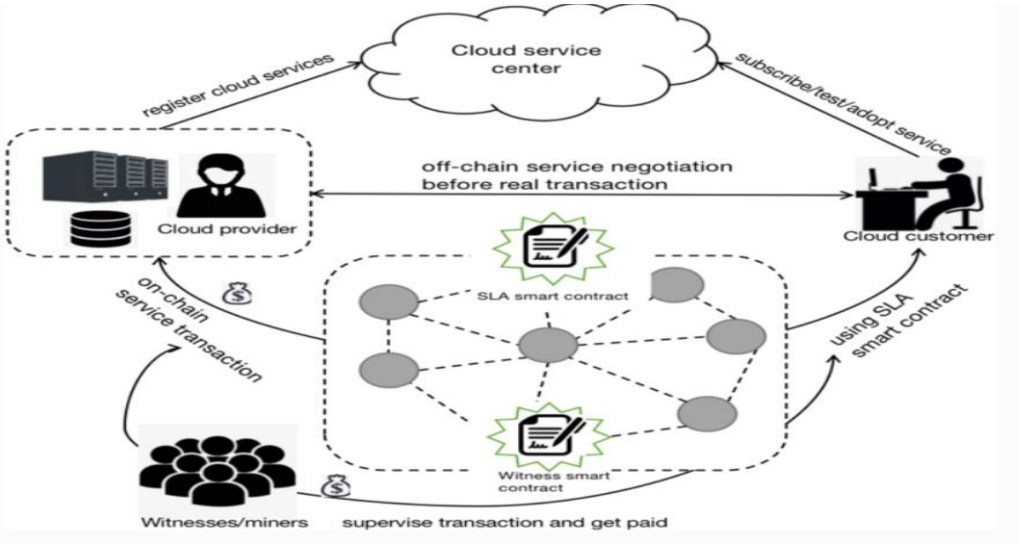


figure 4. Witness-contained cloud service interaction protocol

to eradicate the severe cloud security issues that are faced by the centralized cloud computing architecture they have solved them through introduction of hybrid cloud model with SDN. The new architecture includes the multi-controller layer SDN network layer with blockchain security management layer. This model is based on add-in blockchain security and independent management layer, designed to check the performance of the model through bandwidth occupancy, packet loss and resources availability. In this research the researcher has developed the app and framework of blockchain through the smart contracts, named blockchain-based DCMApp, and used as a resource agreement between the service providers and the customers. In this model most of the users' data is stored locally while some highly confidential data are stored remotely to reduce overhead. The introduction of blockchain in the cloud technology emphasizes on the

decentralization interaction without the need of third-party venders, but the drawback of this model is it can be still expose to the internet, it cannot precise the wrong operation and can do the payment for every operation online. (Jiyi Wu, 2021)

L. Xie et al proposed a new framework which is based on semi-decentralized trust-based model on blockchain technology in vehicular IOT environment in 5G network. This framework works on proof of stake and proof of work to discriminate between the right and malicious traffic while operating. (Jiyi Wu, 2021).

Technical analysis:

Blockchain-based cloud transactions

Cloud computing services relies heavy on privacy and trust-based transactions. Obviously untrusted cloud computing platform isn't reliable and won't ensure safe transactions among entities. To deploy software in the threat free environment a researcher named Zhou et al developed and proposed a clean room security protocol (CSSP) which is actually an agreement with the blockchain framework and usually working in the SaaS environment. The benefit of this protocols is to speed up the transaction process, increase productivity, and protect both service provider and user information. Improve the speed of implementation of software, monitoring of the software and also checking for the malicious content if found in the software the necessary is also taken against them. Smart contract used in this framework works very perfectly in term of detection of malicious content or activity that happen in between user and service provider process. Researchers have developed yet another model for credibility check of the service provider. Joint cloud usually checks the credibility of the cloud service providers by integration of the service requirement and credentials contributors. (Foster, 2002)

This model is proposed based on blockchain model for credibility evaluation of the providers in the Joint cloud environment, while ensuring the non-tampered of the data. This framework also develops reward and punishment mechanisms for both honest and dishonest entities. If the service provider and user both follow the rules and six attributes of the already described services rules they will be rewarded otherwise if found in fraud or malicious activity they will get punished. However, this framework failed to describe implementation of credit card framework based on blockchain, and also fails in dealing with malicious activity from user side because it merely focusing on cloud service providers. (Jiyi Wu, 2021)

Xie et al proposed another framework ETTF by utilizing a peer blockchain protocol, that are generally used in real-time transaction. This model works on three peers' framework, the global blockchain generating peer, the global validation peer and general peer along with two protocols. Protocols are PBP and E-commence Consensus Algorithm (ECA). Thus, ensuring the large secure and trusted transactions and achieve a much higher throughput compared to other framework while maintaining the security and privacy. This framework provides the instant transaction based on blockchain, while eliminating the latency and throughput. It also reduces the delay time. thus, ensuring the fast transaction using blockchain technology in cloud computing compared to other technologies this framework has been used in the ecommerce industry quite a lot. (Foster, 2002)

Cloud outsource is considered as the trending model in service provider these days. A novel framework is developed in other to maintain privacy and security in the cloud environment between user and the service provider. The framework named as Bpay. This framework provides the verification protocol along with the top-down inspection method. In BPAY the outsourcing payment method is divided into four phases, service execution, checking, payment and compensation. This trend of outsourcing payment method is new and thus provides the data integrity on both sides, on a user side as well as on cloud service provider side. To ensure the smooth, and fair transaction through this framework they use the dual verification process. Hence protecting the interest of both the service providers and client by introducing deposit transaction scheme. But BPAY is considered as complex blockchain framework and usually used in complex cloud services framework. (Gao H, 2021)

Blockchain-enhanced resource allocation

Blockchain has been considered as the best model for decentralized and distributed trust framework. But the survey has concluded that it utilizes and consumes so much energy and resources that it cannot be used in the hybrid cloud as an efficient system. Cloud mining technique is used to purchase services from the cloud services as considering new alternative to blockchain due to less resource allocation and considerable efficient. Blockchain applications performance optimization with cloud mining, researcher proposed game theory to cover the interaction between service providers and miners. For this work Alternating Direction Method of Multipliers (ADMM) algorithm has been used.

This model deals with multi-providers and multi-miners in order to resolve the resource competition and resource allocation to them. But this model has also a flaw because node scale is very small compared to the task it would execute so the practical implementation of this framework also looking dark. Transaction between the miners and the cloud service providers must be studied in detail with respect to allocation of resources, and wealth distribution among the entities. (Jiyi Wu, 2021)Two scenarios have been discussed one is to allocate resource to the miners and miners should compete for resources. Thus, in this case the miner will offload the huge tasks to the cloud edge services and will end up maximizing the benefit for both cloud providers and miner.

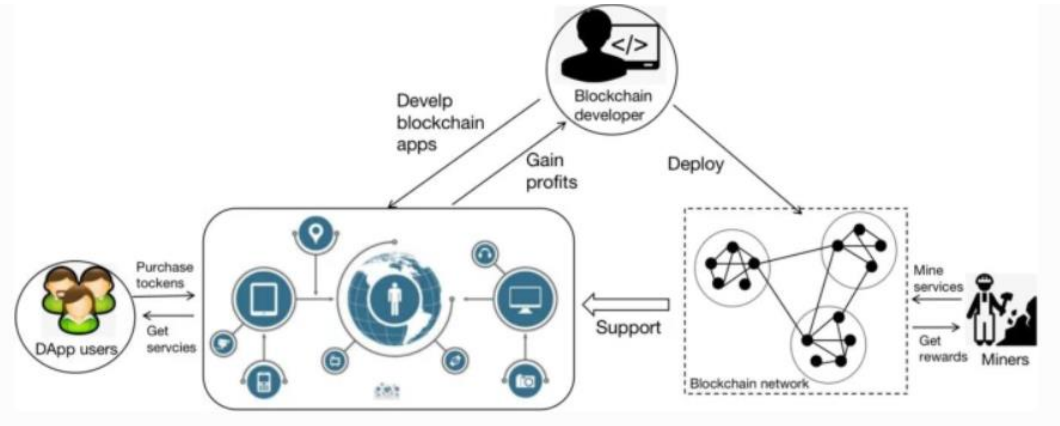


Figure. Blockchain-based cloud-edge business ecosystem

This model helps miners to work in resource constraints environment effectively while achieving their goals. The resources are purchased or borrowed in whatever sense one could take but they are doing it economically but they are losing their integrity and privacy. Framework of virtualization has been introduced here to allocate resources by implementing anonymous login and virtual resource management strategies. This research mainly contributed in resource allocation credibility and simplify the information gathering by distributed consensus mechanism. Resource scheduling problem is solved in this technique through traditional system factors performance, credibility of blockchain nodes and NFV- Mano system. (Foster, 2002)

Trust enhanced cloud virtualization

Docker is considered as best tool for virtualization, due to best resource utilization service of operating system without put extra burden on the system. Cloud users often get offended by the malicious content due to poor authentication knowledge and get victim of it. This tool also has weak mechanism compared to attackers in term of dealing with them. To improvise the quality of

the Docker virtualization a blockchain trust model named Decentralized Docker trust has been proposed. This proposed model provides the basic facility of verification and help in avoid Denial of service attacks on the network. Research has also contributed in determining the threats in the DDT framework and analyzed the weakness of the trust-based system. (Jiyi Wu, 2021)

Identity and access management

Identity and access management is the collection of tools, processes and techniques used to manage the identities of the individuals, their authorization, authentication and roles within or across the organization. identity and access management keep the roles and identities which are associated with the individuals, and also provides the decision makers with the means of control of resources. Key objectives of the IAM are to increase productivity, security, reduce downtime, cost and repetitive tasks. User is a natural player in the IAM system who holds the digital identity and wish to conduct a transaction. The identity provider is the entity that manage identity information. User get identity from the identity provider. While the third player is service provider who is providing the services to the users on the bases of information provided by the identity provider and user. (Soltani, 2021)

Identification is considered as the first step in recognizing the identity of the user while interacting with service provider. Identification process starts when user claim an identity. The subject can be identified through ID, or username. The verification of the identity process is also called identity proofing. The authentication is the process of verifying the identity of the entity by allowing them access to the system resources. This operation is performed by giving password, credential information or biometric identification to give access to the confidential information of the user. Multifactor authentication deals with four tiers of identification. (Soltani, 2021) First one is done through putting the password, or pin details to get authenticated. The second one is done by asking the mobile number or the secret that identity possess. While the third is the actually looking at the identity that what actually it is by using biometrics, or face recognition. Fourth and the last tier of authentication deals with the attributes such as time and location of the user. Another trending and modern approach is the continuous authentication in which authentication process is carrying out continuously without bothering the user to get disturbed from the process while interacting with the service providers. The continuous authentication usually guarantees the identity of the user that

has been provided at the start of the work. Authorization is the process of giving right to the entity to access the resources. (Soltani, 2021)

Self-Sovereign identity

Self-sovereign identity model is the upgraded version of centralized and federated identity model. In the SSI model the user keeps his privacy by having cryptographic keys using Hellman key-exchange protocol. This particular model gives liberty to the user to have full control on data and decide freely upon sharing the data with anyone. The identity holder is a role that have either one or more claims. User has the autonomy of controlling the identity and administration of the identity providers should consider him as the main player. To ensure the user interoperability SSI model must be transportable and should not be confined to a single location. SSI system can be defined by the ten rules below. (Soltani, 2021)

Existence: *"Users must have an independent existence."*

Control: *users must have the liberty to control their data and identity in their desired way. Everything will be discussed clearly there must be a difference between ownership and control of the data. user may have control over the claim about identity issued to them but this not necessary that they own the claim.*

Access: *users have legal right to have access to their identity. They must not be anyone in between to restrict the user from getting access to his own data. But it does not mean that user can alter the data or claims that are associate to the data but rather they would be known if there are some claims that are been altered by someone else.*

Transparency: *identity holders should be known to the all algorithms and must be clear about procedure and processes. Open-source framework is famous to give good insight look to the user about algorithms rather than algorithms which prefer vendor lock-in.*

Portability: *the transportability of the data is must in SSI model. Identity holder can transport it data from one location to another without being restricted. This gives longevity to the data and data should not be kept in the third-party territory even though if they have best intention for identity holder.*

Consent: *user must have consent of the data that are related to him. He has all rights to know about where his data would be used.*

Persistence: *identity data must be long-lived and should have a longer life span until and unless it is disseminated by the identity holder itself. There must be room for data modification, removal or adding new information related to identity.*

Interoperability: *identities should be used as widely as possible. The data should be used by different entities and there must not be any restriction. This supports the durability and availability of data to different departments such as jurisdictions, and architectures.*

Protection: *SSI model basically emphasizes on the protection of data of the identity holder. The data must be protected and user has every right to use it in the protected way.in case of any issue between identity holder and network the decision would go in favor of identity holder to ensure protection. SSI architecture is designed in the decentralized form to eradicate the monopolies of the service providers and avoid possible censorship.* (Soltani, 2021)

Minimization: *the user has all rights to disclose as minimum data as it is required for the task completion. The discloser of the data should be minimized through privacy techniques, selective discloser and proof-based mechanism.*

The major difference between the identity holder and the identity owner is identity holder which has possession and owns claims of the issues but may not be the owner of the claims, while the owner is the user, or organization of the identity. (Soltani, 2021)

Architecture of the Self-Sovereign identity model

Decentralized identifier:

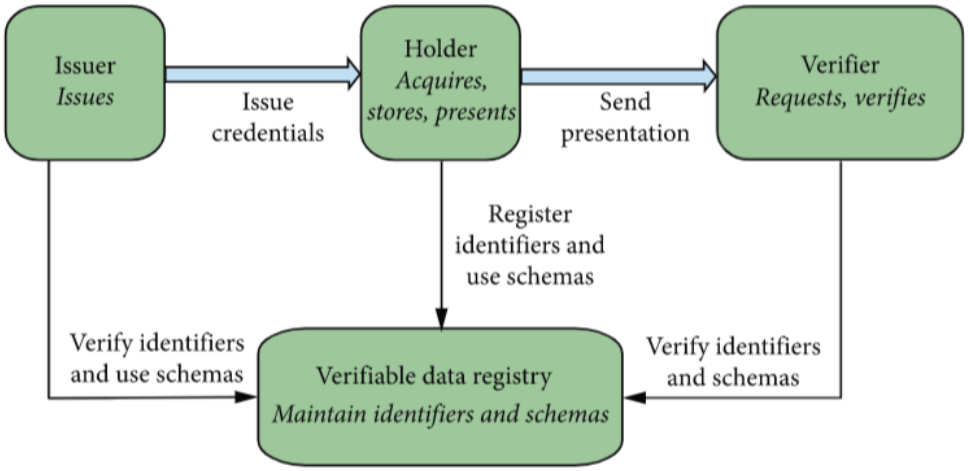
Decentralized identifier are the key components of the SSI model that are developed by the W3C. DID has unique numbers and cryptographic identifier scheme similar to universally unique identifier scheme. It does not rely on the centralized identifier and hence that's why it is called decentralized identifier. Its addresses are managed and generated by the cryptographic keys. The ownership of DID can be proved through digital signature thus making it unbreakable security-wise. DID is the combination of three parts, scheme, method and identifier. DDO (DID descriptor object) is machine readable file and contains various information about the DID subject. These include the information of the endpoints, cryptographic keys, authentication parameters and

timestamp. Service endpoint established connection with the DID subjects. DID subject is identified by the DID which is defined by the DDO. Both DID and DDO would have a key-pair which will be stored in the record. DID could use the specific method to be implemented on the specific network through scheme. This includes four steps, creating, reading, updating and deleting pf the DID using CRUD. There are about 90 methods of the DID. (Alex Preukschat, N.D)

DID has three types one is public DID which is used for unknown-able number of parties. The pairwise DID is used for the interaction in which DID is known to at least one entity either user or service provider. While the third type is N-wise DID which are used by the N-number of entities including its subjects. DID auth protocol is used by the identity owner for its client application in which client uses its mobile device to prove to a service provider that it is in control of a DID. This protocol replaces the user's name and password and establish a secure communication between the user and the service provider. The DID comm protocol is used to communicate securely and privately in peer-to-peer network. This protocol supports manual authentication procedure. (Soltani, 2021)

Verifiable Credentials:

A verifiable credential is an interoperable data structure suitable for representing cryptographically verifiable and tamper-proof claims. Holder is an entity that controls one or more variable credentials. Issuer creates new variable credentials. Verifier is used to verify the credentials that are being sent to them. Online website or ecommerce store that are asking credentials from customer is the example of the verifier. Verifiable data registry is the unit which assists in creating and verification of identifier, keys, verifiable credentials schemes and other information to create variable credentials. Variable credentials could be the URL of the issuer, URL of the subject, and URL that are used for the identification of the credentials. This also required the digital signature to verify credentials. (Soltani, 2021)



Verifiable credentials actors.

Decentralized Public Key Infrastructures

Public key cryptography is the basis for cryptographic operations and it consists of a set of services, tools, processes and technologies which assist the process of cryptographic operations. PKI X.509 or PKIX is the most widely used PKI certificate model in which the digital certificate X.509 is created by central certificate authorities. Public key is given a particular identity by certificate through binding. To assist the management of addresses and identifiers, the DNS registrars and ICANN were created along with certificate authorities. (Soltani, 2021)

Pretty Good Privacy protocol is a decentralized trust prototype which depends on web of trust unlike traditional PKI models and it was implemented prior to the emergence of blockchain. The growth and popularity of PGP model was hindered by the security fears linked with the practice of long-term keys and some unrequited usability and key administrative challenges.

The decentralized public key infrastructure, suggested by the Rebooting of the Web of Trust, provides a distributed trust architecture whose safety and reliability of the organization as a whole can’t be compromised by a single entity. Unlike traditional PKI, the DPKI doesn’t depend upon registration authorities or central certificate authorities.

With the help of blockchain, a distributed key-value data store, the DPKI architecture can be realized. The block-chain which stores the public keys immutably, the data once it is written cannot

be modified or deleted. Any changes to the data are auditable by all members of the network because of the public and transparent nature of the blockchain. There is a difference between trust in a blockchain system as a clear, undeniable and protected storage platform and the confidence placed in the data like public keys stored on a blockchain and it should be distinguished. To let a certain unit to check itself and call the ownership of a public key and a particular identifier, a verifier can rely on DPKI, but that doesn't mean that a verifier is trusting the entity's identity. (Soltani, 2021)

It should be kept in mind that not every security breach and concern about traditional PKI is because of its essential design but reasonably it's due to the insufficient application of security controls in PKI placement or the users' deficiency of appropriate key management. These types of issues are not just confined to PKI but can also be found in DPKI models. It is true that certain features of DPKI may possibly be dependent on blockchain but several tasks such as cryptographic signing operations require secret keys which must be carried on in very safe off-the-blockchain setting. The DPKI systems which offers crystal clear and open structural design, the CAs in the PKI prototype have also plunged towards superior clearness by offering services like certificate transparency logs through which, on CA website, all issued certificates are published. Nowadays, in many industries including health, defense and in banking, many PKI enterprise systems have been installed and its management and deployment have become easier due to rapid growth of PKI ecosystem. (Soltani, 2021)

Blockchain and Distributed Ledger Technology

Both the blockchain technology and the distributed ledger technology, provide a cryptographically dispersed, and secure databank of information. The blockchain support a bulk of cryptocurrencies including Bitcoin. The blockchain reduces the use of CAs and the transaction through it takes place in P-to-P method without the reliance on a central trusted party.

The applications of blockchain nowadays aren't just limited to financial use cases. It has rapidly developed into other industries such as smart homes, supply chain management, healthcare, energy, legal, voting, storage, identity management and corporate registry.

The blockchain technology was first introduced with the introduction of Bitcoin white paper in October 2008 by Satoshi Nakamoto and its procedure was officially launched on the third of January 2009. The block-chain system is composed of an absolute ledger of cryptographically diced transactions. A block is formed by a group of transactions collected together. Each block has a header that references the cryptographic hash of the previous block in block-chain. (Alex Preukschat, N.D)

Blockchain systems can be distributed into two classes of public and private which are further separated into two types of permissioned and permission less. As the categories names indicates, a public blockchain is wholly accessible to all public members to read while on other hand, a private blockchain is reachable only to the authorized members. A permissioned blockchain can be read by all parties but here the writing authorization is only given to some units while permission less blockchain permits anyone to write. Examples of public permission less blockchain are Bitcoin and Ethereum.

One of the most integral components of blockchain is consensus protocol which guarantees that only those transactions are added to the blockchain which are valid. Bitcoin relies on proof-of-work consensus protocol. In this procedure, the members, of the network, are given a difficult mathematical riddle to solve in order to verify the legitimacy of the transactions they aim to submit to the network. (Alex Preukschat, N.D)

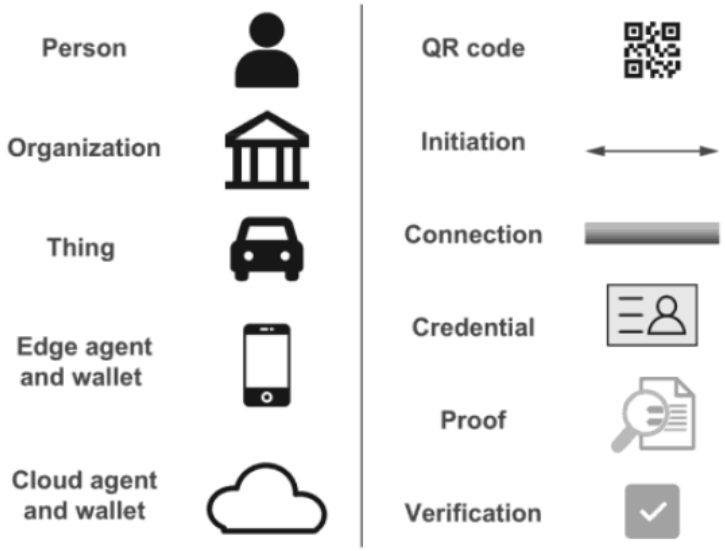
The benefits of blockchain to attain IAM requirements are; first, as the nature of blockchain is decentralized, it lets data and spread-out identifiers to be stored in a ledger and not with a single central unit. Second, as it is indisputable, the transactions can only be attached to the ledger and the integrity of it can be confirmed by all members of the network. Third, public blockchain makes sure that all the records are always visible to all. Fourth, since its nature is distributed, it is very hard to affect the integrity and availability of data.

Many kinds of operations like notarization of identity data, registration of decentralized identifiers, and storage of admittance and agreement records can be shared on blockchain. In contrast to the public keys infrastructure, where the public keys are exchanged directly through secure channel, with strong reliance on certificate authorities, the blockchain, having decentralized nature, allows for unit identifiers and cryptographic keys to be registered on irreversible and distributed system reachable by any entity that want to obtain the authentic public key of the identifier. Due to its

distributed nature the data integrity and security should be preserved. Blockchain having smart contract would enable them to develop a sophisticated and decentralized identity management system. Thus, this management system would be used in the cloud to store data, transaction between the client and the service provider an order to keep the data integrity, privacy and security. (Soltani, 2021).

Digital agents

There is a difference between the physical and digital wallet, the physical wallet is directly operated by the owner while the set up the credentials, issuers send credentials to them, receive credentials by the verifier and verify it. Thus, wallet can be moved between pocket to pocket and purse to purse. While the digital wallet is supported and operated by the software. The software module is called digital agent. The agent basic function is to ensure that only you can use the wallet, protect it from cyber criminals and allow use while putting the cryptographic keys to them. In SSI infrastructure the digital agent has second job which is to set up secure communication between the users, talking to each other and sharing credentials with each to provide the environment for private communication. This is done through decentralized and private messaging protocol that works with the clients. Clients communicates with each other through two different kinds of agent. One is the local agent or edge agent which communicate with the local network or identity holder while cloud agent operates in the cloud either work in the private cloud or in the standard cloud. So, the cloud agent can store the record, financial record, file, data and photo of the identity owner. The data here would be stored in the encrypted form and then will be used to communicate with the cloud service providers. Secure data storage works as the backbone of the digital industry while working with digital data management to maintain and keep the data of all kind of identity owner.



Sample set of the items that are used in the SSI communication

Analysis of approaches

SIMS is the short form of self-sovereign identity management system which help is a blockchain-based privacy-preserving identity management system based on a ZK-SNARK (Zero-Knowledge Succinct Non-Interactive Argument of Knowledge). SIMS can be implemented in any environment in the bitcoin, Ethereum, Hyperledger, using ZN-SNARK approach. The data can be transmitted and encrypted through smart contract between the identity owner which is user and the cloud service providers. When this protocol is used it assists in preserving privacy of the information under self-sovereign identity management system. More and more application can be built on this approach, transaction and work history can be encrypted and transported through implementation of this protocol. Data protection in the cloud environment is generally looking very hard and complex due to intervention of different actors such as data owner, data consumer, Cloud service providers etc. (Bandara, 2021)To use the hybrid cloud model it would involve all stakeholder into one platform and involve them to make policies for data protection. Data owner would have every right to protect his data either through encryption, sticky policy or whatever it takes him to provide protection and privacy to his data. so, SSI which has basic feature of the data protection, identity management and data ownership when combined with Blockchain using ZN-SNARK protocol would enable the data owner to successfully protect the data and transmit

information between the client and the service providers. The CSP must enforce the policy of the data protection across the cloud network to play the role on the other end. (Bandara, 2021)

Hyperledger indy is a public distributed ledger made for the decentralized identity. Indy permits the user to manage their own SSI and allow them to read the content of the ledger. The ledger is using protocol named Plenum Byzantine Fault-Tolerant to do transaction in the ledger. Pairwise protection is given by the asymmetric key and Pseudonymous identifier to ensure the privacy of the data, avoid identity correlation, and connection is being made top secured between the client and the cloud. Indy uses DIDs as the primary keys enabling long-term digital identities without centralized registry services. This approach could be used by the user to get the pairwise protection through asymmetric keys in the hybrid cloud computing services. (Bandara, 2021)

Sovrin is the open-source decentralized identity model on permissioned distributed ledger technology, basically this is used by the trusted institutes such as banks and universities. The basic contribution of this model to the privacy is that user is allowed to have as many identities for privacy purpose. Different identifier is managed by the different asymmetric key pairs and are often not linked with each other. Sovrin identifier itself is managed by the identifier itself using the decentralized identifier. DID has the data structure containing the user identifier, cryptographic public key and other meta data to transact with the identifier.

For self-sovereign identity systems Portable Trust is biometric based authentication and blockchain storage. It has biometric-based authentication model that is autonomous, open-source and permission less. This makes the user to store their personal information in a very safe and secure manner which can only be accessed after a biometric verification by the same user. To be used in mobile phones it combines a permission less blockchain with identity and key attestation. This work is mainly focused and based on self-sovereignty. (Bandara, 2021)

A method for securing biometric information registration and access is Horcrux protocol. The generalization of this protocol is because of the fact that two or more biometric shares can be stored across mobile devices and some personal storage providers that are redundant but safe. The owners of biometric data can give permission to others to get access to the data without having need of involving third party. By creating multiple departmental nodes like organizations and governments, blockchain helps enable an environment which results in mutually formed

distributed consensus making sure that the record of the data is in difference places which makes it resistant to attacks.

Allowing identity owners to control their personal identity and corresponding keys and data is made possible by uPort. The owners of identity and data can authorize others to have access to data, signing of documents, controlling and sending of values on blockchain, interacting with application and smart contracts as well as encrypted data. All those enterprises which uses uPort can create identities for new employees and customers, can establish a Know-Your-Customer process, can build environments which are secure access controlled, can reduce liability by holding little sensitive information and can establish such type of environment where identities have specific roles. (Bandara, 2021)

A protocol which is of open source and general in nature for facilitation of identity records for personal and non-personal entities is termed as Jolocom. This protocol is designed to operate on Ethereum which is a public blockchain infrastructure. All personal information is stored off-chain and is in full control of the user while the blockchain has all the digital identifier information. This protocol is currently in use of Ethereum test network as a prototype.

The utilization of blockchain technology by a mobile-app based identity system to create a safe and secure protocol for storing encrypted personal information along with sharing verifiable claims about personal information is done through Sora Identity. In this system, encrypted with a cryptographic key which is owned by the user, hash values of a user's personal information are published in the blockchain, and this system is built on top of Hyperledger Iroha permissioned blockchain. Users can share their personal information to various organizations and institutions such as banks, according to their own will. (Bandara, 2021)

Blockchain as a Service

Baas borrows concept from cloud computing. This model ensures using blockchain as a computing service or blockchain components used for cloud services for applications. The key aspect of blockchain used as a service is to emphasize customers to focus on their business by avoiding technical hinders that would come in their way while handling tasks. (Keke Gai, 2020)Cloud has already established three models such as SaaS, IaaS and PaaS but to cope with increasing demand of customers using cloud they had to focus on using blockchain as a service, security as a service and process as a service to feed the demanding customers. to enhance the interoperability of

services blockchain and smart contracts are being used in the cloud computing. Cloud computing has diverse nature by providing the services to the organizations. Blockchain chain will be used in the backend services for BaaS. BaaS enable the customers to get the cloud-based services and get blockchain support. Alibaba using the blockchain services in the transaction tracking system, consortium government and smart contract to the customers in term of services. BaaS provide the time saving, low operational costs, economical and control optimization and integration. Blockchain being implemented in the cloud computing to provide good customer services for transaction, security and integrity of data in the online world. Infrastructure of the Blockchain as a service is provided by the service providing organizations and codes are available for the open source. Some of the parts of the BaaS are under observation and need good research to carried out on that to enhance few technically challenging difficulties specially communication, data synchronization and consensus. two characteristics has been displayed from the research stated below.

- Blockchain implementation in the cloud environment is itself a challenging task which are well explained by BaaS. ODP manager provides scalable and adaptable service scope ranging from establishment to configuration and maintenance.
- Cloud service providers manage to run the operations and infrastructure of the blockchain computing source by delivering customer agile service in terms of hosting blockchain application and blockchain functions. (Keke Gai, 2020)

When there are multiple cloud services providers involved in the integrated cloud management model there are some issues with data transactions and usually have lack of data control from either party that could easily be managed and eliminated by the implementation of blockchain data-tracking system. Blockchain platform should be maintained and established by the cloud service providers. They will look into operations such as performance enhancement by designing perfect API, and risk mitigation by providing security.

BaaS industrial implementation

BaaS emphasizes on the interconnection between the logical and physical business activities. Due to revolution of blockchain and emerging trends in blockchain, all major cloud service providers eyeing on the performance of blockchain in the cloud environment. IT companies such as IBM, Microsoft, and Amazon all are deploying blockchain features in their respective cloud service

models. Oracle BaaS are targeting the payment and logistic department, while IBM BaaS attempting to provide services for vehicular systems. (Keke Gai, 2020)

Microsoft Azure a cloud platform that provides the deployment of blockchain which supports Ethereum, and Hyperledger fabric for the deployment and configuration of the blockchain network. the user of Azure could only configure few portions instead of going into full technical details. The data backup for on and off-chain services are provided by the Microsoft. Azure currently supports only the single point configuration of the blockchain and consortium blockchain deployment is still under observation. (Keke Gai, 2020)

IBM cloud platform provide the public cloud which are used by the users to deploy blockchain on it. IBM only accept the solution provided by the Hyperledger fabric while rejecting the widely used Ethereum which is considered as the biggest disadvantage of it. But the advantage of being implemented BaaS in IBM is it provide life-cycle management for users that could benefit them which could help which can ensure a reliable outsourced data management for users. IBM BaaS relies on the secure container. IBM Baas supports on-promise configuration of the blockchain for their user. Another distinctive feature that IBM BaaS has is having secure and reliable cloud environment.

BaaS under Exploration

Baas being explored in the cloud environment and comparing it with the fog environment. Samaniego et al a researcher has discussed that while running on both platforms the blockchain as a service could have higher computation and storage capacity in cloud than fog environment while having latency time was longer. But he also concluded that the cost of communication between the IoT devices and servers were low in the Fog environment compared to the cloud. Another aspect of the blockchain integration in the cloud computing has remove the fear of trustless third-party due to decentralization. Also, it has improvised the trustworthiness of the stakeholder and the interaction between stakeholders can be done irrespective of whether they are trustful or not. (Keke Gai, 2020)

Service agreement will be signed with the service provider to restrict activities of the CSP. Trust concern and integrity of data has always been headache for users and they were least relaying on the CSP to provide the integrity and security for their data. CSP can provide data integrity by offering transparent operations on the distributed ledger. But here four optional solutions are

presented by the Singh at el. Are authenticated trustful environment, access control strengthening, eliminating recentralization by introducing CSP federation and fourth is enhancing user controllability over the data. the reliability and availability can be improvised by the introduction of dynamical Reliability Block Diagrams to make master and slave in the Hyperledger. This integration resulted in high reliability and availability of the BaaS in the cloud environment. Third party identity management system has been eliminated by introduction of Blockchain as a service ID management system. This ID could be made by providing virtual ID and signature ID information. The verification could be done between the user the verification identities through BaaS ledger. (Keke Gai, 2020)

Another model of BaaS has been explored and discussed by Chen et al. that was expended to a server less architecture. This model is inspired from the big data open architecture and comprises of four layer which are component, service, infrastructure and logic layer. Zhang et al. proposed the smart contract based secure bailing approach in the blockchain as a service model in ride hailing services. The rule of the smart contract can be used and deployed in other system too that could enable other services. MaaS (mobile as a service) can be achieved through implementation of the Smart contract rules.

Consideration for Blockchain Security: Challenges

Settlement of Blockchain

Blockchain should be one because it is the sequential connection of the generated blocks, a blockchain can be generated temporarily if the two peer succeed in mining the answer for generating the block at the same time. So, the block that is not chosen by the majority peers in the bitcoin so the continuing the mining will be meaningless. Bitcoin will choose peers who have mining capability of more than 50% automatically. So, the security problem may occur here because if “attacker” god forbid get above 50% mining capability he can take control of the blockchain and can slip in falsified transaction by removing the original one. (Park, 2017)There is also some malicious way through which attackers can get gain to the blockchain by only having 25% of operating capability while pretending that they have 51% operating capability. Thus, this false state put the blockchain into serious danger of getting wrong statistics from peers and this could go in hands of cyber criminals which can exploit the whole system. Nonetheless, the mining pools are actively participating in mining thus increasing the probability of mining. So, anyone

who exceed the threshold of 50% would get full control over the blockchain and can utilize the data control of the blockchain making it unreliable and vulnerable to cyber security threats. Due to these risks anyone can dominate the blockchain by control the security of the bitcoin thus hindering its progress in the economic factors due which the price is very much volatile. (Park, 2017)

Security of Transaction

Different transaction form can be created from the different codes generated in programming languages. A bitcoin contract is a method of testing bitcoin against the existing authentication and financial service. Using the scrip this contract is generated by multiple signature technique called Multisig. But as the complexity of the scripts increases the possibility of the improper transaction has also been increased with it. Thus, this increased the issue of having wrong scripts lead towards the security and integrity of the data and transaction and it gives a look of unprotected system. Different model of bitcoin is suggested to the bitcoin contract type transaction to verify the script accuracy. (Park, 2017)

Security of Wallet

A bitcoin address is the combination of public key and personal keys in the form of hash value. The locked value of the hash should be unlocked with the public key of the address along with personal key. The wallet of bitcoin stores the information of personal key of the address which help in generation of the unlocking script. So, the information that are stored in the wallet are extremely confidential so if we loss that information it will leads towards the loss of bitcoin since it is useful for using bitcoin. So, securing the wallet of bitcoin is an essential aspect of securing the blockchain from hacking attacks.

Wallet security should be utmost priority in order to secure bitcoin through multiple-signature technique called multisig. The transaction only happens if there is more than one signature and it is considered as redundant security checking for the wallet. If the multisig is set in an online bitcoin wallet, the configuration would require the signature of the online wallet site and owner's signature so if the transaction is requested from wallet, it would require signature of both, so the malicious transaction will be prevented in that case due to since the owner's personal key is not stored, even when the online wallet site is taken over by a hacking attack. Multisig also evolving into the

transaction service by withdrawing money from the bitcoin wallet through two-factor authentication or use biometric data for customers. (Park, 2017)

Some cold-storage wallet, offline wallet is available to dodge the cyber-attacks on paper-based bitcoin that are not connected to internet may be called a physical bitcoin. To reduce the risk of online transactions scams hardware type wallet has been introduced in the bitcoin. The hardware type wallet named as trezor stores the key information in a temper-proof storage system that are connected to the computer via USB. This hardware type of wallet is only used when the signed transaction is transferred through the key by the authenticated user. The rest of the time storage remain like in a cold state while it is only connected when required a transaction. This is kind of more secure but has issues of non-user friendliness and loss of storage. (Park, 2017)

Blockchain security case studies

Authentication

In order to hack the blockchain the attacker always try to get access to the person key of the owner that are usually stored in his mobile phone or laptop device. Attacker sometimes uses the ill-meanings of installing malware on the mobile device of the user to intrude to his system and leak personal key information to hack bitcoin. To protect the personal key information some researcher has suggested the token system for the approval of transaction. (Gaetani1, 2017) Authentication measures should be strengthening to enhance the security of the storage that contain bitcoin. A two-factor authentication is considerably a good method for strengthening authentication.

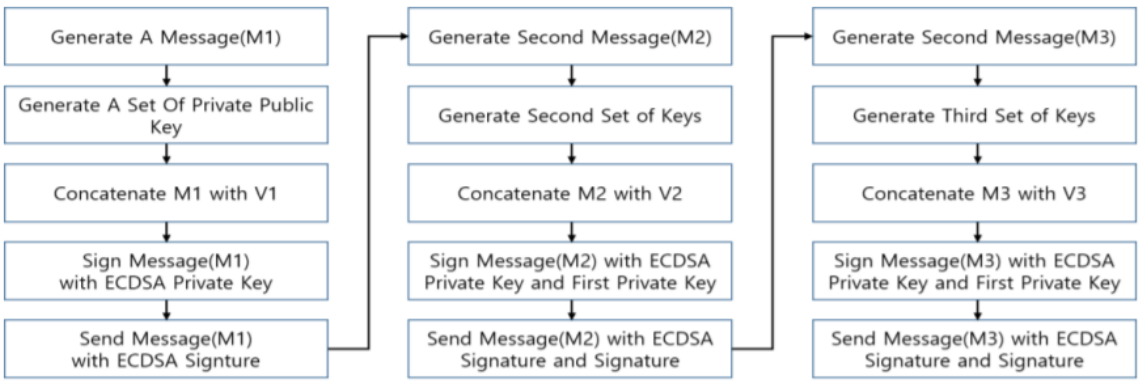
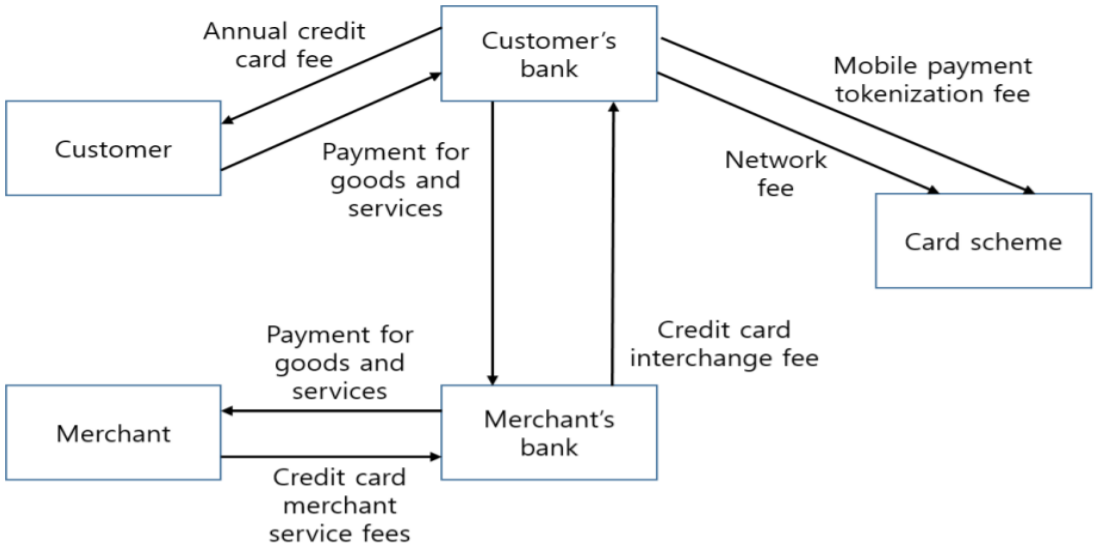


Figure 3. ECDSA two-party signature.

Improved blockchain

Blockchain holds complex system and customers are using it from different locations so the points targeting for security attacks are as many as the customers. the blockchain and the bank account are connected to each other and interact with each other to set a fee for shopping and buying goods.



General payment process system

Conventional transaction is done as a peer-to-peer transaction that would help in reliable and verifiable transaction thus costing much lower than the traditional transaction due to no involvement of third party. The transaction would be independent of the boarder and would be take less than other transactions that will surely get effected due to border involvement. The conventional transactions are very much hackable due to its data is stored on the central server while that of the blockchain gets away with it due to now central server involvement. To hack the blockchain transaction the attacker must hack the distribution system and alter 51% of the peer-to-peer. (Park, 2017)

Challenges of Self-Sovereign Identity

IAM models are continuously evaluated and the self-sovereign identity model is the latest evaluation. There are various shortcomings in already existing digital space and the self-sovereign identity model attempts to address it but the problem is that it also faces some challenges which needs research, discussion and exploration. Following are the challenges:

Standards for Data Management and Wallets: For better working of self-sovereign identity model, the user experience should be taken into consideration and by using user experience, standard protocols, practices and policies along with data management, and data exchange should be clearly defined and implemented. The SSI model which supports an open ecosystem, there's value given to user interaction, data management policies and standards. Regardless of the direction taken to tackle this challenge there should be an approach towards using a system which is user-centric and focuses mainly on the problems faced by the user and works on it to nullify it and this user-centric system should be aligned with Privacy by Design and Security by Design principles.

Key Management: As we know that in traditional identity models, the providers of identity are mainly responsible for the management of secret keys and identity data and therefore it must address these weaknesses, technical requirements, risks and dangers associated with the task. Same is the case with SSI model, these risks which are associated with the responsibility of the task are placed on user's shoulder. There are numerous examples of users losing their cryptographic keys, making them lose their valuable information and unrecoverable funds. This challenge of key management is addressed in way that the reliance should be shifted towards decentralized key custodians. (Soltani, 2021)

Consent: The consent given by the user for the data sharing and identity data should be meaningful, clear, unambiguous, given freely and specifying clear decisions. As the user is requested many times to give consent to privacy policies and data sharing practices, where the user is bombarded with privacy notifications, has led to the consent fatigue. In consent fatigue the user gets tired of giving consent again and again and thus it affects the user's decision making and its clarity. To address this challenge automated decision making should be entertained and research around consent management is valuable. (Soltani, 2021)

Access: Distributed ledger technology is the backbone of many SSI systems. Some DLT systems are of public nature which allows any entity to read and write into the ledger while some are permissioned which allows only selected authorized entities to read and write into the ledger. Both the permissioned and public systems should be carefully designed so that the attacks on it are defended and the risks are minimized. (Soltani, 2021)

Accountability and Governance: For the identification and addressing of malicious behavior and dishonest entities, it is of utmost importance to make policies and procedures and it is also important to make the decentralization possible for the better working of SSI systems. One of the drawbacks of SSI is that some of its implementations makes the selected few entities very powerful and controlling, making these entities the weakest point of the network while some other implementations are of more decentralized, machine readable and programmable governance framework.

Trust in Data: Such methods should be carried out which builds the trust among entities and the trust in data and these should be carefully designed. For the validation and authentication of data a trusted authority outside of the blockchain network should be used.

New Technology Adaptation: As SSI is a new identity model so it requires some modifications to the existing system architectures. To make SSI successful it is necessary that such discussions which involves suitable technology stacks, operational procedures and deployment practices, should be carried out. It is of utmost importance that user experience including the user interaction from the operator's perspective should be given maximum attention. Such type of steps should be avoided which led to the downfall of Pretty Good Privacy model. (Soltani, 2021)

Investments and Commercialization: As it is a new model with a growing ecosystem, it is facing some problems because of the limited knowledge on the revenue model, unknown user acceptance and unknown risks so any entity wanting to adopt SSI must design a plan that supports the investments and tackles the risks involved in operation of such system. (Soltani, 2021).

Challenges of blockchain in cloud computing

Scalability

The scalability is huge issue in progressing of blockchain technology due to limited storage available on the nodes. As we know every block has a specific storage and it also needs to store every transaction, so it become very hard to process and store many transactions at a time. blockchain capacity restricts blockchain to seven transactions per second and it also takes some time to create new block. Thus, with huge number of transactions, transactions would get delayed eventually. Scalability issue is the highest in blockchain.

Regulations and Laws

Legal issues have been risen in the deployment of blockchain due to lagging legal supervision. There should be clearly stated laws that would be implemented and followed while implementing the blockchain characteristics.

Governance

Blockchain has vast implementation in terms of governance and implementation and can be expected to transform government rules and laws. Helps in improving the security of the government structure, making it less complex and provide transparency to it. Blockchain deals in distributed system that needs a specific governance policy with immediate solution.

BLOCKCHAIN: HARNESSING CLOUD DATA SECURITY

Blockchain technology is a cutting-edge technology that assures data security and prevents unauthorized access to data. Each block of the blockchain contains three parts: one that stores data, another that hashes the data, and a third that serves as a reference. Despite the latest development, adversaries are still capable of determining the key from the hash code. Several techniques have been used to address the challenge of data security in the cloud. The BTM technique, on the other hand, assesses users' trustworthiness based on a data structure that was deliberately provided to them. Despite this, the malicious user was able to recognize the structure and read the data from the blockchain. Similarly, several ways to make advantage of blockchain and encrypt data with different keys based on the user's profile. These aren't dynamic methods. Similarly, the technique does not use distinct key/scheme selection strategies depending on the level of data.

All of these factors influenced the researcher's decision to develop the proposed RSFSA model with blockchain encryption for cloud data security. Moreover, to address the cloud provider supply chain difficulties, a novel system is also proposed, called Multi-Cloud Hybrid Cloud Supply Chain Blockchain with Smart Contracts (C3HSB). Both the models will be covered in subsequent paras.

RSFSA based blockchain encryption

The method initially undertakes RSFSA analysis with the help of different access traces created using these two. Furthermore, the technique predicts that FLAG will limit user access. Similarly, the technique generates a service result and a blockchain with the appropriate number of blocks based on the quantity of the data, with the resultant chain being adjusted based on the user's grant as specified in the user profile. To provide a result to the user, each block of data has been

encrypted using dynamic methods. Figure 1 depicts the functional architecture of the proposed RSFSA model.

The real-time service-centric feature sensitivity analysis algorithm looks at how sensitive particular features are. The rate of adversary attacks encountered by any feature in the cloud data was used to define its sensitivity, as well as the attribute taxonomy. The approach divides the characteristics into groups based on their sensitivity. The produced sensitivity class has been employed in the development of blockchain and data encryption.

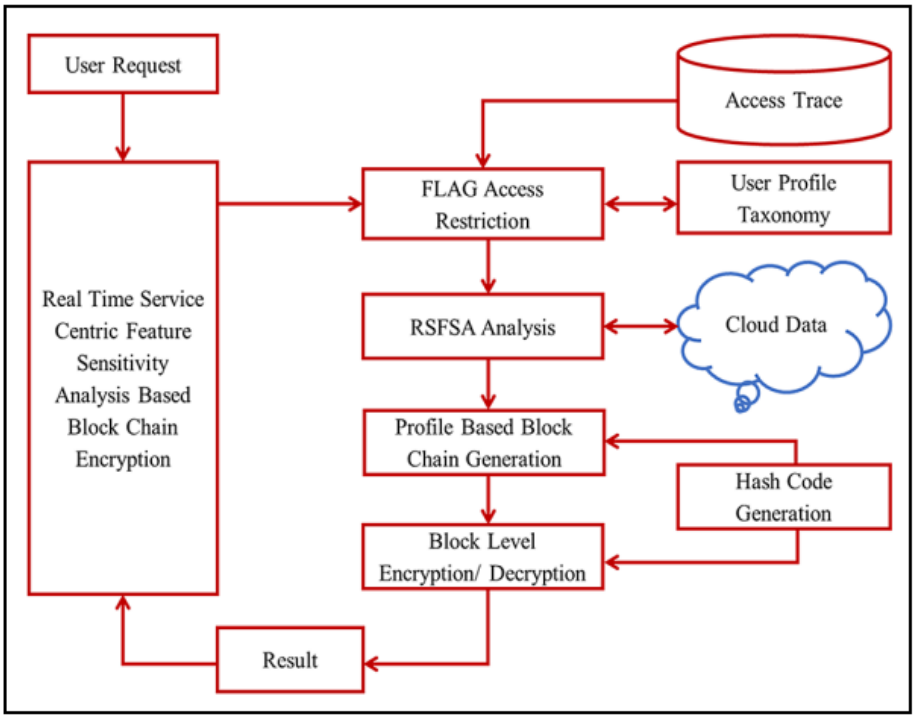


Figure 1. Functional architecture of proposed RSFSA model.

The suggested RSFSA algorithm was tested to see how well it will handle service hijacking attempts, malicious users launched data-stealing attacks and authentication attacks.

Service hijacking attack

It was tested by producing a K number of assaults aimed at hijacking the service by making service requests in the manner of a legitimate user. The suggested method confines all threats to a frequency of N/K, where N is equal to K. In all stages, the approach checks the user's trust.

Data stealing attack

The suggested RSFSA algorithm produces 100% tampering in this scenario since it validates the access grant on all levels. In terms of access limitation, the algorithm is capable of generating better results.

Authentication

The suggested RSFSA method is capable of validating the user profile for service access and limiting the attack from the very beginning.

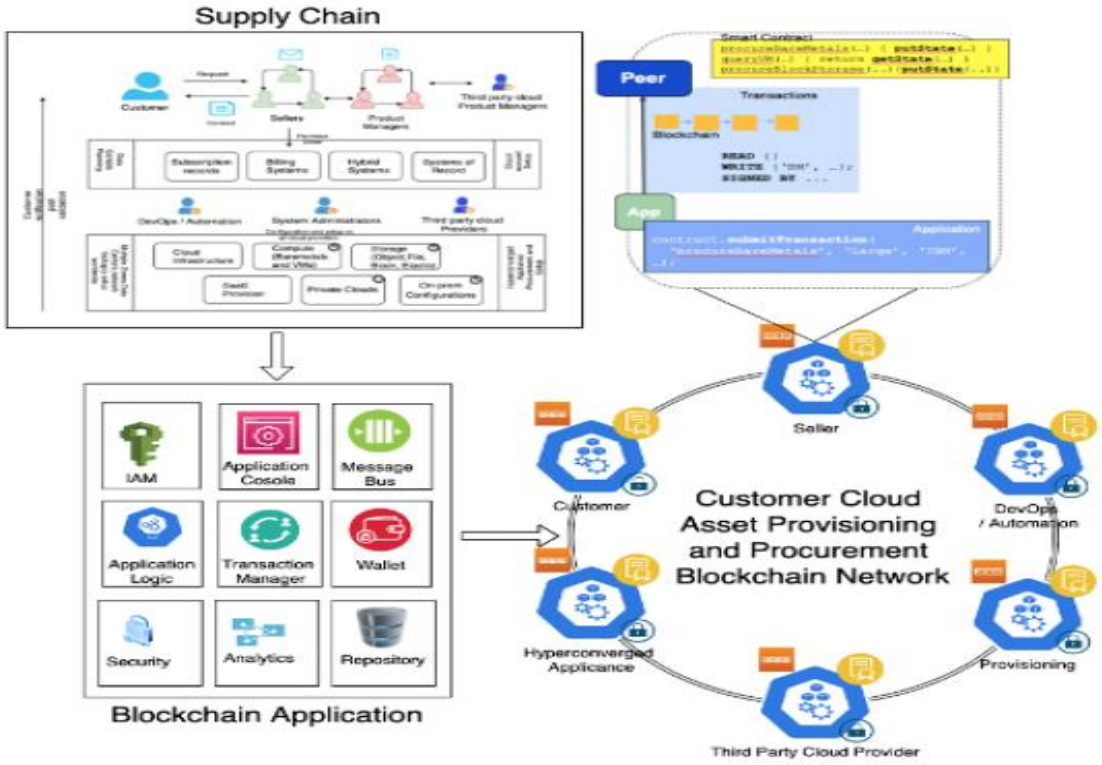
C3HSB

This technique is unique in that it scales the C3HSB blockchain topology outlined previously. Many components of the fulfilment process remain manual in many hybrid cloud supply chains in the business. This generates unnecessary delays, and providers may find it difficult to explain and obtain customer agreement for alternate alternatives.

With the following significant contributions, C3HSB tackles many of these problems by automating basic aspects:

- Using IBM Blockchain, designed and implemented a permission Blockchain-as-a-Service (BaaS) network.
- For C3HSB, created smart contracts and a blockchain application, as well as validated supply chain flows.
- Smart contracts that have been demonstrated can assist minimize reaction time and promote transparency, particularly when difficulties arise during the deployment orchestration of complicated cloud topologies.

C3HSB is a decentralized distributed architecture (see Figure 2) that uses blockchain and smart contracts to let cloud providers and their suppliers navigate their supply chains and engage customers with verifiable and immutable data. The provisioning of computing, storage, and network resources in the cloud was evaluated. The usage of IBM's permitted BaaS platform, which is provided by various cloud providers, was also considered. This makes it easier to build and implement a blockchain network by including built-in features for policy governance, smart contracts, security, and compliance, among other things.



In another case, C3HSB smart contracts are used to automate the execution of extra business procedures. The application executes process Notification in the customer's smart contract to alert the client about the situation when the system is back-ordered at a data center and the customer needs to be asked to choose other choices if the wait time is unacceptable. The client modifies the

order by designating an alternative datacenter with proven availability, as well as a different chipset or storage performance attribute, and the procurement process continues seamlessly with the amended terms and any cost implications.

The level of openness and access to data provenance that the framework gives to customers is unrivalled. Furthermore, the platform may take this a step further by notifying the sales leader via a smart contract that decides that the original purchase contract has changed and has to be amended. The speed with which this exchange can transpire, and the resultant non-repudiated transaction that meets all legal requirements as well that can be put in place was previously unimaginable in existing supply chain workflows. All peers have a vested interest in seeing that the order is fulfilled as soon as possible. It also aids companies in keeping their inventory current by allowing them to optimize their relationships with downstream suppliers and partners. An application was then created to execute the business logic in smart contracts to conduct activities on the blockchain ledger, such as creating, transferring, or updating cloud provider assets. For example, updating the status of bare metal and storage purchases. Identity and Access Management (IAM) was set up to authenticate people and systems, as well as analytics to warn on problematic third-party providers, and so on. It was observed that all transactions are logged in the blocks, as expected.

Conclusion:

To make the trust-based credentials system between user and the cloud service providers blockchain based authentication has been recommended. To enable the privacy efficient credentials verification while maintaining user control and privacy over the data this research report has suggested self-sovereign identity with blockchain system. Both the cloud service provider and the user would trust each other on financial terms and would ensure that both have a common goal. Identity owner and cloud service provider both contribute to the blockchain infrastructure. One user has separate identity that would be verified by the identity issuer at keep the proof of the verification of the credentials. The credentials and proofs would be then stored on the mobile device of the user, in digital wallet. The proof of the credentials is stored in the blockchain system. The Identities that are issued to the users are stored in the distributed IDs or DIDs on the blockchain. Identity verification system called Casper has been used to verify identities in the financial institutes. But the system should need to be extended beyond identity verification. In this platform the storage of the credentials takes place in the mobile device of the user acting as a digital wallet which would ensure the security and privacy of the data. SSI provides the correct degree of decentralization that supports the user centric identity model. For identity operation, transaction, secure data, and user authentication should be focused in future to understand the degree of centralization. SSI must have multiple interaction with blockchain and should need to ensure that both would bring the revolution in term of cloud and client-based operation whether its storage, operation or processes. With the implementation of both the scalability, operational costs and performance of many businesses would evolve.

In short, this report has discussed user data privacy, identity management systems, cloud services, privacy issues in cloud services, lack of user control, data integrity, confidentiality, and availability, different model of blockchain that could be used in the cloud technology to enhance privacy and security of the user data. SSI model architecture, design of SSI, blockchain identity model, enhancement of the user control, cloud-based infrastructure and trust model, challenges of self-sovereign identity system, and analysis of different approaches has been done in detail.

Appendix

DID: decentralized identifier is an identifier that does not need centralized registration authority due to it is generated cryptographically

Decentralized identity management system: that has decentralized approach of managing identity to avoid centralized control.

Claim: it can be made by either entity or the subject itself

DPKI: Decentralized public key infrastructure that relies protocols such as DID and verify public key encryption without using traditional certificate.

Digital identity: having particular digital content or credentials representing an entity.

Identifier: a value to a real world

SSI: self-sovereign identity that enable the identity holder to hold one or more identities.

CSP: cloud service provider

Bibliography

Alex Preukschat, D. R. (N.D). *Self-Sovereign Identity: Decentralized Digital Identity and Verifiable.* Shelter island USA: Amazon.com.

Amal Ghorbel, M. G. (23 January 2017). Privacy in cloud computing environments: a survey and research challenges. *The Journal of Supercomputing volume 73, pages2763–2800 (2017)*, 2763–2800.

Bandara, E. (2021). Casper: a blockchain-based system for efficient and secure customer credential verification. *Journal of Banking and Financial Technology (2021)*.

Foster, I. K. (2002). The Physiology of the Grid: an Open Grid Services Architecture for Distributed Systems Integration. *Global Grid Forum* .

Gaetani1, E. (2017). Blockchain-based Database to Ensure Data Integrity in Cloud Computing Environments. *In Proceedings of the First Italian Conference on Cybersecurity (ITASEC17), Venice, Italy.* Venice, Italy.

Gao H, H. W. (2021). he cloud-edge-based dynamic reconfiguration to service workflow for Mobile ecommerce environments. *QoS prediction perspective*, 1–23. .

Jiyi Wu, J. C. (2021). Blockchain-based trust management in cloud computing systems: a taxonomy, review and future directions. *Journal of Cloud Computing*.

Keke Gai, S. M. (2020). Blockchain Meets Cloud Computing: A Survey. *IEEE COMMUNICATIONS SURVEYS & TUTORIALS*.

Park, J. H. (2017). Blockchain Security in Cloud Computing: Use Cases,Challenges, and Solutions.

Soltani, R. (2021). A Survey of Self-Sovereign Identity Ecosystem. *https://www.hindawi.com/journals*.

Song, J. (2021). *How BlockChain Can Help Enhance The Security And Privacy in Edge Computing?* arXiv.

Tebaa, M. (2015). Hybrid Homomorphic Encryption Method for Protecting the Privacy of Banking Data in the Cloud. *International Journal of Security and Its Applications*, 61-70.

Wenjuan Li, a. L. (2020). Trust Model to Enhance Security and Interoperability of Cloud Environment. *Hangzhou Normal University, Hangzhou, Zhejiang 310012, China* (pp. 69-79). Zhejiang China : springer.com.

Zhou, Z., & Zhang, H. (2013). Prometheus: Privacy-aware data retrieval on hybrid cloud. *2013 Proceedings IEEE INFOCOM.* Turin, Italy: IEEE.