DATA PROTECTION THROUGH DATA SECURITY-AS-A-SERVICE USING BLOCKCHAIN ENABLED PLATFORM

Dr Magesh Kasthuri, Hitarshi Buch,
Krishna Moorthy and Vinod Panicker

Wipro Limited, India

ABSTRACT

Data access is inevitable in today's world and it is prone to threat attacks and hence data security is utmost important for any enterprise to handle industrial solutions. The economics of data being used across the industries rapidly growing in current digital world so the potential data related threats is also rapidly growing. Data security is an integrated solution component for any Enterprise solution but with the growing demand on data security and potential threat handling, Data Security as a Service (DSaaS)f is a new model widely accepted in modern age architecture in Blockchain and Big Data world combining the power of cloud based security services, decentralized network in Blockchain and tamper-proof ledger management. Any Enterprise Security architecture comprises of how data is handled in a secured way and how integration between services (consumers/producers or API interaction or any middleware services) handles data between them. Hence it is inevitable to that future technology adoption should include Data Security-as-a-service for zero-trust solution design complying with compliance and security standards for industry.

KEYWORDS

Data Security, Blockchain, Decentralized Ledger, DSaaS, Data Loss Prevention (DLP), User and Entity Behaviour Analytics (UEBA), Cloud Access Service Broker (CASB), Certificate Management, Key Management.

1. INTRODUCTION

Data has become the easy target for the cyber attackers due to its spread and availability. The data explosion or data breach has become a new criminal activity impacting business, government, and individuals. Protecting sensitive data is the key priority for the digital enterprises as innovative ways are being found to compromise the systems, attack the security, and steal data. Having a cloud-based offering is not a solution considering its centralized systems, interfaces and accessibility still gives a vulnerable target. Also, some of the companies are not ready to accept their business data or personal data is stored in the cloud or third party provider (SaaS). Hence, handling of data securely is the key in digital transformation is the need of this hour.

Data security is a practice and technique to safeguard the sensitive data in the digital ecosystem. It includes protecting the data from unwanted actions through unauthorized access, data theft and corruptions throughout the life cycle of a data. It aims to maintain the data confidentiality, reliability, availability, and data integrity. It encompasses the organizational policies, country specific policies and procedures to protect the data from destructive forces.

David C. Wyld et al. (Eds): MLDS, NECO, SEMIT, IBCOM, SPPR, SCAI, CSIA, ICCSEA - 2021
pp. 127-143, 2021. CS & IT - CSCP 2021 DOI: 10.5121/csit.2021.111811

In Cloud platforms, security and compliance feature for Federal Risk and Authorization Management Program (FedRAMP), US Department of Defence Architecture Framework (DODAF) and UK's Ministry of Defence Architecture Framework (MODAF) are supported by native services and during application design, supported cloud services for these compliance framework to be chosen for data protection activities.

This article provides foundational blocks of designing a DSaaS solution, possible DSaaS requirements at various states of data, DSaaS capabilities required, use cases and possible reference architecture for DSaaS by leveraging Blockchain solution. The use of Blockchain technology does not fully remove the inherent risks of data security as it needs to be pro-actively managed.

2. LITERATURE REVIEW

Subashini et al [1] explains the real-time common threats to data applications in cloud platforms like AWS or Azure and how it can be approached in different architectural models like Multi-cloud or hybrid cloud solution design. They have discussed how data at rest and data in motion is important in architectural decision for cloud application design.

Deyan et al. [2] in their survey paper explains how data privacy is important in various security and compliance requirements including country restrictions like regulatory policies and how it can be addressed in cloud platforms like Azure, Amazon and Google Cloud in both backend database platforms and in front-end application services. The author explains some key problems in data security including public access restrictions and data management issues.

Vladimir et al. [3] talks about some of the common data security issues in cloud architecture in a given domain solutions like Healthcare and Financial services and how common data security policies like General Data Protection Regulations (GDPR) helps in customer personal data security policies and adoption of customer identity access management (CIAM) solutions. The authors explains a data protection approach for Data Loss Prevention (DLP)

Ravi et al. [4] proposed a solution approach for integrating security services for data protection in cloud platforms including application, platform and infrastructure security services and how to address non-functional requirements for application design. The authors proposes an efficient User and Entity Behaviour Analytics (UEBA) framework for cloud data protection.

Eman et al. [5] has done detailed survey for data protection techniques in cloud platforms and how Software as a Service (SaaS) solutions like Microsoft Active Directory (AD) or Amazon Cognito along with Key vault and certification manager helps in data protection in two-way handshake during application integration services.

Vijay et al. [6] proposed a new-age service model in cloud platforms for enhanced application protection and how cloud services can handle data security in backend data store and integration services. They have also explained Five pillars of Well-architected framework covers data security as a pillar in enterprise application design in cloud platforms. The authors introduces Cloud Access Service Broker (CASB) method for cloud data security.

Mohammed et al. [7] introduced a new-age solution called SECaaS which means security-as-a-service framework which can be applied on cloud-based application design for data protection and how this can be an cloud agnostic solution to be used for any hyperscalers in cloud application design.

3. KEY COMPONENTS OF DSAAS

In this context, conceptualizing Data Security as a Service (DSaaS), that provides various data security requirements, data security capabilities, services, policies, procedures, and associated use cases forms an important design consideration for any Enterprise. Key Components of DSaaS are shown in below Figure-1.

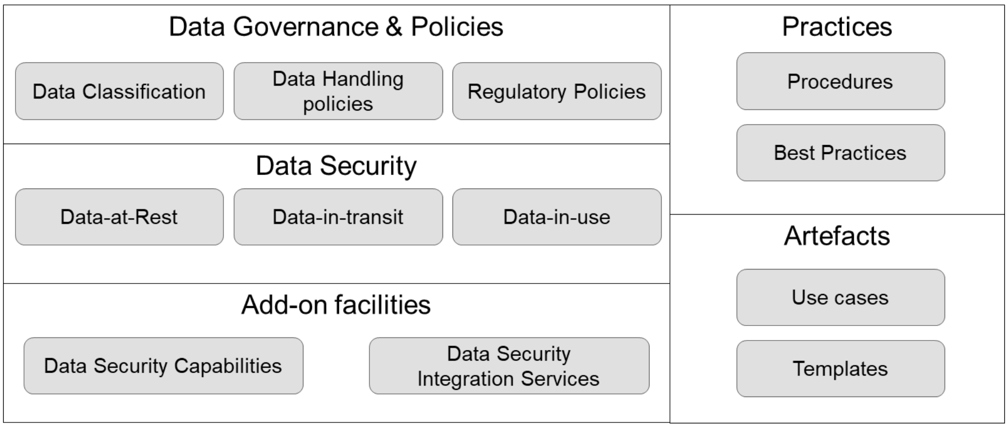

Figure 1. Key components of DSaaS

Blockchain platform can act as a main enabler for DSaaS in which a Distributed Ledger Technology providing an increased cyber resiliency and maintains ledger integrity because of its decentralized architecture, implementation of enhanced security frameworks for tamper-proof transaction, access patterns with no single point of failure (SPoF). The data is stored in blocks and connected with chain of blocks; thus, attacking a specific block does not affect the other blocks and the attacker needs to tamper all the blocks, but then detection is evident. The encryption and cryptography solution that Blockchain applications use to manage the data or transactions blocks protects individual transactions or records and the entire ledger. Thus, Blockchain proves to be a holistic capability to serve DSaaS requirements.

4. DATA SECURITY REQUIREMENTS

With external attacks account for the majority of data breaches suffered, Attackers take advantage of the fact that firms are interconnected and reliant on many other components of a broader business ecosystem. Emerging Data Protection regulations and Data privacy and protection compliance like General Data Protection Regulation (GDPR), General Data Protection Regulation (CCPA), Health Insurance Portability and Accountability Act (HIPAA) and Payment Card Industry Data Security Standard (PCI DSS) makes tough for organizations to identify what data falls under the umbrella of such regulations.

CXOs across different industries indicated that Personally Identifiable Information (PII) and Intellectual Property (IP) as the top two data types compromised in a breach followed by theft of payment and credit card data. Malicious internal incidents are on the rise and attackers are compromising employee credentials to data or tamper with it. Masquerade as insiders (with legitimate access privileges) in their efforts to access sensitive data causes potential data threats. Lot has changed with enterprises embracing cloud and cloud based services and increasing risk in data protection in such an environment. DSaaS (Data Security as a Service) like encryption as a service and certificate management as a service are now used by organization to improve their security posture in the cloud.

Considering the current exposure to the threats, the data security requirements are explained using its three states of data in a digital ecosystem. The three states of data are: **Data-at-Rest, Data-in-Transit or Data-in-Motion and Data-in-Use.** Any data can be exposed to threats when it is at rest, or in-transit or in-use. It requires a protection layer and robust data security solution to enable this requirement. There are multiple approaches to secure the data at various juncture and encryption plays a major role in enabling this.

4.1. Data-at-Rest

Data-at-Rest is generally termed as one of the states of digital data, where data is not moving, and the data is not being transferred or accessed. It is a stable state and physically stored when compared to the other states and it can be stored in cloud, end points(computers, PCs, mobile devices) file servers, tape drive, hard drive, computers, any physical devices, or archive storage/document management systems, etc. in any form.

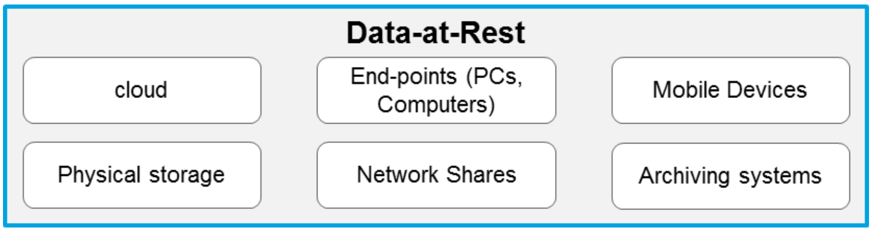

Figure 2. Data-at-Rest services

The data-at-Rest scenarios are increasing concern to organisations, business and government. So focus is to provide data security requirements for data-at-Rest to avoid malicious attacks and theft of physical storage media. In general, these data storage areas are highly protected with various defensive layers including security zones, firewalls, anti-virus layers and physical securities. However, these are not impassable.

The data security requirements start from where these data gets stored i.e., storage location, storage type, storage media and access requirements. Additionally, another critical point is to understand how it is classified based on the data classification requirements. So, key considerations on data security at rest are:

- **Secured, physical location:** Physical location where data is stored is one of the critical considerations whilst adopting data security at rest.
- **Size of the network:** Size of the network plays a major role here if a Blockchain network is not too large or well distributed and it becomes a potential risk for the attack.
- **Secured, physical location with classification of data storage at classified locations:** Ideal to consider multiple locations for multiple types of data based on the data calcification outcome. But it is also a challenge considering multiple copies of data can be available in PCs, Storage Devices, and Mobile devices. With the distributed architecture of Blockchain, it is feasible to store data at chunks across the locations.
- **Storage:** Data Storage on mobile devices is a key challenge considering use of these mobile devices are very common nowadays. Hence data protection that is stored within these mobile devices is another challenge to consider.
- **Infrastructure set-up:** A key parameter that influences the Data-at-Rest is its infrastructure set-up and right-levels of secured server farms with right patches up-to-date. It includes redundant, highly available distributed architecture to have a reliable storage. In addition to this, having data controls on cloud storage is good, however the actual encryption keys are

owned by the storage provider and hence control is not with the company. Applying right sets of protection layers for the storage media also a critical requirement to avoid the potential risks.

- **Access Policies:** Considerations to be applied with strong access policies to prevent these risks, as longer the data remains unused in storage, the more likely it might be at risk. Additionally, admins of permissioned Blockchain networks can hamper the blocks e.g., rewriting blocks history, delete resources etc.
- **Blockchain users backing up the Private keys in physical media:** Theft of these private keys is another big challenge as these keys are used for critical operations in the Blockchain ecosystem. So, having a secured key governance practices are key for this.
- **Automated approach to evaluate the security requirements:** Automated process to evaluate the security risks considering automated data classification framework and data protection evaluation to identify potential risks.

To prevent these data being accessed or stolen, companies apply security protection measures with additional layers of defence such as data encryption, password protection or both. The security options used for this type of data are broadly referred to as Data-at-Rest Protection (DARP). The companies should consider right data security requirements for the data-at-rest as part of the preventive Data Loss Plan along with country specific data protection regulations as applicable.

4.2. Data-in-Transit or Data-in-Motion

Data-in-Transit or Data-in-Motion or Data-in-Flight is termed as data which is travelling from one point to another through public or private communication channels such as messaging, mailing, sharing, chats, cloud, collaboration tools (eg. Microsoft Teams, Zoom meeting), applications etc. It is data moving in a network i.e. opposite in meaning for Data-at-Rest. In the current digital world, data is shared across in multiple ways to collaborate each other. Since data is moving in a network and it contains many nodes, where different access points are connected to the same network, hence data-in-motion needs to be protected.

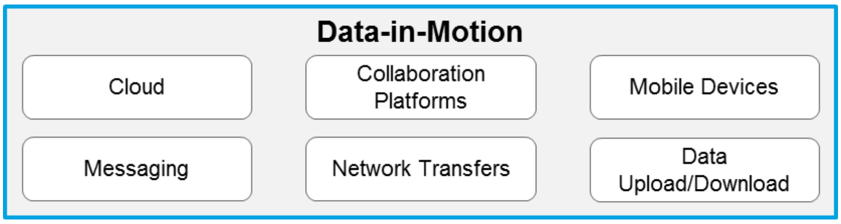

Figure 3. Data-in-motion services

Key considerations for Data in motion are:

- **Need for a Transport layer security:** Transport layer security to be considered for all the sensitive data that is being traversed in the network using SSL/TLS protocols. However, the challenge is, there are infinite number of mechanisms and means of channels for data sharing. Examples include email encryption using PKI (Public Key Infrastructure), File Transfers using Secured File Transfer Protocols (sFTPs) and HTTPS.
- **Prevent all:** Prevent accessing the sensitive information even from the root users/administration users.

- **Data transfer/sharing mechanisms:** Moving sensitive data from one location to another using USB drives, uploading to cloud storage, web contents etc.
- **Secured access from individual workstations vs End-user attacks:** It is common mechanism to leverage end-user access points for malicious practice as it is an entry point. The attacker will gather user credentials to infiltrate the network
- **Data Leak Prevention:** Identification of the purpose of retrieving sensitive data whilst it-in-transit is a challenge. Also, unable to identify potential risks when data is accessed by the endpoint or at receiving end as control of transport layer is no more applicable.
- **Private Key management:** The recent cyberattacks have proved that they are achieved through stealing end-user keys and then using those to enter the network, instead attacking the server farm directly through malware/virus/physical assault. So, the private key management is the key consideration for data protection.
- **Automated approach to evaluate the security requirements:** Automated process to evaluate the security risks using suspicious activities in the network, diagnose potential threats and proactively improve the security.

4.3. Data-in-Use

Data-in-Use can also be called as active data in the context of being worked in a database or accessed by an application. In general, any data is opened by either application or users, for its consumption or any treatment, then we can consider the state as Data-in-Use. In this state, data is more vulnerable considering data is decrypted whilst processing the data. The requirement is to have additional controls prior to access the data. The examples include working/accessing the Applications (on premise, cloud, and mobile) and Database.

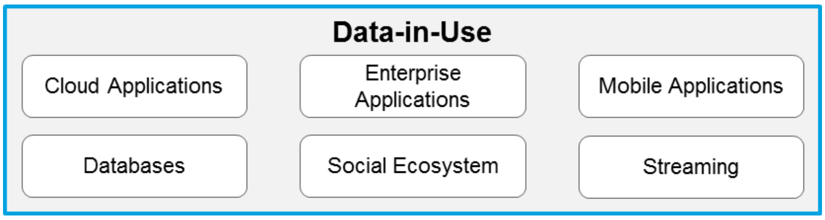

Figure 4. Data-in-use services

Key considerations for Data-in-use are:

- **Authentication/Identity Requirements:** To ensure right user is identified and authenticated to access the data/platform.
- **Authorization/Role based access:** On authentication, user/program should consider right persona and sufficient privileges to process/access/operate the data leading to focus on data visibility and it is usability.
- **Difficult to control the data after it is being accessed:** After users are authenticated and authorized; the user can take a photo and use for malicious purposes. These controls are difficult to put in place.
- **User code-of-conduct:** It is a challenge to apply user code of conduct considering human element in the public domain/digital ecosystem and remote usage environments.
- **Encrypting Memory:** Encrypting memory to be considered as it prevents data accessibility from malicious users but with the trade-off of performance due to additional activity of encryption and decryption.
- **Data Sharing:** Rights to download the documents or sharing, forwarding etc is being widely used and needs to be considered as part of data security requirement.

- **Automated approach to evaluate the security requirements:** Automated process to evaluate the security risks using suspicious activities in the network, diagnose potential threats and proactively improve the security.

Compromising of data-in-Use enables access to encrypted data-at-Rest and data-in-Motion, through accessing the private leys of data-at-Rest or data-at-Motion and thus manipulating the original content will lead into data security risks.

5. RELEVANT USE CASES

Relevant use case have been identified where data security service is traditionally used to prevent data leak. Since data security implementation have move forward from financial services to Digital document management, Digital Authority and Signature, Digital Identity solutions, Smart contract solutions for Supplychain services, Digital record management using Distributed ledger services, Decentralized network services for workflow management and security services, Data security as a service adoption to various industrial usecases have been emerging for the last five years across Healthcare, Manufacturing, Retail and Non-banking domain in Financial services like Claim processing and verification in Insurance, Investment management in Capital Markets to name a few.

There are various usecases in Healthcare and Medical service Industry being explored to use Blockchain platform. In addition, post COVID19 situation, Gartner research says that there are some key reasons for Blockchain based Healthcare application being evolved where data security is the primary concern as explained below:

- With disruption in Financial services, urgent transaction in cash liquidity is important to expedite business transactions such as purchase of medical devices, realization of purchase order etc.,
- Improve trust in supply chain processing and track assets with high transparency.
- Digital Document signing and asset management to be suitable for Insurance Claims processing.
- Rapidly processing Supply chain services in invoicing, purchase order approval, Zero touch payments and trustless party handling
- Using digital currency and crypto-currency for Supply chain financial activities.

Some of the popular usecase scenario where Data security as a Service can potential create higher impact to business risk management are discussed below.

5.1. Use case – Electronic Medical Record (EMR) Services

Electronic Medical Record (EMR) or Electronic Health Record (EHR) are very important to be handled safely and data operations with EMR is a costly operation and important for the Healthcare Industry. When a patient is admitted for medical emergency, understanding the patient’s medical history is very important and prima-focus step before starting the treatment and hence data handling with utmost speed with data privacy handling is always a demanding scenario in Healthcare industry.

Growing nations are handling these Medical Records in central repository which can be accessed across different hospitals. These kind of EMR has popularly demanding as it can be used for Hospitals for patient history and Insurance companies for claim verification.

Hence Data Security as a Service (DSaaS) is one of the most sought out solution approach for EMR with Blockchain based solution, where we can implement Smart contract based solution to enable alerts based on fluctuations in these readings so that hospitals can remotely take active measures to handle patients efficiently.

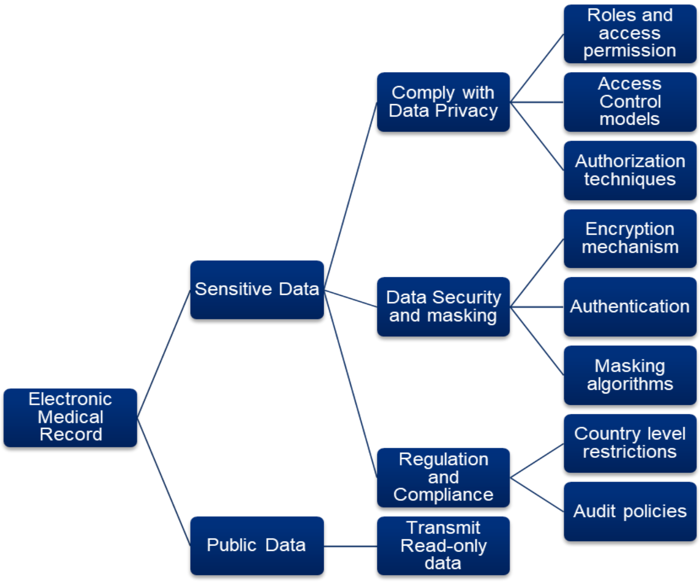

Figure 5. Data Security path for EMR

This can be achieved with a Blockchain based solution, where IoT transmission between hospitals across different locations can trigger events like search patient records, patient treatment history, doctor's advice during previous ailment etc., Based on these events, actions can be associated for automating the approval process in order to ensure the entire EMR transmission is automated completely without manual intervention and at the same time without compromising security in openly transmitting records to unauthorized parties using Blockchain security services.

5.2. Use case – Retail Supplychain Services

Implementing a Retail Supplychain solution involves handling people (governance), process (activities and information flow) and technology (using blockchain smart contracts to accelerate the workflow or using faster data encryption and decryption mechanism for both Data at rest and data in transit).

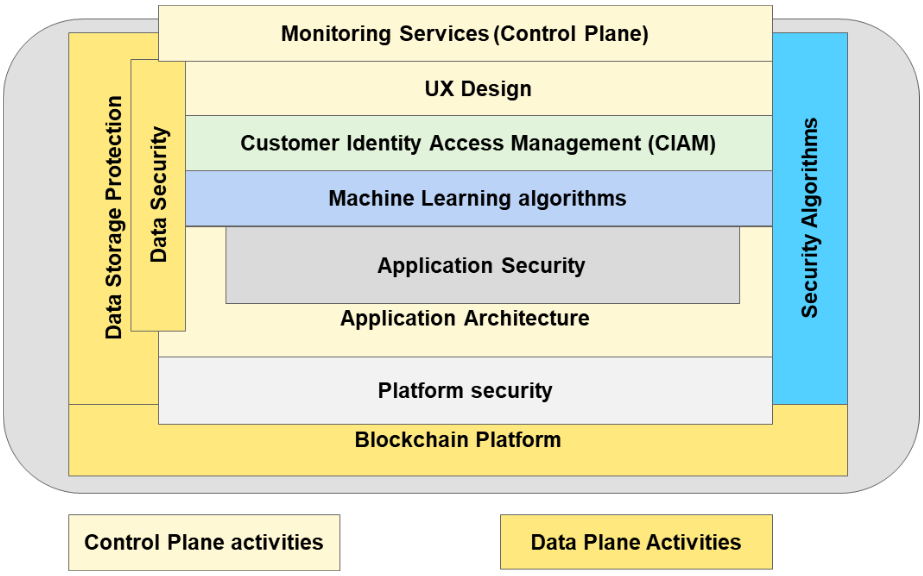

Figure 6. Data Security integrated in Retail Supplychain platform

Operational activities in a Blockchain platform are typically divided as Control Plane and Data Plane as shown in Figure 6. Control Plane is the driver which can be used to create and manage any cloud resources. Data plane is the operational activity which handles the capabilities of the resources created/managed by Control plane. These kind of resource management through Control plane and the Control plane helps to control and manage security services, auditing, policy driven activities, activity logs and resource hierarchy services.

During Retail Supplychain process, multiple parties are involved in purchase order, invoicing, payment approval, payment transfer etc., and hence in a traditional approach it takes considerable time to execute the entire Supplychain process. Using a Blockchain platform, helps to accelerate the process by involving multiple parties (node approver and owner) to execute the entire workflow operation. But Data Security could be a roadblock to address in such case. Hence implementing a Data Security as a Service platform integration with private Blockchain platform can help to develop a zero-trust security integration for the entire Supplychain process.

6. DSAAS CAPABILITIES

Data security or protection can be provided by ensuring that a specific set of technology services and patterns are used in conjunction with blockchain to ensure that data breaches are prevented. The factors based on which this can be ascertained are:

- Data and documents are the most important asset that any enterprise application needs to protect. The level of protection to be enforced must be determined based on how confidential and sensitive the data asset is to the organization.
- Another factor that needs to be considered is the stage at which data security needs to be implemented. As described earlier data security mechanisms for data-at-rest will be little different from techniques required to protect data-in-use.
- Finally, irrespective of the data security mechanism used effective monitoring for proactively detecting data breach will always be required

In order to fulfill the enterprise grade data security requirements, DSaaS will require specific set of components and services as explained below.

6.1. Data Storage Protection Services

Data and document storage requirements require techniques that protect data-at-rest. Blockchain enabled data storage techniques would comprise of:

i) **Blockchain for Data Storage:** Data that does not contain confidential or personal information and is smaller in size (<10KB) can be directly stored on blockchain. The append-only ledger of blockchain ensures that data cannot be altered unless it undergoes the validation and consensus in-built into the blockchain network. Depending on the blockchain platform being used; there can be limitations on how data can be queried and may have to be mitigated by using an offchain read-only storage.

ii) **Blockchain for Proof of Existence:** Depending on the blockchain platform being used there can be limitations on how data can be queried. When data and/or documents are of larger size then traditional centralized storage mechanisms such as databases and filesystems must be used, which are susceptible to breaches as the content can be changed without getting detected. In such scenarios, storing the digital fingerprint of data asset as a one-way hash in blockchain is the recommended approach. This ensures that digital representation is always available on blockchain which can act as ir-refutable timestamped proof of the state of the data asset. A verification service will be provided so that the actual data or physical document can be verified against its proof recorded on blockchain. Additional metadata such as ownership, reference to the physical storage and access rights will also be maintained in blockchain for easy tracking of transaction.

iii) **Data Sharding:** For highly sensitive data / documents storing it at one central location can pose a high security risk. Therefore, sharding of data and documents into multiple parts and storing it on distributed nodes is recommended. Technology components like IPFS (Inter Planetary File System) provide a protocol and peer-to-peer network for storing data in a distributed file system. Each file is assigned with a unique identifier in a global namespace and then stored in a distributed manner, which can be reassembled on demand. The fragmented parts ensure that the document cannot be tampered with to protect the data.

iv) **Data Encryption:** Data-at-rest in production-grade systems is maintained in encrypted format. Symmetric encryption using industry standard algorithms (AES, DES) will be used for storing data on blockchain and when using sharding techniques to further strengthen data security. Asymmetric encryption will only be used for data storage when the data stored is intended to be consumed by specific stakeholder.

6.2. Data Usage Protection Services

There are several proven techniques for securing data-at-rest and there are several layers of security that can be implemented above the data / file storage layer. But protecting data-in-use is more complicated considering that its usage is spread across all the components of a solution. Some of the protection services recommended for DSaaS solution are:

i) **Secured Multi-Party Computation:** SMPC is a technique of distributed computing and cryptography which enables entities (individuals, applications or devices) to work with data while ensuring that the data and/or encryption keys are kept in a protected state. Multiple entities can participate in handling these confidential data. SMPC provides a new model for protecting data-in-use by strengthening the traditional security mechanisms. SMPC also helps in mitigating data residency issues by eliminating the single-point-of-failure risk because of the ability to “split” the confidential data or cryptographic key into multiple parts that can be re-assembled on-demand at runtime when a data transaction is executed. SMPC and blockchain are complementary technologies and it can be applied to all data lifecycle stages.

ii) **Decentralized PKI:** When data is being exchanged or used between different parties or systems, then its security is heavily reliant on PKI (Public Key Infrastructure) usage for authentication and asymmetric encryption of data. For production-grade systems, the public and private keys are supposed to be procured from a trusted CA (Certificate Authority). But these centralized CA are also susceptible to breaches which can lead to key theft and impersonation. Therefore, usage of Decentralized PKI on blockchain based framework like Web of Trust model is recommended approach for complete transparency and security. The different capabilities of the DPKI system that will be available on blockchain will include:

***ii-a) Identity Registration* –** Smart contract-based registration of UUID (user's unique identifier) and their public key. The rules governing registration and renewal of identifiers will be transparently maintained on blockchain.

***ii-b) User-controlled key generation* –** For the registered identity, the generation of public-private keypair will be initiated by the user. Private keys must be generated in a decentralized manner under user's control. User may authorize an agent to manage keys on their behalf.

ii-c) Master key and sub-keys – Each user or entity whose identity is registered will be assigned a master private key and sub-keys from the master key. The sub-keys will be maintained as that identity's metadata and used for transactions signing related to that identifier, whereas the master key controlled by the user will be used to manage the identity's metadata. To ensure that master key is secured and can be recovered, SMPC (Secured Multiparty Computation) will be used to shard and store parts of the master key.

ii-d) Public key availability – All ledger participants will get automatic access to the identities public key which can be used for asymmetric data encryption

ii-e) Identity Revocation – Identity can be deleted or revoked only by the identity holder or via a workflow process that required multiple entities to provide their approval

iii) **Blockchain based ACL:** Access Control Lists or registries managed via smart contracts on blockchain will provide security rules to be enforced by last-mile security agents. The advantage of maintaining these rules on blockchain is that any change to the access levels for data or transaction type will always be available on blockchain as an audit trail. This technique also ensures that data leakage is prevented at the blockchain layer because smart contract-based authorization controls who can read or write to blockchain.

iv) **Multi-Signature Accounts**: Critical, high value transactions will be protected by ensuring that multiple signatures are required before the transaction is accepted and processed further in blockchain. This technique can be used for both onchain and offchain data. In case of offchain data, the smart contract events will be triggered when the required signatures are in place based on which the offchain execution can be controlled.

v) **Smart Contract Oracles**: Since smart contracts are not allowed to directly communicate with systems external to blockchain which is required to ensure that the smart contract output is deterministic. Therefore, smart contract oracles are required to supply the external data to blockchain. Data security rules implemented on blockchain can require validations from an external system. This can be enabled by initiating the Oracles component that listens to smart contract events and automatically fetches external data and supplies it to smart contract. This pattern ensures that smart contracts can obtain external data on-demand and make informed decisions.

vi) **Blockchain based UEBA**: User and entity behaviour analytics or UEBA arose out of the malicious behaviour by users and other entities. Self-sovereign Identity (SSI) based Blockchain trust networks have a critical role in ensuring that UEBA can be done responsibly keeping user privacy as top priority. UEBA uses machine learning and algorithms to strengthen security by monitoring users and other entities, detecting anomalies in behaviour patterns that could be indicative of a threat. A SSI based Blockchain network can be used to present proof of normal behaviour of user rather than collect raw user data in a UEBA. This is a proactive approach to security and gaining visibility into user and entity behaviour without compromising on user's privacy at any point in time.

7. Reference Architecture

DSaaS (Data Security as a Service) architecture will comprise of all the components and services mentioned in previous section in a layered architecture. The services can be leveraged by any kind of solution irrespective of whether blockchain is used not. As shown in Figure-1, the DSaaS Architecture comprises of well-defined microservices and underlying components.

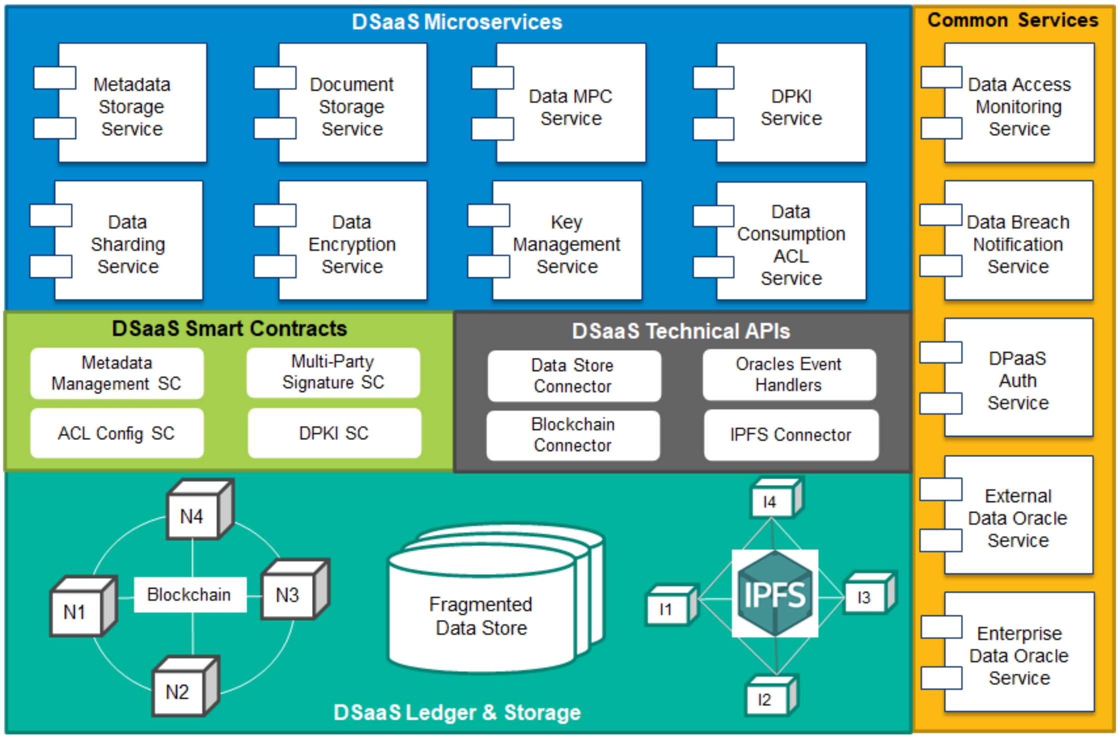

Figure 7. DSaaS Reference Architecture

The reference architecture comprises of the following of key building blocks that are segregated into different buckets.

7.1. DSaaS Micro Services

Microservices architecture will be followed to build APIs with the right granularity and cohesive functionality and the components in the architecture are:

Metadata Storage Services: These are generic APIs which will provide ability to the client application to store metadata related to their assets and process in blockchain so that data protected at rest with complete provenance and audit trail. This can be used to store small sized data, or it can be used to store the reference and digital representation of large sized data or document

Document Management Service: For protecting large sized documents this API can be used to store the physical document on IPFS and its metadata stored on blockchain. Documents can either be uploaded as an attachment or a shared folder be specified for upload of very large documents.

Data Sharding Service: Highly sensitive large sized data and/or digital assets will be protected by using this API that performs data fragmentation and stores in an encrypted token form in the file system. The metadata and its associated access rights are maintained on blockchain using which the data is reconstructed by decoding and combining all fragments as shown in Figure-8 below.

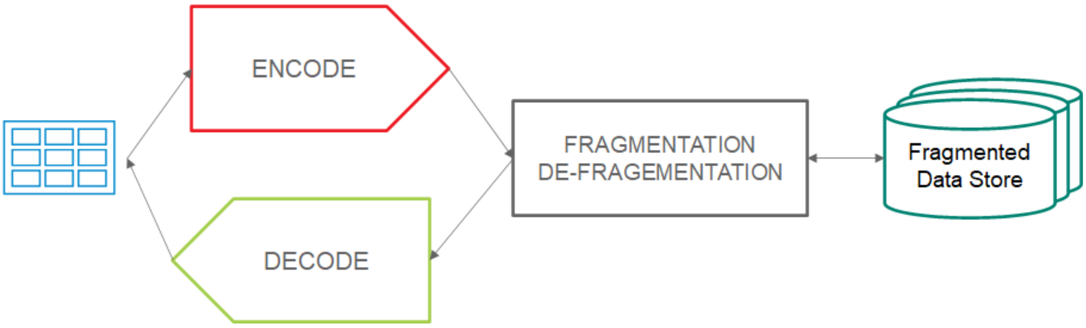

Figure 8. Data Sharding / Fragmentation

Data Encryption Service: This API will provide options of using symmetric encryption algorithms, which can be used to encrypt / decrypt data before storing it on blockchain or in the fragmented file store.

Data Multi-party Commutation (MPC) Service: This set of APIs enable the usage of multi-party computation capabilities ranging from enrolment of entities that would participate in the computation process to low level cryptographic functions as well as interactive and non-interactive functions required for effectively using MPC to protect the data-in-use.

DPKI Service: This set of APIs will help client applications to leverage the Decentralized PKI capabilities which involve identity registration, key generation, revocation etc. by leveraging the corresponding smart contracts deployed on blockchain.

Key Management Service: Generic APIs which will allows other microservices and client applications to be able to securely generate public-private key infrastructure using blockchain.

Data Consumption ACL Service: This API will allow setting up of registry with roles and permissions on blockchain that would be used for authorizing access to different types of data based on its sensitivity, confidentiality etc.

Common Services: APIs which will cover the cross-cutting concerns that cover monitoring, notifications, authentication / authorization as well as Oracles functionality to connect to external and enterprise data sources.

7.2. DSaaS Smart Contracts

The smart contacts layer represents the storage and logic deployed on blockchain which is required to provide functionality for data storage functionality and advanced functionality related to SMPC, DPKI microservices. Each smart contract will comprise of multiple functions with the appropriate role-based authorization checks.

7.3. DSaaS Technical APIs

This set of APIs are fine-grained technical APIs which are required for interacting with blockchain, IPFS and other data stores. This will also include the smart contract event handlers which will trigger the retrieval of external data from Oracles.

7.4. Blockchain

This is a distributed ledger network that forms the core of the DSaaS offering. A minimum viable number of nodes will be required to maintain this blockchain network to ensure high availability. A permissioned blockchain platform will be used so that it provides the required smart contract functionality as well as higher performance.

7.5. IPFS (Inter Planetary File System)

This component will provide a decentralized network to store and maintain the physical documents. Each document stored in IPFS will be allocated a unique hash identifier which will be maintained on blockchain.

7.6. Fragmented Data Store

This component will comprise of multiple logical and physical partitions in which the sharded data will be stored in encrypted format.

8. DSAAS ENABLED IMPLEMENTATIONS

The previous section describes the reference architecture comprising of the building blocks of a DSaaS framework. It is pertinent to outline how this reference architecture can be used to fulfil the data security requirements of the use cases described above. The actual realization of the architecture may vary based on the functional architecture, so the best fit DSaaS component is outlined in this section.

In the Electronic Medical Records (EMR) user scenario, some of the key data protection services required for this use case are:

Metadata Storage Service: Protecting the PII (Personally Identifiable Information) data of patients is one of the key requirements so health records against the patient's UUID on blockchain. For highest level of data security, the data fragmentation services to shard and store patient's health records can be leveraged.

DPKI Service: Identities related to patients, health care providers so that public key infrastructure is generated and distributed securely to each user.

Data Consumption ACL Service: Each actor in the EMR use case will have specific data access rights. Therefore, a default ACL limiting access to ledger data will be setup. Additionally, dynamic update of ACL will also be enabled for scenarios where patient provides their consent and agrees to share their data with a particular health care provider.

Document Management Service: The detailed medical comprising of pathology tests, X-ray, CT-scan reports etc. will be maintained by the health care provider's in electronic format. If physical access to the document needs to be made available, then the document storage services can be used for maintaining such documents in IPFS. If the documents are centrally managed, then their unique digital fingerprint will be maintained on blockchain via the metadata storage services.

In this use case, we design a solution which is PII compliant and uses Data storage protection services leveraging blockchain to provide DSaaS capabilities with zero-trust security platform for integrated solution.

In the Retail Supply chain User scenario, Supply chain is a vast area and hence the data security requirements also vary depending on the functionality being addressed. Some of the key data protection services required for this use case are:

DPKI & ACL Services: The identity and access management requirements for this retail-centric use case will require these components.

Data MPC Service: Large orders and high value payments will leverage the MPC services to ensure that the transaction signing key is sharded and secured to prevent fraudulent transactions by obtaining multi-party confirmation.

Data Encryption Service: Confidential data transactions in supply chain such as trading agreement, purchase order management, invoice generation and settlement etc. will leverage the encryption services to ensure that the data can be read and processed by the authorized recipient only.

Data Oracle Services: Supply chain logistics and financial transactions may require data for validation or processing purposes from external sources such as retrieving the latest forex rates etc. Similar order and product management will require validations from data quality perspective from the enterprise business applications. This can be achieved by using data oracles services to interact with external and/or enterprise data sources.

In this use case, we develop a solution which uses Customer Identity Access Management (CIAM) like PingIdentity or Gigya and other pluggable interface components of DSaaS reference architecture like microservices, connectors and API services for integrated Retail Supplychain solution design.

9. FUTURE SCOPE OF WORK

Data security as a Service can be enhanced in future as cloud agnostic solution and blockchain agnostic approach to be used for any cloud platforms and using any blockchain platforms to enhance its features with cloud native service integration for better solution approach for enterprise application design.

Also, these features can be developed in future using Microservices architecture along with polyglot database for better flexibility to integrate with multiple application and reuse the asynchronous communication services for internal and external application integration.

10. CONCLUSIONS

Data Security is the key to strengthen the Enterprise architecture when handling workflow operation involving multiple parties. For various industries like Fintech (Insurance, Payment, Investment banking) and Healthcare (EMR/EHR, Medical Retail Supplychain operations), it is important that an efficient pluggable DSaaS integration is incorporated which can help in business agility, cost efficiency and improved Governance and security compliance.

As shown in the reference architecture of DSaaS solution, there are many pluggable components which can be used to integrate a business agile solution for integrated Enterprise security to enable industry level compliance like FedRamp, PII, TOSCA, PCI or HIPAA compliance service functions.

REFERENCES

[1] Subashini, Subashini, and Veeraruna Kavitha. "A survey on security issues in service delivery models of cloud computing." Journal of network and computer applications 34.1 (2011): 1-11.

[2] Chen, Deyan, and Hong Zhao. "Data security and privacy protection issues in cloud computing." 2012 International Conference on Computer Science and Electronics Engineering. Vol. 1. IEEE, 2012.

[3] Getov, Vladimir. "Security as a service in smart clouds--opportunities and concerns." 2012 IEEE 36th Annual Computer Software and Applications Conference. IEEE, 2012.

[4] Kumar, P. Ravi, P. Herbert Raj, and P. Jelciana. "Exploring data security issues and solutions in cloud computing." Procedia Computer Science 125 (2018): 691-697.

[5] Mohamed, Eman M., Hatem S. Abdelkader, and Sherif El-Etriby. "Enhanced data security model for cloud computing." 2012 8th International Conference on Informatics and Systems (INFOS). IEEE, 2012.

[6] Varadharajan, Vijay, and Udaya Tupakula. "Security as a service model for cloud environment." IEEE Transactions on network and Service management 11.1 (2014): 60-75.

[7] Hussain, Mohammed, and Hanady Abdulsalam. "SECaaS: security as a service for cloud-based applications." Proceedings of the Second Kuwait Conference on e-Services and e-Systems. 2011.

AUTHORS

Dr. Magesh is a Distinguished Member of Technical Staff at Wipro. Magesh holds a Ph.D in Deep Learning and Genetic Algorithms. He is a senior member of IEEE and has published more than 50 articles in OpenSource For You, PC Quest, Cutter Business IT Journal and other notable international journals. He has also published around 480 thought leadership articles on AIML, Blockchain, and Cloud on LinkedIn with the hashtag #shorticle.

Vinod Panicker is a Chief Architect & Distinguished Member of Technical Staff with over 19+ years of software development experience. Vinod currently leads the Blockchain initiative for the Cybersecurity & Risk Group at Wipro. He is an expert in open source and crowd sourcing platforms. He was the Lead architect for Wipro’s inner sourcing platform, Open connect and helped it scale seamlessly to over 30K users.

Hitarshi has 20+ years of experience in IT architecture, consulting, design and implementation using blockchain, API, SOA, BPM and Java/J2EE technologies. He has experience in IT transformation and modernization initiatives and has provided enterprise-wide SOA-based solutions.

In his current role, at Wipro Technologies as a Chief Architect in Service Transformation at Wipro, he leads the Center of Excellence initiatives as part of the Blockchain practice. His charter involves applied research and building technology assets around blockchain protocols such as Hyperledger, Quorum, Besu, Corda, Multichain, Hedera etc.

Krishna Mty is a Lead Architect and Distinguished Member of Technical Staff in Wipro and part of Integrated Digital Engineering and Application Services.