

Blockchain-based Cyber Forensics Chain of Custody Management with Hyperledger Fabric

Nischitha T S¹, Dr. Purohit Shrinivasacharya² and Diya S³

¹⁻²Siddaganga Institute of Technology/Information Science and Engineering, Tumakuru, India

Email: {tsnischitha09@gmail.com, purohitsn@gmail.com}

³CDAC/Cyber Security Group, Thiruvananthapuram, India

Email: diya@cdac.in

Abstract—The current solution in cyber forensics for storing and sharing case records is done through mail because the shortage of trustable and reliable forensic data sharing mechanism. Blockchain has the potential to improve chain of custody by putting the case at the center of the system and enhancing case data protection and interoperability. Blockchain is a decentralized framework providing a cryptographic guarantee of data integrity, security, privacy, and smart contracts for data access. This paper introduces various concepts that are used in Hyperledger Fabric, proposes a permissioned blockchain system solution for storing and sharing case records using the Hyperledger fabric framework. Police investigators have full control of their case information and authorize the forensic analysts that can view case details by using grant and revoke access mechanisms and also presents the challenges faced, and issues to resolve, before the successful adoption of Hyperledger Fabric in chain of custody management system.

Index Terms— Blockchain, Fabric Framework, Chaincode, Certificate Authority (CA).

I. INTRODUCTION

Digital evidence management systems are built to capture all details of cases, the efficiency desires to rise even further strongly. The data for status (e.g., evidence collection, analysis) of the case is tremendously useful in respect to digital crime. The growth of internet users and cyber-physical devices rises cybercrimes and research for improved digital forensics for cyber cases. The greatest challenge in digital evidence chain of custody management is securing the internet-enabled device's security [7]. At present, numerous systems have been intended to make use of mobile devices to deliver case mobile-centric amenities to higher officers. Although the current digital evidence management system can have potential advantages, it faces certain obstacles and restrictions with the most technological limits and organizational constraints.

In the digital forensics' domain, identification and differentiation among forged and original events while investigating a crime with digital media are the limitations. These aspects are prominent for preservation and information integrity [9]. Another challenge is proving the incident's validity and authenticity [3]. Additionally, collecting digital evidence remotely is vulnerable while acquiring evidence on network sources including handling data nodes from the discovery of remote evidence recovery, reconstruction, and authentication [10]. This process of cyber forensics chain of the custody management is insecure while gathering, storing, evaluating, and preserving digital evidence [11, 15]. Since the involved database system is incomplete to sustain the confidentiality, and integrity of the digital evidence as well as the chain of custody events during collection,

storing, transferring, interpreting analysis of the evidence while investigating to solve the crime incident [4]. The modern mechanism of decentralization blockchain provides strong solutions to digital evidence management system security and trustworthiness issues. Another interesting aspect of blockchain technology is disregarding intermediaries. In recent years blockchain has been used to dig new paths after cryptocurrencies. This support attaining transparent, secure, and immutable investigations that were difficult to modify, in consort with secure storage and encrypted preservation [12]. Combination of blockchain and a chain of custody, including all aspects of privacy, the validity of safety, and access to reach provenance, transparency, and traceability of investigation [5].

II. BLOCKCHAIN

Blockchain is a decentralized, trustworthy distributed ledger on a peer-to-peer network that consists of a list of chronologically ordered blocks. Every block consists of a hash of a previous block and hence forming a chain. The first block in the blockchain is the genesis block and the block before a given block is called as its parent block. A block contains a header and body as in Figure 1. The block header consists of

- Version: Specifies the endorsement rules.
- Hash of the previous block: This is to ensure that without changing the present block's header, it is impossible to modify its preceding block.
- Time stamp: Time at which the block was created.
- Merkle root: Computed hash of the overall transactions, ensure that without changing the present block's header, it is impossible to modify its communications stored inside a block.
- Nonce: 4byte unique number used only once in the communication.

The core components of blockchain architecture:

- Node/Peer: Device (computer) existing inside a network that has a copy of each transaction.
- Transaction: Exchange of information among the two addresses.
- Block: Group of transactions that is distributed among all the nodes in the network.
- Chain: A series of blocks arranged in chronological order.
- Miners: Before adding any transactions to a new block node, selected nodes carry out the block verification.
- Consensus: It is a mechanism through which all of the blockchain network's peers come to an agreement about any transaction in the blockchain.

III. CYBER FORENSICS

Cybercrime Investigation mainly involves collecting and analysing digital evidence which links criminal activities to the person committed the crime. The biggest challenge during the investigation is to ensure the integrity, authenticity, and auditability of digital evidence as it moves along different levels of hierarchy in the chain of custody. Huge quantity of information generated by billions of devices on the Internet makes the process more challenging. Thus, there should be some mechanism to ensure integrity and authenticity of digital evidence to make sure the admissibility of the collected evidence before the court of law. Thus, guaranteeing the authenticity and legality of processes and procedures used to gather and transfer the evidence in a digital society is a real challenge if the processes are recorded in papers as currently doing. Cybercrime case investigation includes processes starting from the time of reporting the case to the preservation of evidence after presenting the case before the court. These processes occur at different places and the crime analysis itself can be done by various agencies and by LEAs of different jurisdictions. The lab at which the case will be analysed varies with respect to the nature of the reported crime and the expertise available in the labs. Unfortunately, one of the issues faced by cybercrime investigators is the lack of strong linkage between the acquisition team and the analysis teams since the former happens at the scene of crime and the later happens at the analysis labs based on the suggestions from the court.

IV. RELATED WORK

This section will highlight the process involved, algorithm, and remarks on existing blockchain technology solutions for securing cyber forensics chain of custody process as in Table I.

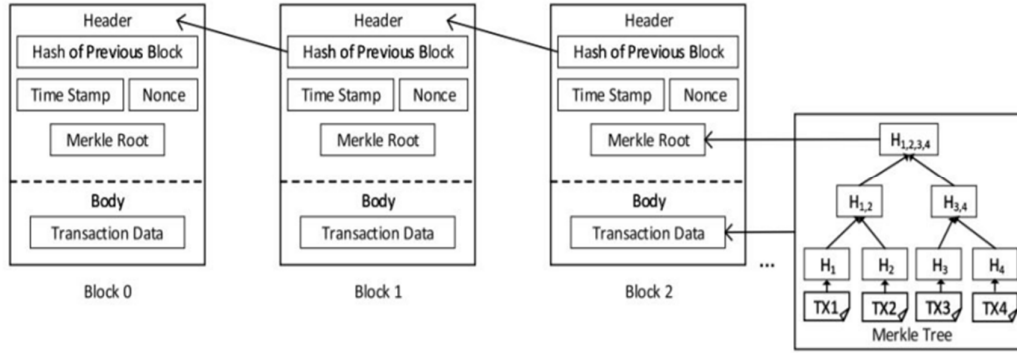


Figure 1. Structure of Block

TABLE I. PAPERS OF RELATED WORK

No.	Processes Involved	Algorithm	Remarks	Ref.
1	Presented Blockchain - Chain of Custody (B-CoC), an evidence log has designed as the private network on Geth using Ethereum nodes. Validator nodes are fixed and also available to perform consensus endorsement policy. Proof of Authority (PoA) protocol adapted on Geth for ease of complexity and deployment. Smart contract has implement using solidity programming language.	Istanbul Byzantine Fault Tolerant (IBFT) consensus algorithm.	Cannot handle multiple stakeholders.	[2]
2	Proposed blockchain based three layered digital forensic framework for IoT environment to solve centralization of existing forensic investigation. Participants are device user, investigator, device manufacturer and service provider. Ethereum private network platform used as blockchain layer. IoT devices communicates each other in an environment, generate block for every communication exchange, when crime occurs participants can verify with their own shared ledger.	Proof of work permitting partial branching.	Gas consumption (block transaction charge) and time complexity.	[6]
3	Presented the Ethereum architecture, investigator collect and generate hash of the evidence. Store the hash of evidence, location, time and date of crime scene. Transfer of stored information among other participants based on the values of destination address, state of evidence, permission level, date and time which are in the smart contract.	Proof of consensus algorithm.	No authenticity of participants since, public blockchain.	[1]
4	Smart lock is created for validated transaction using hash of investigator's node IP address, evidence name, evidence content using Keccak-256 hashing algorithm. A private Ethereum fork created over google cloud among nodes of authorized organization who wish to be part of chain of custody. Nodes communicate using Remote Procedure Calls through web3.js.	Raft consensus algorithm.	Cannot handle multiple investigators.	[8]
5	Ethereum blockchain network among two nodes of authorized law enforcement agencies with Geth tool. Smart contract compiled using Remix IDE and solidity programming language, then deployed on nodes in the private network. Administrator, owner, creator and investigator are participants of the chain of custody network, who transact the case details as per the access coded in the smart contract among different law enforcement agencies.	Proof of stake algorithm.	No guaranteed authenticity, auditing platform.	[14]
6	Law Enforcement Chain scheme has implemented on Ethereum public test network for its feasibility and efficiency. Geth mini server, MetaMask on browser to communicate with Geth via Http, four virtual machines on vmware, file folder "node", puppet to generate the genesis block, smart contract is executable code written on Remix which executes automatically on every node in network according to triggering transactions like store, access, update by investigators.	Proof of authority consensus algorithm.	Not fully support information auditing and transparency.	[13]
7	Proposed blockchain-based multimedia forensics investigation using Sawtooth. This system facilitates trusted organisation participants to do transactions like entree, add, create through blockchain Decentralized Application.	Proof of Elapsed Time (PoET).	Less transaction speed with central lmdb database.	[16]

V. ARCHITECTURE

A high-level architecture of a system is depicted in Figure 2. The initial network configuration and provisioning of necessary credentials access to the users who govern the system are done by the blockchain operator. All the peers from the law enforcement agencies are nothing but docker containers of image hyperledger/fabric-peer,

orderer, and ca. As described, each application module is pluggable, so ExpressJS server with JavaScript language used for smart contract and backend code to run representational state transfer application user interface. The framework Angular 11 is used to design the user interface. The communication is performed using JSON token with representational state transfer calls. For the given scenario single channel named 'organisationChannel' and two law enforcement agencies are sufficient. The third LEA can also be successfully added in between when the network is running and gets joined with the same channel. Level and Couch are two databases supported by the fabric. The Couch is more resilient than Level and also provisions indexes and images, disparate Level database. Hence, the couch is applicable to store all details of case.

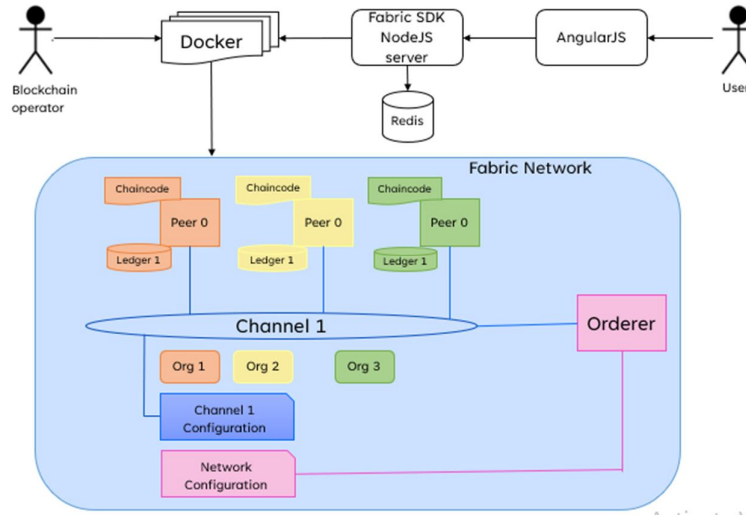


Figure 2. Architecture

Couch database supports here properly since simple use cases are present here. The world state and transaction log are combined into the ledger. The couch includes a world state. Transaction log stores complete communications starting from the block named genesis. It can be found on local system at `/var/lib/docker/volumes/net\_peer0.org2.example.com/_data/ledgersData` (depends upon the installation path). In Couchdb docker containers it is by default available at `/var/hyperledger/production/ledgersData /chains/chains/mychannel/`. It is configurable in `docker-compose-organisation-net.yaml`. The Redis key-value DB also stores officers' credentials - username and password for forensic analysts' credentials. Whereas other details of the forensic analysts are stored as user attributes using fabric SDK. The patient record has some personal and investigation fields described here in the `initLedger` file. This is the initial data when the network gets up. Similarly, two forensic analysts from every two organizations are created.

VI. METHODOLOGY

The HLF comprises many components, has a complex architecture. And these components can be applied and used in many various ways. The following describes how HLF components are applied.

A. Smart Contract and Chaincode

A smart contract in hyperledger fabric defines the transaction logic that controls the lifecycle of a business object contained in the world state. Multiple smart contracts are packaged into a single chaincode which is then deployed to a blockchain network. The rules between different organizations are defined in executable code by the smart contract. Smart contract invoked by Fabric SDK to generate communications and then executes ledger alteration. All smart contracts defined in chaincode are made available to applications, by deploying it. Figure 3 shows the application includes primarily three contracts wrapped into a single chaincode - each role (Investigator/Forensic analyst/Admin) invokes its very own smart contract.

- **AdminContract** – The admin's contract invokes when a user selects the admin role and logs in. The contract includes defined methods that facilitate admin for the creation of the other two roles and an asset case to the ledger. Admin can also view all the cases presented throughout the system.

- InvestigatorContract - The case interacts with the ledger by invoking this contract. The contract contains the logic that is required for the investigator. For instance, only the investigator can update/view the personal details and password via the methods defined in the contract. And also, the investigator contract contains methods to grant and revoke access to a forensic analyst.
- ForensicAnalystContract – Methods defined in the contract of forensic analyst provide permission to update/view the case investigation particulars.

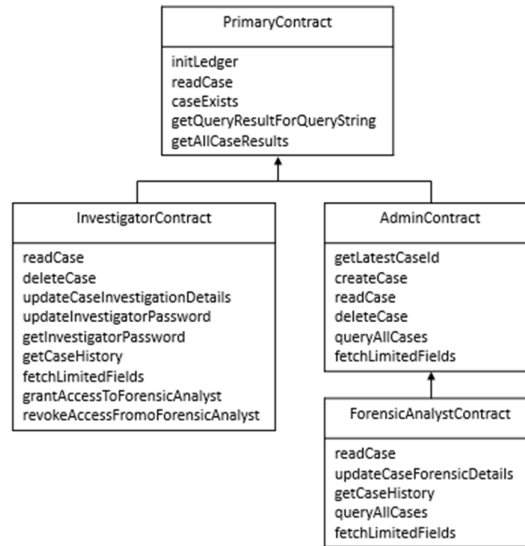


Figure 3. Smart Contracts Hierarchy

The ideology for the three smart contracts is to make sure that the right role has the right access to the data on the ledger. Only the investigator has the right to update the case details, grant/revoke access, and update the password. These methods can be invoked by neither the admin nor the forensic analyst. And similarly, the methods of the forensic analyst cannot be accessed by the investigator or admin. The read case method is common for all, but the data which is retrieved by the contract is different for each role admin just has the access to case name, the forensic analyst has the access to the case investigation details and the investigator has the access to the whole object.

B. Endorsement Policies

The chaincodes in Hyperledger fabric have an endorsement policy that specifies the channel nodes that execute the chaincode methods and approves the outcomes to the ledger and make the communication valid. The policies define the peers which will validate and agree/reject the execution of communication. During this process, the nodes validate the transaction, and the committing node ensures the transaction contains the essential number of confirmations which is configured in the policy.

The Endorsement Policy is defined in the configtx.yaml. Each organization contains its very own endorsing peers as specified in the YAML file in the path &org1/Policies/Endorsement similarly to org2. The rule specifies the roles who are endorsers. In the YAML file, all the peers of the organizations are endorsers. A Chaincode-level endorsement policy is defined in YAML file in the PATH Application: &ApplicationDefaults/Policies/Endorsement 'MAJORITY' specifies that only when a majority of channel members approve a chaincode definition then the definition committed to the channel.

C. Endorsing Members of Organizations

In path &org1/Policies/Endorsement the rule OR ('org1MSP.peer') requests one signature from the peers of org1. Other values can be:

- Requests one signature from each role
AND ('org1MSP.peer', 'org1MSP.admin')
- Requests one signature from either one of the roles
OR ('org1MSP.peer', 'org1MSP.admin')

D. Specifying the endorsing policy

In path &ApplicationDefaults/Policies/Endorsement other types can be a signature with the one of the following rules:

- A transaction is committed only if admin of the both the organization approve.
AND ('org1MSP.admin', 'org2MSP.admin')
- A transaction is committed only if either one of the admin of the both the organization approve.
OR ('org1MSP.admin', 'org2MSP.admin')
- OutOf(1, 'org1MSP.admin', 'org2MSP.admin') which is equivalent to
OR ('org1MSP.admin', 'org2MSP.admin')

E. Ledger

A Distributed ledger in Hyperledger Fabric- one that is logically singular, but has many consistent copies distributed throughout a network comprises of two parts - the world state which holds the current values of the objects, the other part is the blockchain which is the history of the transactions which resulted to the current world state. In the application, the ideology of a ledger can be thought of as one logical database in a Hyperledger Fabric network. In reality, the network contains multiple copies of a ledger for each organization's peers – which are kept consistent with every other copy through consensus. The ledger in the application is being utilized in the following ways

- Firstly, the world state is the database where all the case data is stored. All cases that are queried and updated are stored in the world state. All the query transactions are retrieved from the world state.
- Secondly, the transaction log which is a sequential structured interlinked blocks of log files, where each block contains a sequence of transactions, and a transaction represents a query or an update to the world state. The getHistory API specifically uses the transaction log to retrieve the history of an asset.

F. Identity

For a participant of a law enforcement agency, the participants first need to prove that can be are trusted to make transactions in the blockchain network. An identity for each participant issued by a trusted authority is required to be recognized in this network. The MSPs are the trusted authorities - which can be found in configtx.yaml. An Organization Certificate Authority (CA) dispenses certificates to each of its participants. These certificates are digitally signed by the CA and bind together the participant with the participant's public key with a set of permissions. As a result, if one trusts the CA (and knows its public key), it can trust that the specific participant bound to the public key included in the certificate by validating the CA's signature on the participant's certificate. In the application either CAs(default) or cryptogen can be used to issue identities in the network:

- Cryptogen: Cryptogen is a tool that creates certificates and keys. This tool can be used during development and testing. The tool quickly just creates the required crypto material for the law enforcement agencies. The configuration files for all the LEAs and the orderer are defined in cryptogen directory, the YAML files contain the necessary information for the tool to create the crypto material for the LEAs and the Orderer.
- CAs: Each law enforcement agency has its very own CA server that creates the identities using client of the server to prove that the identity belongs to their LEA and can be trusted. All of the identities created by a CA run of the LEA share the same trust of the root CA. The application uses CAs to allows registration and enrolling investigators/forensic analysts with the Fabric SDK which cannot be done using Cryptogen. The configuration files of the CAs of the LEAs and orderer defined in org1CA, org2CA and OrdererCA respectively.

G. Membership Service Provider (MSP)

Hyperledger Fabric is a permissioned blockchain network, all the participants of the network, in our application the peers/forensic analysts/investigators need to prove their identity to join and perform transactions in the blockchain network to prove that these participants can be trusted. Hyperledger Fabric uses a public key infrastructure (PKI) to verify individualities. The fabric uses Certificate Authorities (CA) to prove identity, the CAs generate both encryption keys for each identity which formulae a key-pair using which the participants can prove their identity and be recognized in the network. The MSP is the one that verifies the private keys of the participants by matching them with the stored public key. For instance, in the application the peers of the hospital endorse a transaction using its private key, the membership service provider on ordering service is where public key stored which is used to validate the private key devoted valid transaction. CAs are used to create trusted identities that are recognized by the network (peers/forensic analysts/investigators). Membership

service providers are required for defining LEAs which are reliable by the network participants. Membership service providers allocated the participants of the LEAs by conventional roles and authorizations in a network. In applications, all the LEAs that join the network can perform read/write transactions in the network. These permissions are configured in configtx.YAML and for participants of these LEAs to perform a transaction in the network, CA should be issued an identity that is recognized in the fabric network. And be a participant of any of the LEAs - MSP is how the participants are linked as a member of the LEA, this is when the public key of the participant is stored in the LEA MSP.

VII. CONCLUSIONS

Hyperledger fabric is a promising blockchain framework that covers the ideas of policies, smart contracts, and the provision of secure identities to make the records controlled. The proposed methodology qualifies interoperability among various law enforcement agencies. Forensic analysts can track the history of a case easily. Investigators do not need to carry case history files and will be significant improvements in digital records. A forensics chain of custody scenario belongs to a private and closed blockchain and this result can magnificently accomplish that it is an inspiring framework. It delivers a consistent and protected solution in handling criminal field records.

REFERENCES

- [1] H. Lone and R. N. Mir, "Forensic-chain: Ethereum blockchain-based digital forensics chain of custody," *Scientific and Practical Cyber Security Journal*, vol. 1, No. 2, pp. 21-27, 2017.
- [2] S. Bonomi, M. Casini, and C. Ciccotelli, "B-CoC: A blockchain-based chain of custody for evidence management in digital forensics," arXiv:1807.10359, 2018.
- [3] A. Dey, S. Bhattacharya, and N. Chaki, "Software watermarking: Progress and challenges," *INAE Letters*, vol. 4, no. 1, pp. 65-75, 2019.
- [4] T. Rochmadi and D. Heksaputra, "Forensic analysis in a cloud storage with live forensics in windows (drive case study)," *International Journal of Cyber-Security and Digital Forensics*, vol. 8, pp. 292-297, Sep. 2019.
- [5] Y. Xiong and J. Du, "Electronic evidence preservation model based on blockchain," in *Proc. 3rd International Conference on Cryptography, Security and Privacy (ICCSP)*, Jan. 2019, pp. 1-5.
- [6] Ryu, Jung Hyun, Pradip Kumar Sharma, Jeong Hoon Jo, and Jong Hyuk Park, "A blockchain-based decentralized efficient investigation framework for IoT digital forensics," *The Journal of Supercomputing* 75, no. 8, 2019.
- [7] T. K. Dasaklis, F. Casino, and C. Patsakis, "SoK: Blockchain solutions for forensics," 2020, arXiv:2005.12640. [Online]. Available: <https://arxiv.org/abs/2005.12640>.
- [8] Ahmad, Liza, Salam Khanji, Farkhund Iqbal, and Faouzi Kamoun. "Blockchain-based chain of custody: towards real-time tamper-proof evidence management." In *Proceedings of the 15th International Conference on Availability, Reliability, and Security*, pp. 1-8. 2020.
- [9] L. Ahmad, S. Khanji, F. Iqbal, and F. Kamoun, "Blockchain-based chain of custody: Towards real-time tamper-proof evidence management," In *Proceedings International Conference on Availability, Reliability and Security*, pp. 1-8, Aug. 2020.
- [10] S. Bourouis, R. Alroobaea, A. M. Alharbi, M. Andejany, and S. Rubaiee, "Recent advances in digital multimedia tampering detection for forensics analysis," *Symmetry*, vol. 12, no. 11, p. 1811, Nov. 2020.
- [11] D. Barrett, "Cloud based evidence acquisitions in digital forensic education," *Information Systems Education Journal*, vol. 18, no. 6, pp. 46-56, 2020.
- [12] D. Li, W. Liu, L. Deng, and B. Qin, "Design of multimedia blockchain privacy protection system based on distributed trusted communication," *Transactions on Emerging Telecommunications Technologies*, vol. 32, no. 2, p. e3938, Feb. 2021.
- [13] M. Li, C. Lal, M. Conti, and D. Hu, "LEChain: A blockchain-based lawful evidence management scheme for digital forensics," *Future Generation Computer Systems*, vol. 115, pp. 406-420, 2021.
- [14] Fu-Ching Tsai, "The Application of Blockchain of Custody in Criminal Investigation Process," *25th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems*, 2021.
- [15] G Jagadamba, R Sheeba, K N Brinda, K C Rohini, and S K Pratik, "Adaptive E-Learning Authentication and Monitoring," *2nd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA)*, pp. 277-283, 2020.
- [16] Abdullah Ayub Khan, Mueen Uddin, Aftab Ahmed Shaikh, Asif Ali Laghari, and Adil E. Rajput, "MF-Ledger: Blockchain Hyperledger Sawtooth-Enabled Novel and Secure Multimedia Chain of Custody Forensic Investigation Architecture," *IEEE Access*, vol. 9, pp. 103637-103650, 2021.