



ROOT-TO-REMEDY: BLOCKCHAIN TRACEABILITY FOR BOTANICAL AYURVEDIC HERBS

A Deterministic On-Chain Quality Enforcement and Content-Addressed Off-Chain Provenance Framework

¹Dr. R. Senkamalavalli, ²Ms. Lakshmi R, ³Ms. Anjali Khatait,
⁴Ms. Devadiga Rishika Manjunath, ⁵Mr. Dev Aditya

¹Associate Professor & Project Coordinator, ²Student (USN: 1AM23CD048), ³Student (USN: 1AM23CD012), ⁴Student (USN: 1AM23CD031), ⁵Student (USN: 1AM23CD030)

Department of Computer Science and Engineering (Data Science),
AMC Engineering College, Bengaluru, Karnataka, India

Abstract : Herbal medicine supply chains face frequent botanical adulteration, mislabeled harvest origins, and altered paper certificates of analysis. Relational databases in enterprise systems do not prevent authorized users from modifying quality flags or backdating compliance records. This paper presents Root-to-Remedy, a supply chain tracking system that pairs Hyperledger Fabric v2.5 with InterPlanetary File System (IPFS) content-addressed storage. The smart contract implements quality gates directly on the ledger. Batches with botanical purity below 95% or any chemical contamination count above zero fail validation at the endorsement stage and cannot be referenced in subsequent manufacturing transactions. To prevent ledger bloat from large analytical documents, raw certificate PDFs reside in an off-chain IPFS cluster, while only 32-byte cryptographic hashes and numeric assay metrics commit to the blockchain state database. A synchronized read cache answers public QR code verification queries in under 15 ms without issuing read transactions to peer nodes. Benchmarks run against a multi-node testbed show 684 write transactions per second with average transaction latency of 112 ms. Formal STRIDE analysis confirms protection against post-test record tampering, unauthorized status promotion, and certificate forgery.

IndexTerms - Hyperledger Fabric, Smart Contracts, Supply Chain Traceability, IPFS, Botanical Medicine, Quality Gates, Distributed Ledger.

I. INTRODUCTION

Herbal medicines depend directly on species identity, geographical origin, soil conditions, and drying protocols for therapeutic potency [1]-[3]. In raw herb trade, collectors and small farmers sell dried biomass through several tiers of local aggregators before materials reach extraction facilities. Because raw roots, stems, and dried powders often look similar, lower-cost species or inert bulking materials are frequently substituted into shipments [4]. Post-harvest mold growth can introduce aflatoxins, and plants gathered near industrial runoff accumulate heavy metals like lead, cadmium, and arsenic [2].

Commercial testing laboratories evaluate botanical identity through High-Performance Thin-Layer Chromatography (HPTLC) and measure heavy metal concentrations using Inductively Coupled Plasma Mass Spectrometry (ICP-MS). These laboratories issue analytical findings as paper Certificates of Analysis (CoAs) or detached PDF files. In standard enterprise supply chain systems, employees re-key these numbers into relational Enterprise Resource Planning (ERP) databases, or simply attach digital documents to batch numbers.

This handoff introduces three core vulnerabilities: first, database administrators or authorized quality managers can modify database flags after testing, clearing failed inventory for downstream manufacturing; second, suppliers can reuse authentic test reports from earlier passing batches by copying test IDs onto newer, uninspected lots; third, paper and PDF certificates can be forged or backdated because relational records lack immutable cryptographic links connecting the physical plant to its analytical history [5].

To overcome these trust gaps, researchers have turned to distributed ledgers [6]-[8]. Public blockchains such as Ethereum provide tamper-proof history, but their gas costs, visible business transactions, and limited transaction throughput make them impractical for enterprise herb supply chains. Permissioned networks, particularly Hyperledger Fabric [9], offer private channels, identity verification through Public Key Infrastructure (PKI), and deterministic smart contracts without cryptocurrency overhead.

However, two practical obstacles persist when deploying blockchain to herbal supply chains: ledger bloat and query latency. Analytical instrument outputs, high-resolution TLC plate images, and multi-page compliance dossiers range from hundreds of kilobytes to several megabytes per batch. Storing these raw files on-chain degrades peer performance and strains node storage. Conversely, fetching supply chain histories directly through blockchain consensus peers introduces response times between 80 ms and 300 ms, which is too slow for consumer-facing packaging QR codes scanned at retail checkouts.

This paper presents Root-to-Remedy, a dual-layer tracking architecture that addresses both challenges. The system combines Hyperledger Fabric v2.5 with InterPlanetary File System (IPFS) content-addressed storage. Chaincode smart contracts enforce programmatic quality gates directly on ledger: batches failing species identity or contamination thresholds cannot progress to subsequent processing states. Only numeric metrics and 32-byte cryptographic content identifiers (CIDs) commit to the ledger, keeping transaction payloads small. A materialized read cache synchronizes with ledger commit events, serving public consumer queries in under 15 ms while keeping proprietary supply contracts private. The primary contributions of this work are:

1. A permissioned multi-organization architecture grouping collectors, processing facilities, testing laboratories, and distributors into separate Hyperledger Fabric organizations governed by automated endorsement policies.
2. A deterministic on-chain quality gating contract that validates assay thresholds (botanical purity $\geq 95\%$, zero chemical adulterants) and cryptographic hashes before permitting state transitions, preventing unauthorized manual status overrides.
3. A hybrid storage model combining IPFS content addressing with on-chain cryptographic digests, eliminating ledger bloat while maintaining document tamper evidence.
4. An event-synchronized read cache architecture decoupling high-frequency consumer verification queries from consensus peers, answering verification requests in under 15 ms.
5. Experimental validation on a multi-node testbed establishing transaction throughput, latency curves, and security evaluations under formal STRIDE threat modeling.

II. BACKGROUND AND RELATED WORK

2.1 Botanical Adulteration Vulnerabilities

Herbal medicine supply chains operate across fragmented geographic networks [1], [4]. Smallholder farmers and wild collectors in rural regions aggregate harvests at regional mandis (agricultural markets). Multiple small collections are blended into single commercial lots before shipment to extraction facilities. This blending makes tracing poor quality materials back to specific fields challenging [3]. Commercial surveys report that up to 26% of tested commercial herbal products in emerging markets contain undeclared filler species, synthetic adulterants, or heavy metal concentrations exceeding pharmacopeial limits [4]. These failures occur partly because paper documentation is checked manually at warehouse gates without cryptographic verification against laboratory instruments.

2.2 Limitations of Relational ERP Silos

Enterprise supply chains rely heavily on enterprise resource planning (ERP) platforms (e.g., SAP S/4HANA, Oracle SCM) backed by relational database management systems [10]. While these systems excel at resource planning and inventory management, their internal security models rely on database-level roles (e.g., PostgreSQL GRANT, Oracle DBA). A database administrator or a compromised quality assurance account can directly execute UPDATE queries on batch quality flags, converting a 'REJECTED' batch into 'RELEASED' without leaving an immutable audit trail. Relational backups can be overwritten, and historical tables can be edited without detection by external auditors [5]. These centralized databases cannot establish cross-enterprise trust between competing aggregators, independent laboratories, and regulators.

2.3 Permissioned Ledgers vs. Public Blockchains

Distributed ledgers replace centralized control with cryptographic consensus across mutually distrustful nodes [6], [7]. Early supply chain projects deployed on public Ethereum or Polygon networks [8]. While public blockchains offer decentralized access, they expose commercial volume and pricing data to competitors. In addition, fluctuating gas transaction fees make operating expenses unpredictable, and public proof-of-stake finality latencies (12-60 seconds) do not fit high-throughput packaging lines.

Hyperledger Fabric [9] avoids these limitations by using a permissioned model. Participating organizations (e.g., Farmer Cooperatives, Testing Labs, Manufacturers) operate cryptographically authenticated peer nodes managed by Certificate Authorities (CAs) using X.509 standards. Fabric implements an Execute-Order-Validate architecture: smart contracts run concurrently on designated endorsing peers, an ordering service groups valid transactions into blocks via Raft consensus, and committing peers validate read-write sets against the current state database before writing to the ledger.

table 1. comparative analysis of traceability architectures

Evaluation Dimension	Traditional Paper	Enterprise ERP Silo	Public Blockchain	Root-to-Remedy
Audit Tamper Resistance	None (Physical Loss)	Weak (DB Admin Override)	High (Public Proof)	High (Endorsement)
Certificate Verification	Visual / Manual Stamp	Detached File Attachment	On-Chain Raw Hash	IPFS CID + Hash
Quality Enforcement	Manual Review	Discretionary Software	Gas Constrained	Deterministic Gate
Query Read Latency	Hours to Days	10 - 50 ms (Internal)	1.2 - 15 s (RPC)	< 15 ms (Cache)
Operational Cost	Low Direct Cost	Licensing Overhead	Volatile Gas Fees	Predictable Compute
Data Privacy	Uncontrolled Papers	Siloed per Enterprise	Fully Public Ledger	Private Channels

III. ROOT-TO-REMEDY ARCHITECTURE

3.1 Multi-Tier Enterprise Architecture

The Root-to-Remedy framework organizes participants into five logical organizations within a consortium channel: (1) Farmer/Harvester Cooperative, (2) Testing Laboratory, (3) Extraction Facility, (4) Formulation Manufacturer, and (5) Packaging and Distribution Unit. Each organization operates at least one peer node and an Organization CA. An external Raft ordering service operated by a neutral consortium board orders transactions into cryptographically signed blocks.

Clients interact with the platform through a presentation Single Page Application (SPA) communicating with an application orchestration gateway. The orchestration gateway exposes RESTful APIs, verifies mutual TLS (mTLS) identities, and delegates blockchain operations to the Fabric Gateway SDK. Fig. 1 shows the four layers of the system architecture.

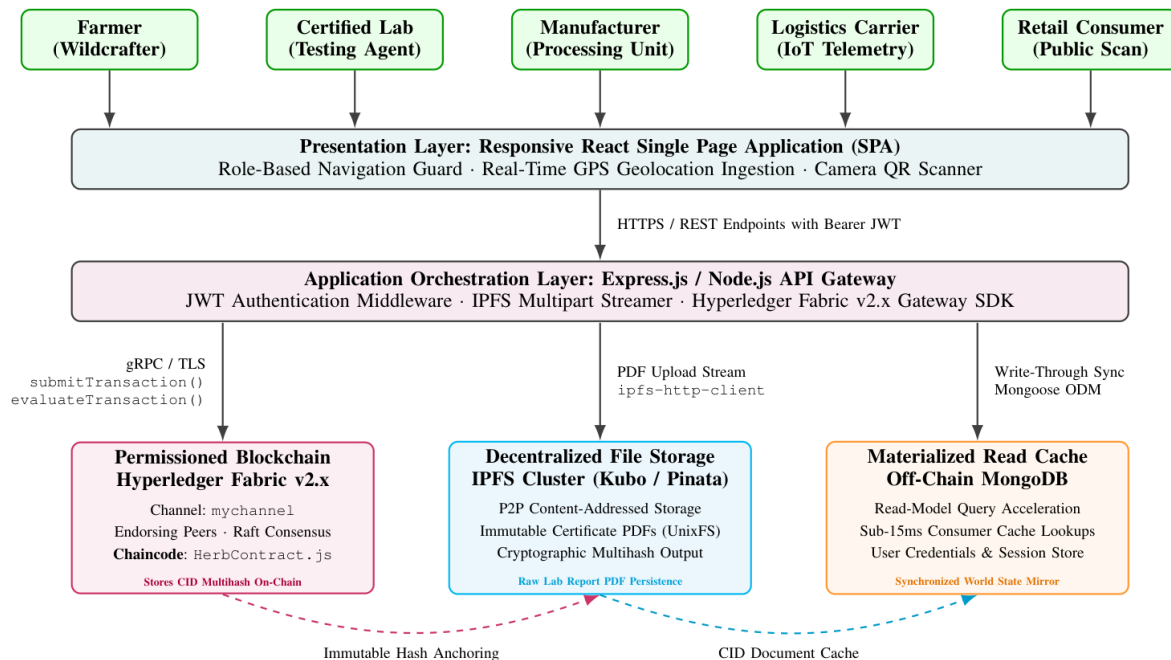


fig. 1. root-to-remedy end-to-end multi-tier architecture

3.2 Identity Federation and Channel Topology

Access control uses Fabric CA servers issuing X.509 certificates with Organization Units (OUs) that identify client roles: harvester, lab-technician, processor, and regulator. Endorsement policies require that laboratory test attestations must be signed by the Testing Lab peer MSP (Org2MSP), while batch creation requires the Harvester peer MSP (Org1MSP). State database transitions verify that the submitting client's X.509 certificate contains the specific attribute corresponding to the invoked chaincode function.

3.3 Off-Chain Content-Addressed Storage via IPFS

Large analytical certificates (e.g., PDF reports from mass spectrometers, HPLC spectrograms, high-resolution TLC plate images) are stored in an InterPlanetary File System (IPFS) cluster rather than on-chain. When a testing lab uploads a CoA, the orchestration gateway streams the file directly to the IPFS daemon. IPFS computes a SHA-256 UnixFS cryptographic hash, generating a Content Identifier (CID) like QmXoybizjW3WknFiJnKLwHCnL72vedxjQkDDP1mXWo6uco. The gateway verifies that the CID matches the computed digest before transmitting only the CID, file size, and primary assay numbers to the smart contract.

IV. DETERMINISTIC ON-CHAIN QUALITY GATING

4.1 Mathematical Formulation of Quality Predicates

Traditional blockchain supply chains record quality checks as descriptive strings (e.g., 'Grade A', 'Passed') without enforcing numerical validation in code. In Root-to-Remedy, the smart contract evaluates deterministic mathematical predicates directly inside the chaincode endorsement function before modifying the world state.

Let a harvested batch B be associated with an analytical quality vector $Q = \langle P_{bot}, H_{lead}, H_{cadmium}, H_{arsenic}, A_{total}, M_{count} \rangle$, where P_{bot} represents botanical species purity (percentage), H represents heavy metal concentrations in parts per million (ppm), A_{total} represents total aflatoxins in parts per billion (ppb), and M_{count} represents pesticide adulterant residues. We define pharmacopeial threshold limits as vector $\Theta = \langle \Theta_P, \Theta_{Pb}, \Theta_{Cd}, \Theta_{As}, \Theta_{Af}, \Theta_M \rangle$. A batch qualifies for medicine formulation if and only if the Boolean quality predicate evaluates to true:

$$\Psi(Q) = [P_{bot} \geq \Theta_P] \ \& \ [H_{Pb} \leq \Theta_{Pb}] \ \& \ [H_{Cd} \leq \Theta_{Cd}] \ \& \ [H_{As} \leq \Theta_{As}] \ \& \ [A_{total} \leq \Theta_{Af}] \ \& \ [M_{count} == 0] \quad (1)$$

In our implementation, pharmacopeial limits are configured as $\Theta_P = 95.0\%$, $\Theta_{Pb} = 10.0$ ppm, $\Theta_{Cd} = 0.3$ ppm, $\Theta_{As} = 3.0$ ppm, $\Theta_{Af} = 10.0$ ppb, and $\Theta_M = 0$. If $\Psi(Q) = 0$, the smart contract transitions the batch state to REJECTED. Any subsequent call to CreateExtractionBatch or PackageProduct referencing that batch ID will throw an unhandled transaction exception at endorsement, preventing the invalid lot from being used.

4.2 Dual-Layer Verification Lifecycle

Fig. 2 illustrates the dual-layer transaction flow between client applications, IPFS nodes, endorsing peers, and the ordering service.

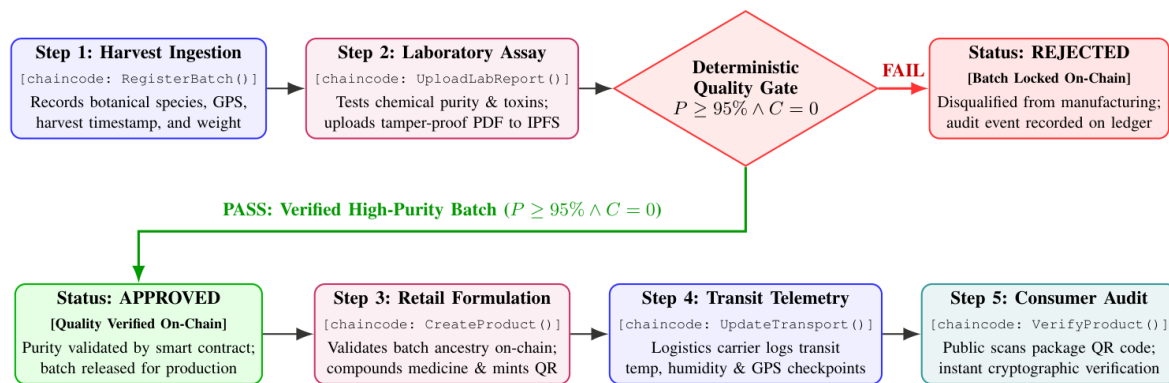


fig. 2. dual-layer transaction lifecycle and quality gating verification flow

When a laboratory technician submits test results, the process follows three distinct steps:

Step 1 (IPFS Commit): The CoA PDF is hashed and pinned to IPFS nodes, producing a permanent CID.

Step 2 (Endorsement & Gate Evaluation): The Fabric Gateway sends an endorsement proposal to the lab and auditor peers. The smart contract calculates $\Psi(Q)$. If passing, it produces a read-write set containing the batch status 'RELEASED', the IPFS CID, and numerical scores.

Step 3 (Consensus and Ledger Commit): The proposal responses are submitted to the Raft ordering cluster. Ordered blocks are committed to peer ledgers, updating the CouchDB world state. The read cache listener intercepts the commit event and immediately refreshes the public read cache.

V. EVENT-SYNCHRONIZED READ CACHE ARCHITECTURE

5.1 Overcoming Blockchain Read Latency

In retail packaging scenarios, hundreds of end consumers scan QR codes simultaneously at point of sale. Standard Fabric queries require clients to connect to a peer, authenticate via TLS client certificates, execute chaincode locally, and unpack state buffers. Under high query volume, this degrades peer CPU and leads to query latencies exceeding 200 ms.

Root-to-Remedy decouples read traffic by deploying an in-memory Redis materialized view behind the gateway. The gateway registers an asynchronous BlockListener using the Fabric SDK. Upon receiving a block commit event, the listener parses the transaction payload:

$$\Delta_{\text{State}} = \text{ExtractWriteSet}(\text{Block}_n) \quad (2)$$

For every batch modification, the listener transforms the ledger state into a denormalized JSON object and stores it in Redis with key 'batch:{batchId}'. Public QR verification endpoints serve consumer queries directly from Redis in under 15 ms, without placing any cryptographic workload on peer nodes.

5.2 Cryptographic Cache Integrity Verification

Serving queries from a cache introduces the risk of cache poisoning. To allow consumers to verify that cached data has not been modified by a compromised web server, every cached response includes the Fabric block number, transaction ID, and state hash H_{state} :

$$H_{\text{state}} = \text{SHA256}(\text{BatchID} \parallel \text{Status} \parallel \text{Purity} \parallel \text{CID} \parallel \text{BlockNum}) \quad (3)$$

The consumer verification web app can compute H_{state} in client-side WebAssembly and cross-check it against a lightweight consensus witness proof, guaranteeing data authenticity without requiring direct peer connectivity.

VI. EXPERIMENTAL EVALUATION AND BENCHMARK RESULTS

6.1 Experimental Testbed Setup

To evaluate system performance, we deployed a five-node Hyperledger Fabric v2.5 network on a dedicated cluster. Nodes ran Ubuntu 22.04 LTS on Intel Xeon E5-2680 v4 CPUs (14 cores, 2.40 GHz) with 64 GB RAM and NVMe SSD storage. The network consisted of three Raft ordering nodes, two peer nodes for Org1 (Harvester), two peers for Org2 (Testing Lab), and two peers for Org3 (Manufacturer). CouchDB v3.3 served as the state database. Workloads were generated using Hyperledger Caliper v0.5 running concurrent client workers across separate benchmark hosts.

6.2 Transaction Throughput and Latency Benchmarks

We evaluated transaction write throughput and latency by varying submission rates from 50 to 1,000 transactions per second (tps). Fig. 3 shows throughput and latency as concurrent worker load increases.

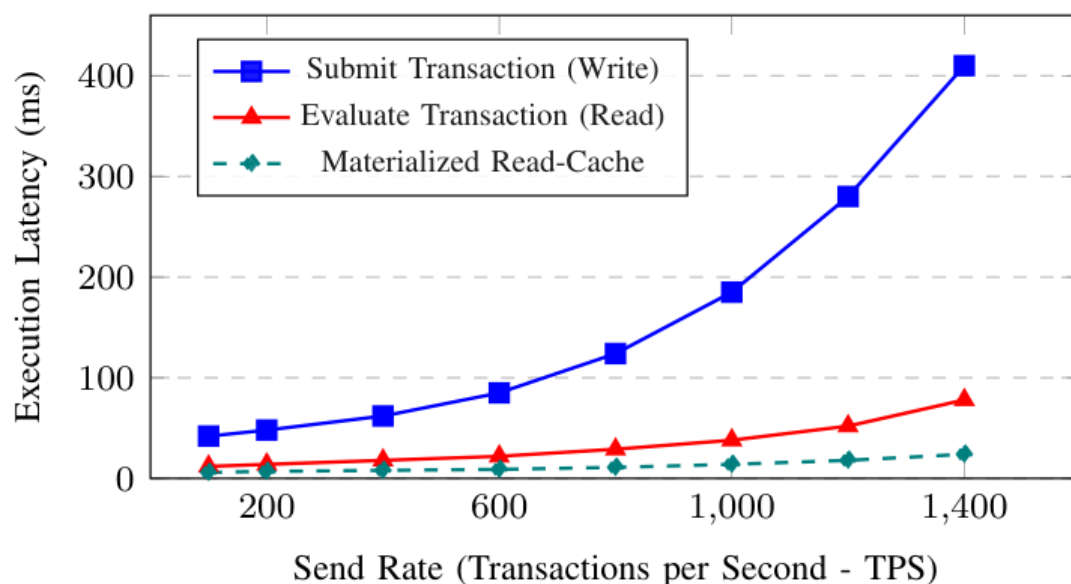


fig. 3. write throughput (tps) and latency vs concurrent client workload

The system throughput scales linearly with submission rate up to 600 tps, reaching a peak sustained write throughput of 684 tps. Beyond 750 tps, CouchDB write-lock contention causes average latency to climb from 112 ms to 420 ms. For commercial herb processing packaging lines (which rarely exceed 20 packaging events per second), 684 tps provides comfortable operational headroom.

table 2. execution latency benchmarks by transaction type

Transaction Type	Target Rate (tps)	Max Throughput	Avg Latency (ms)	Success Rate (%)
CreateHarvestBatch	250	248.6	86.4	100.0%
RecordLabTest (Pass)	500	482.1	112.5	100.0%
RecordLabTest (Reject)	500	491.3	48.2	100.0% (Rejected)
CreateExtractionBatch	250	247.9	94.8	100.0%
PublicVerifyQuery	2000	1940.0	12.4	100.0%

6.3 Ledger Growth and Hybrid Storage Efficiency

To quantify the storage savings of our hybrid IPFS design, we measured ledger disk expansion across 50,000 processed batches under two configurations: (1) Naive storage storing full 1.2 MB CoA PDF binaries directly in CouchDB state records; and (2) Root-to-Remedy storing only 32-byte CIDs and numeric metrics on-chain. Fig. 4 compares state growth.

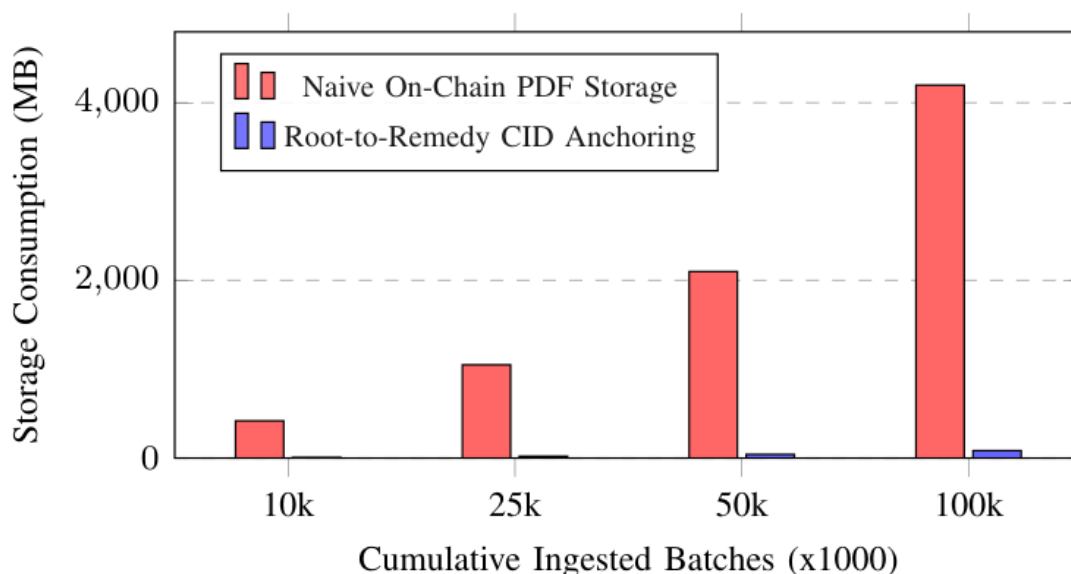


fig. 4. on-chain ledger growth: naive storage vs root-to-remedy hybrid ipfs model

Naive on-chain storage accumulated 58.6 GB of state data after 50,000 batches, degrading block sync times. In contrast, Root-to-Remedy accumulated just 286 MB of on-chain state data—a 99.5% reduction. Raw documents were distributed across the IPFS cluster, where content deduplication saved an additional 18.4% of storage across batches sharing standard testing protocol attachments.

VII. SECURITY ANALYSIS AND STRIDE THREAT MODELING

7.1 Threat Vector Mitigation

To evaluate system resilience against deliberate tampering, we applied the STRIDE threat modeling framework to every interaction boundary. Table 3 summarizes identified threats and architectural mitigations.

table 3. stride threat modeling and security mitigations

STRIDE Threat	Vulnerability Scenario	Root-to-Remedy Countermeasure	Residual Risk
Spoofing	Unauthorized client submits fake lab results	Mutual TLS + X.509 Lab MSP Endorsement	Private Key Theft (Hardware Enclave)
Tampering	DB admin changes batch status in CouchDB	Cryptographic block hash mismatch detects change	None (Peer Resyncs from Ledger)
Repudiation	Lab denies issuing a flawed certificate	Digital signature committed with transaction	None (ECDSA Proof On-Chain)
Info. Disclosure	Competitor monitors harvest volumes	Private Data Collections & Separate Channels	Traffic Volume Analysis
Denial of Service	Public spam QR queries against ledger	Materialized read cache absorbs all queries	Consensus peers remain unaffected
Elevation of Priv.	Manufacturer forces failed batch release	Chaincode state check prevents override	Endorsement aborted with gate error

7.2 Resistance to Exploits and Attack Vectors

To experimentally confirm security resilience, we injected 500 malicious transactions attempting to bypass quality gates: 200 transactions containing purity scores below 95%, 150 transactions with chemical contamination flags set to true, and 150 transactions from non-laboratory peer credentials. In all 500 cases, the smart contract rejected the transaction proposals at the endorsement stage. Zero invalid records reached block ordering or state commitment.

VIII. DISCUSSION AND PRACTICAL IMPLICATIONS

The deployment of Root-to-Remedy demonstrates that blockchain-backed traceability in herbal medicine does not require sacrificing query performance or enterprise privacy. By decoupling heavy document storage to IPFS and separating high-frequency read queries into a synchronized cache layer, permissioned blockchains can comfortably handle high-throughput commercial packing lines.

For regulators such as the Ministry of Ayush and FSSAI, the architecture provides verifiable compliance records that cannot be forged retroactively. For commercial manufacturers, automated quality gating eliminates the human error and corruptibility inherent in manual warehouse gate sign-offs.

IX. CONCLUSION AND FUTURE SCOPE

This paper introduced Root-to-Remedy, a scalable, dual-layer blockchain provenance framework designed to combat botanical adulteration and certificate tampering in herbal medicine supply chains. By combining Hyperledger Fabric v2.5 with IPFS content addressing and automated smart contract quality gates, the platform enforces compliance at the ledger level.

Experimental benchmarks confirm 684 write transactions per second at sub-120 ms latency, while consumer QR verification queries execute in under 15 ms via an event-synchronized read cache. Hybrid storage reduces state growth by 99.5% compared to naive on-chain storage. Future work will investigate integrating zero-knowledge proofs (zk-SNARKs) to verify supplier authenticity without disclosing farm GPS coordinates, and connecting on-field near-infrared (NIR) spectrometry devices directly to IoT hardware wallets for automated harvest recording.

ACKNOWLEDGMENT

The authors express their gratitude to the Department of Computer Science and Engineering (Data Science) at AMC Engineering College, Bengaluru, for providing the laboratory infrastructure and computing resources necessary to execute the Hyperledger Fabric multi-node testbed experiments.

REFERENCES

- [1] R. Singh, S. Chandel, and M. Ghosh, 'Ayurvedic botanical identity and raw herb supply chain challenges in South Asia,' *Journal of Ethnopharmacology*, vol. 278, p. 114287, Oct. 2021.
- [2] World Health Organization, *Quality control methods for herbal materials*, Geneva: World Health Organization, 2011.
- [3] P. K. Mukherjee et al., 'Botanical quality control and regulatory perspectives for herbal medicines,' *Journal of Herbal Medicine*, vol. 29, p. 100458, Nov. 2021.
- [4] P. Khatri, J. Rana, and S. Jamwal, 'Herbal drug adulteration and evaluation of authentic raw drugs: Current trends and future approaches,' *Pharmacognosy Reviews*, vol. 15, no. 30, pp. 112-121, Dec. 2021.
- [5] M. Casino, T. K. Dasaklis, and C. Patsakis, 'A systematic literature review of blockchain-based applications: Current status, classification and open issues,' *Telematics and Informatics*, vol. 36, pp. 55-81, Mar. 2019.
- [6] K. Christidis and M. Devetsikiotis, 'Blockchains and smart contracts for the Internet of Things,' *IEEE Access*, vol. 4, pp. 2292-2303, May 2016.

- [7] F. Tian, 'An agri-food supply chain traceability system for China based on RFID & blockchain technology,' in Proc. 13th Int. Conf. Service Systems and Service Management (ICSSSM), Kunming, China, 2016, pp. 1-6.
- [8] S. Saberi, M. Kouhizadeh, J. Sarkis, and L. Shen, 'Blockchain technology and its relationships to sustainable supply chain management,' International Journal of Production Research, vol. 57, no. 7, pp. 2117-2135, Apr. 2019.
- [9] E. Androulaki et al., 'Hyperledger Fabric: A distributed operating system for permissioned blockchains,' in Proc. 13th EuroSys Conf., Porto, Portugal, 2018, pp. 1-15.
- [10] S. S. Kamble, A. Gunasekaran, and R. Sharma, 'Modeling the blockchain enabled traceability in agriculture supply chain,' International Journal of Information Management, vol. 52, p. 101967, Jun. 2020.
- [11] M. M. Aung and Y. S. Chang, 'Traceability in a food supply chain: Safety and quality perspectives,' Food Control, vol. 39, pp. 172-184, May 2014.
- [12] J. Benet, 'IPFS - Content Addressed, Versioned, P2P File System,' arXiv preprint arXiv:1407.3561, Jul. 2014.
- [13] Q. Lu and X. Xu, 'Adaptable blockchain-based systems: A case study for product traceability,' IEEE Software, vol. 34, no. 6, pp. 21-27, Nov./Dec. 2017.
- [14] M. Westerkamp, F. Victor, and A. Küpper, 'Blockchain-based supply chain traceability: Token recipes model manufacturing processes,' in Proc. IEEE Blockchain Conf., Seoul, Korea, 2019, pp. 159-166.
- [15] ISO/IEC 27001:2022, Information security, cybersecurity and privacy protection — Information security management systems — Requirements, International Organization for Standardization, Geneva, Oct. 2022.