ICCCN, 31 July 2019

Data Management in Supply Chain Using Blockchain: Challenges and A Case Study

Hanqing Wu*, Jiannong Cao*, Yanni Yang*, Cheung Leong Tung*, Shan Jiang*, Bin Tang*, Yang Liu †, Xiaoqing Wang †, Yuming Deng †

*The Hong Kong Polytechnic University, Hong Kong, China

† Alibaba Group Holding Limited, Hang Zhou, China

Table of Contents

Big Data Era

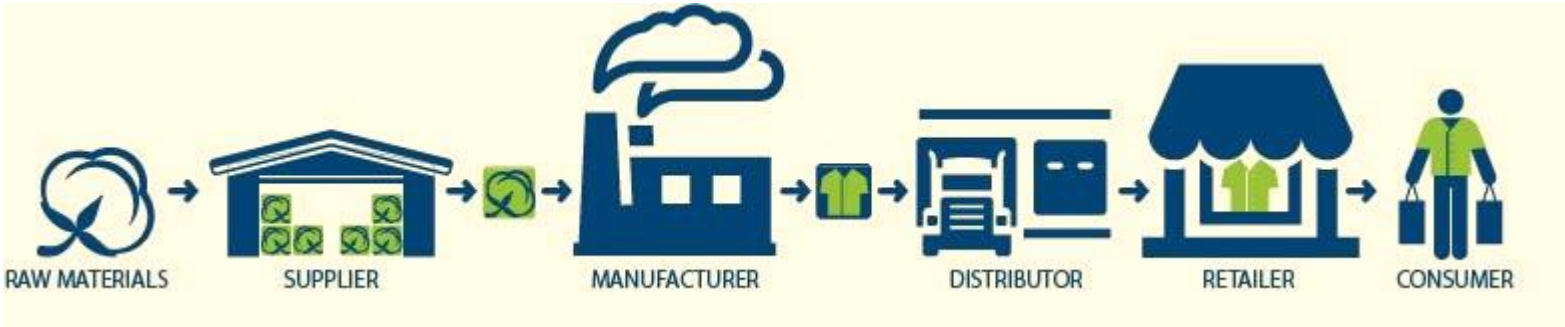
Numerous big data applications benefit human beings

Data Management

- Open data
- Big data trading
- Big data collaboration

What is Supply-Chain Management?

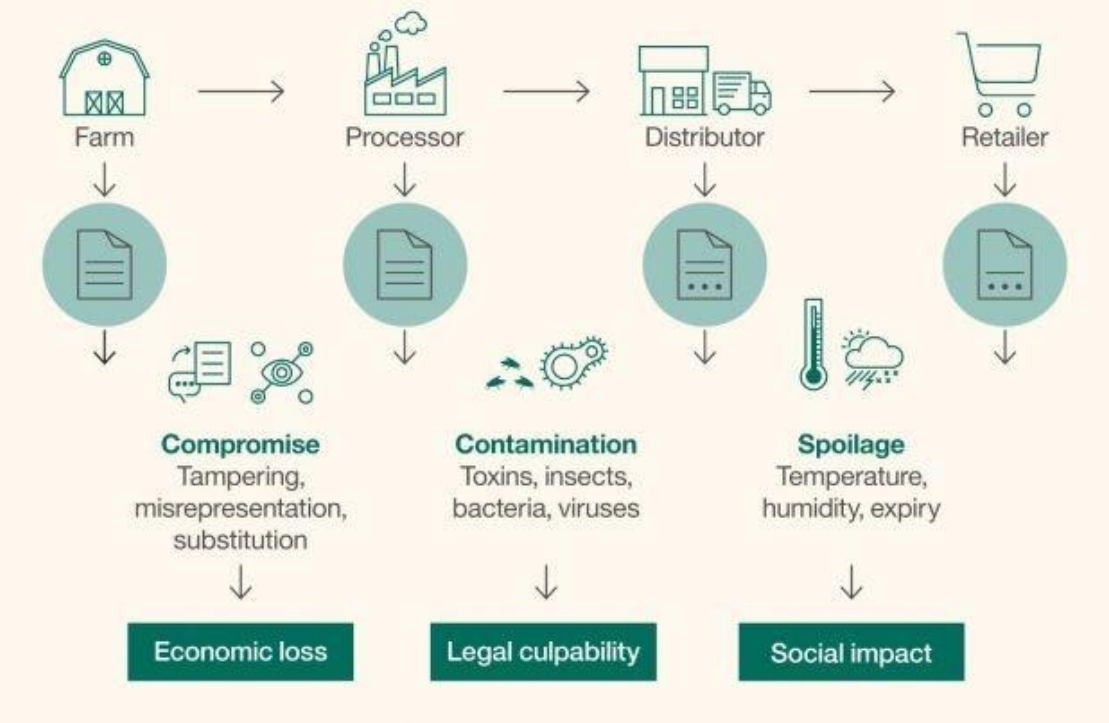
- **Supply Chain (SC)**: a system of organizations, people, activities, information, and resources involved in moving a product or service from supplier to customer.



- **Supply-Chain Management (SCM)**: the management of the flow of goods or services, involves the movement and storage of materials, inventory, and finished goods from origin to consumption.

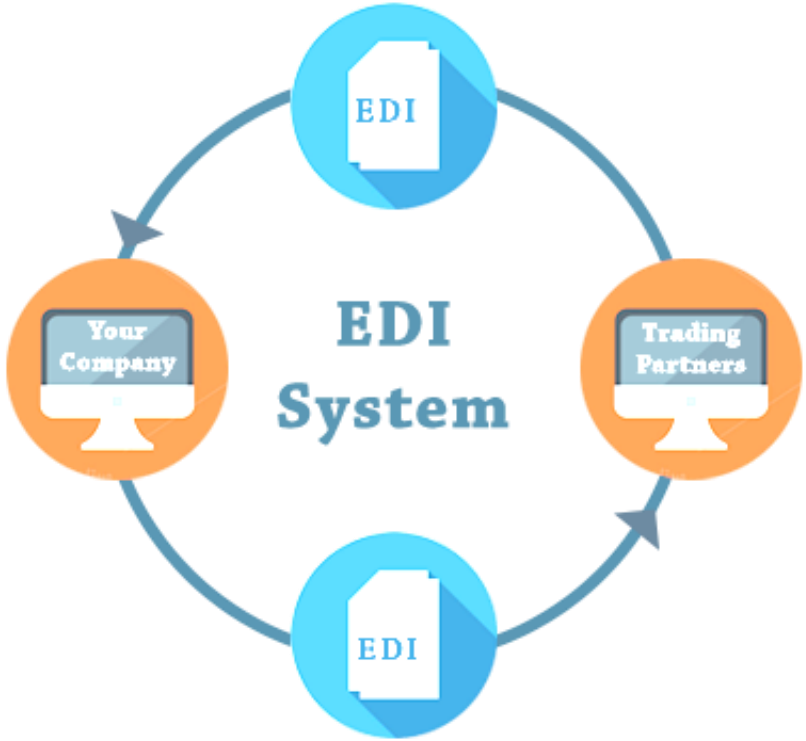
Data Sharing & Management for Supply Chain

Food safety, especially traceability, needs data management and sharing among stakeholders



Problems of Current Solution

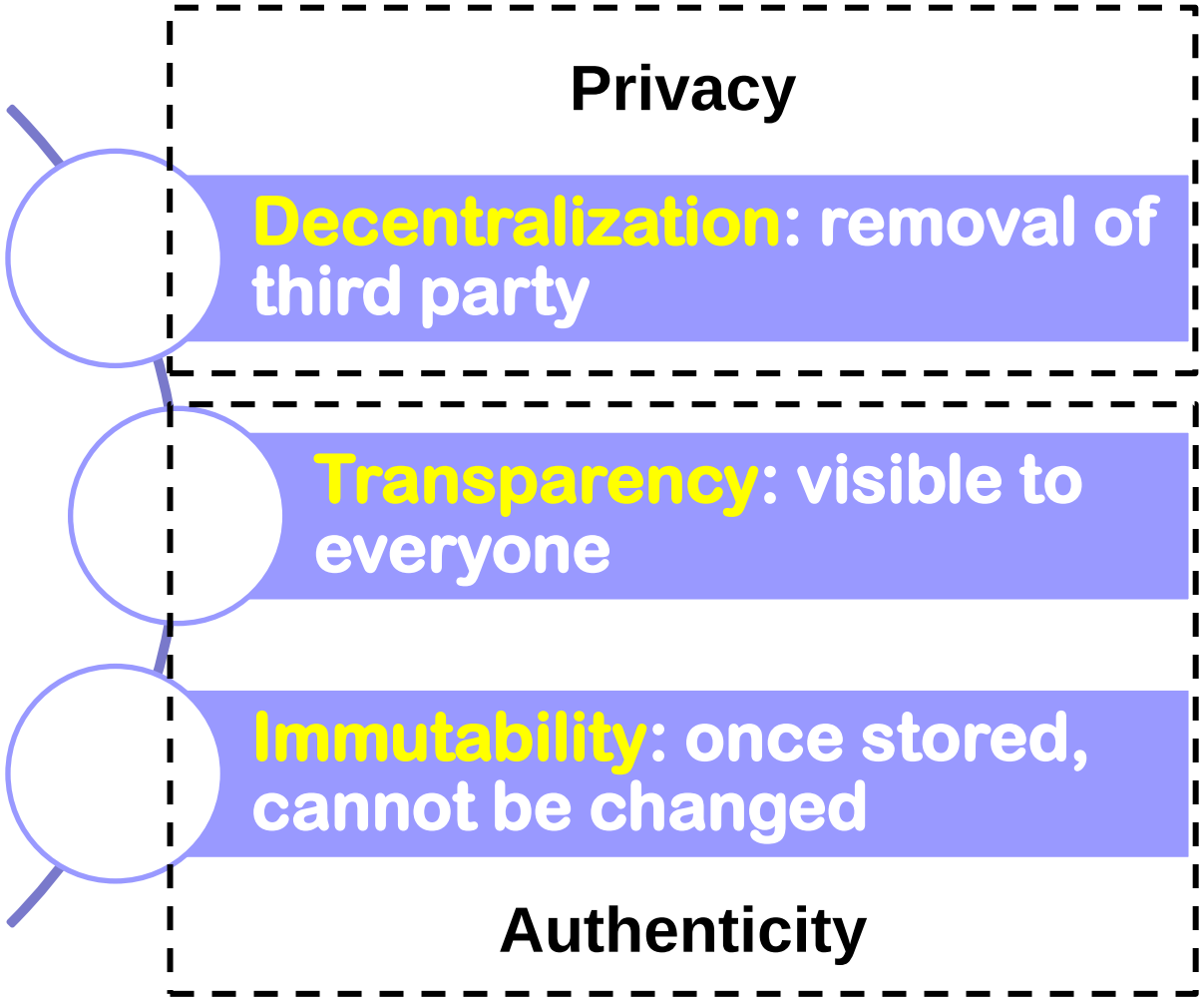
- Scaling EDI: trading partners have their own EDI transaction sets.
- Overcoming bad data: transactions are affected or suspended due to data related anomaly
- Achieving transparency: better visibility of information is critical for SCM.



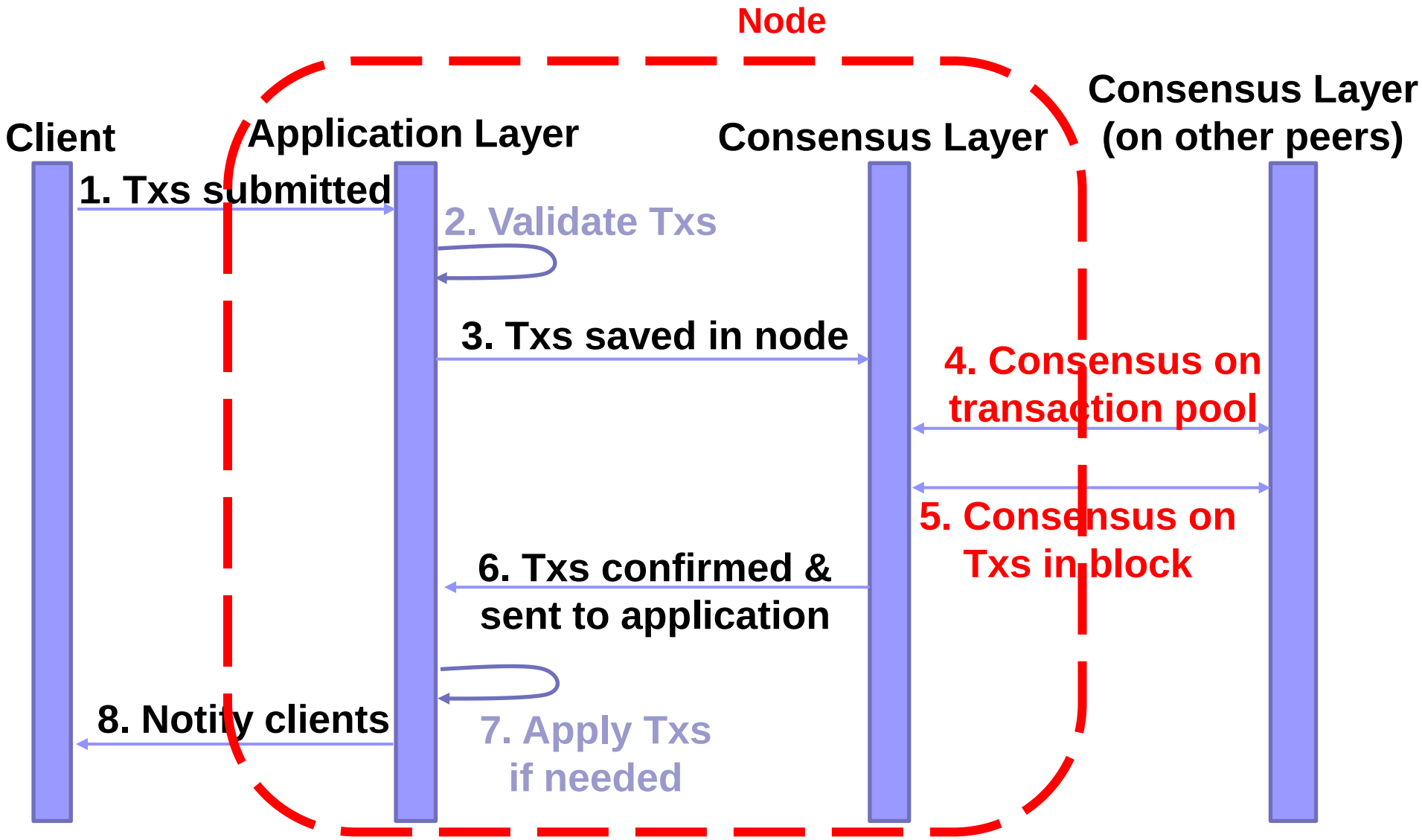
Blockchain

Blockchain

A technology of **distributed ledger** for **trustless** data storage with **auditability**



How Blockchain Works



What Can Blockchain Help?

- Information Management
  - Provide product information transparency
  - Provide immutable and reliable information storage
  - Provide decentralized and scalable information sharing
- Inventory and Asset Management
  - Provide proactive inventory management
  - Provide digitalized asset exchange, pledge and mortgage

Technical Challenges in Blockchain for SCM

- Scalability
  - Network Scalability
  - Storage Scalability
- Throughput
  - BTC 7 TX/s vs Visa 24,000 TX/s
- Fine-grained Access Control
  - User Identity Data: anonymity, pseudonymity
  - Transactional Data: privacy, authenticity
- Data Retrieval
  - Efficiency and Reliability

Challenge 1: Scalability

- Network Scalability:
  - How can the Blockchain network scale with the increase number of nodes/participants? Consensus.

Properties	Cryptocurrency (PoW, PoS, ..)	Distributed System (Raft; PBFT; ..)
Strategy	Lottery-based (most)	Voting-based (most)
Real-world app.?	√	√
Proof in theory?	?	√
Fault Tolerance	Byzantine	Crash / Byzantine
Finality	Poor	Perfect
Throughput	Low	High
Latency	High	Low
Scalability	Good	Poor

Consensus Criteria

Finality: will all nodes always agree on the same single state?

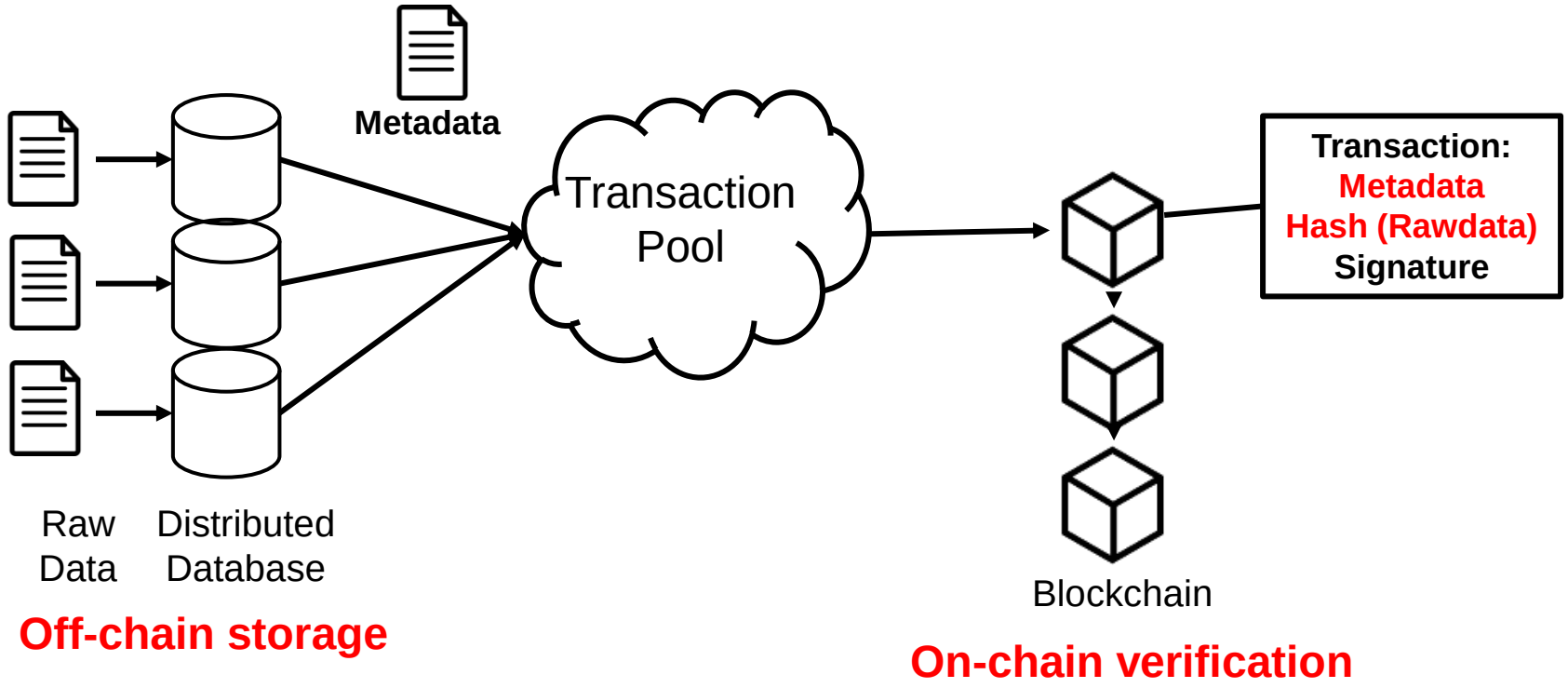
Latency: how long does it take from data submission to confirmation?

Throughput: how much data can be processed per unit time?

Scalability: how does number of nodes affect system performance?

Challenge 1: Scalability

- Storage Scalability:
  - Off-chain Storage & On-chain Verification



Challenge 2: Throughput

Bitcoin

Ethereum

~7 TPS **~30 TPS**

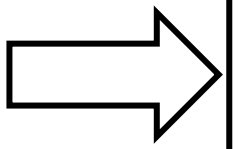
Low Throughput

Far from enough

VISA

~2,000 TPS

Bitcoin throughput 1.7 Kb/s, can hardly handle raw big data



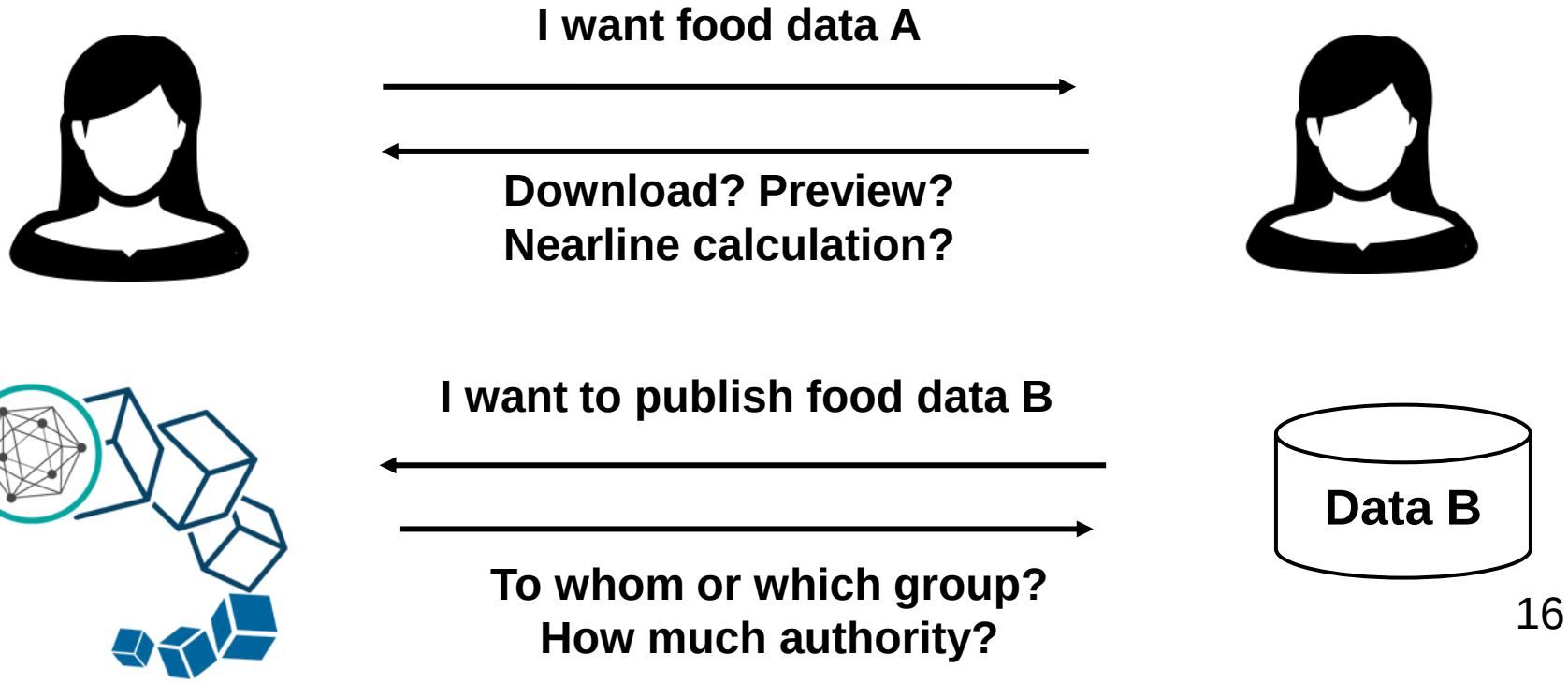
Combine Off-chain storage and On-chain verification

Challenge 2: Throughput

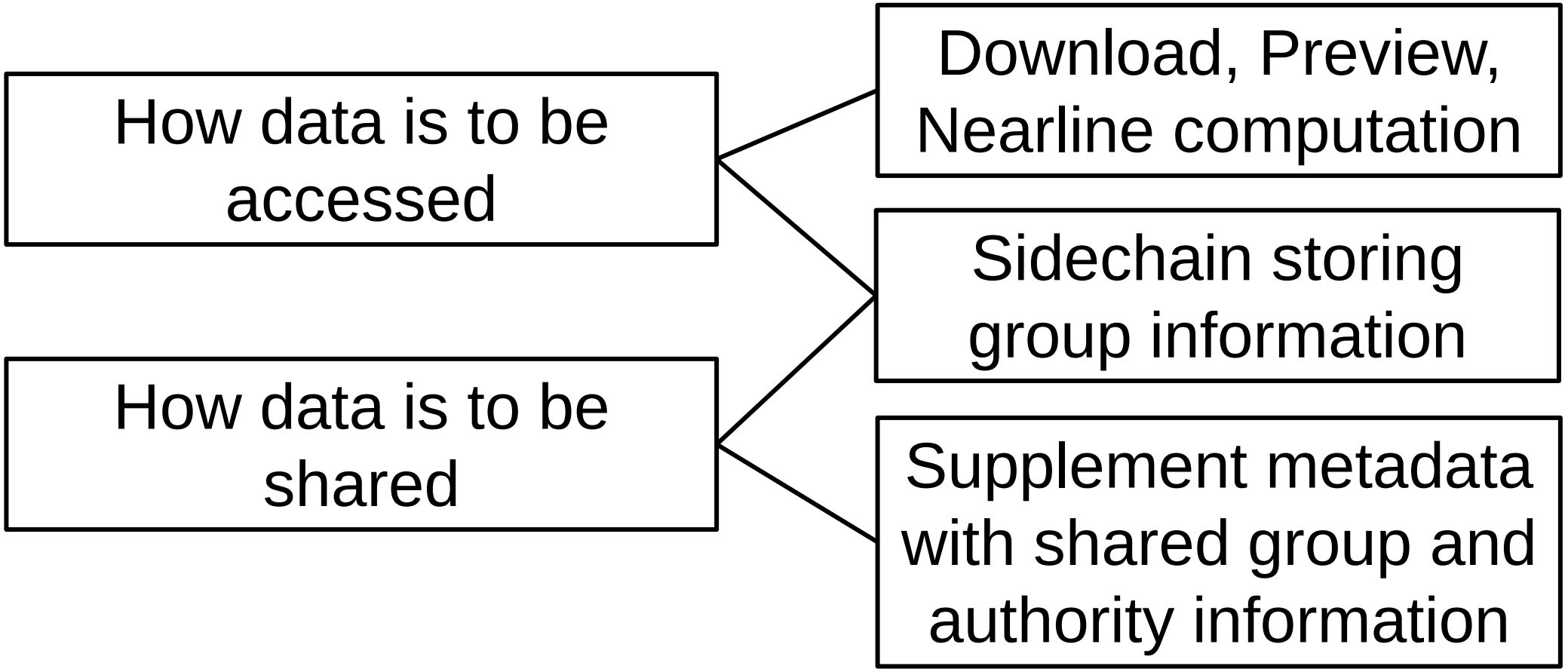
- Improve consensus of traditional distributed systems
  - Raft in R3CEV Corda (Usenix ATC'14): making Paxos practical
  - Algorand (SOSP'17)
- Adopt sharding from distributed database
  - ELASTICO (CCS'16)
  - OmniLedger (S&P'18)
  - RapidChain (CCS'18)
- New data serialization methods
  - DAG rather than chain: IOTA Tangle, Swirld Hashgraph
  - Microblocks: Bitcoin-NG(NSDI'16), ByzCoin(USENIX Security'16)

Challenge 3: Fine-grained Access Control

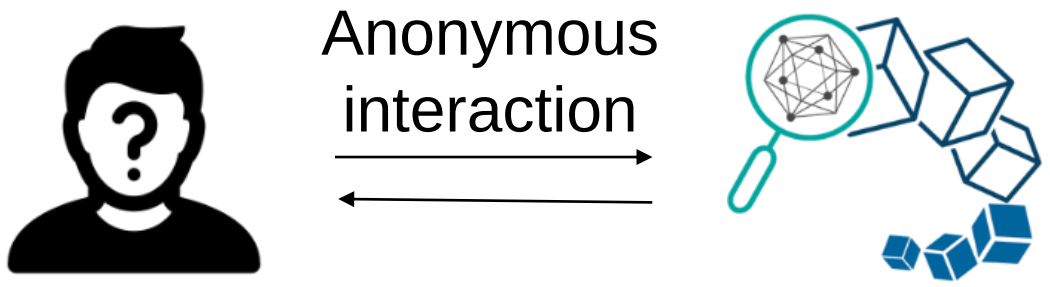
- Fine-grained Access Control
  - Transactional Data: privacy, authenticity
  - User Identity Data: anonymity, pseudonymity



Challenge 3: Fine-grained Access Control



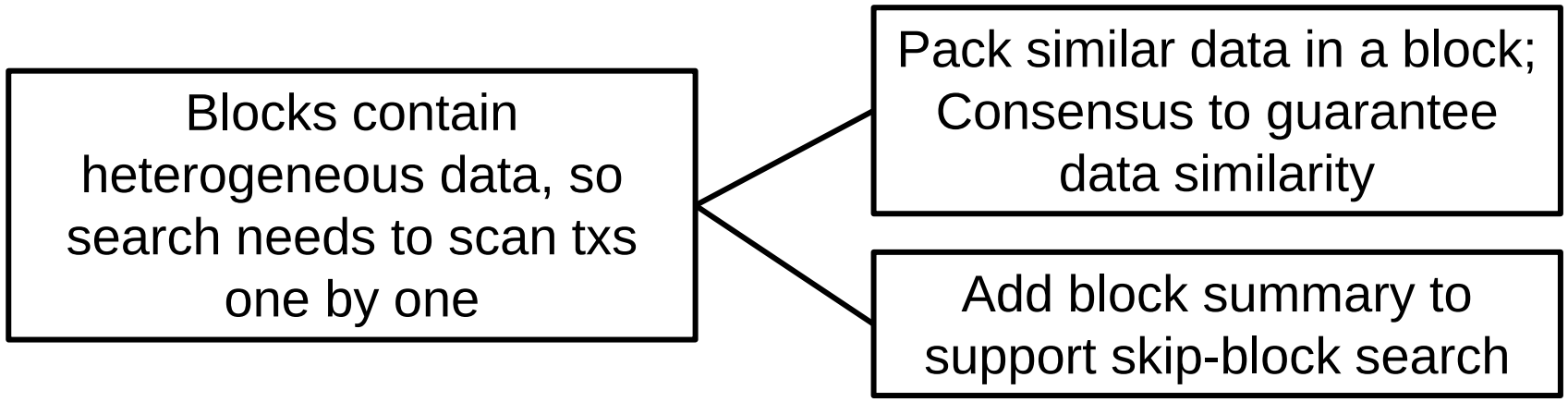
Challenge 3: Fine-grained Access Control



- Existing blockchain solutions are not enough for anonymity
  - Pseudonymous (Bitcoin): once revealed, forever exposed
  - Zero knowledge proof (Zcash): high computational resources
  - One-time ring signature (Monero): identity can be deduced

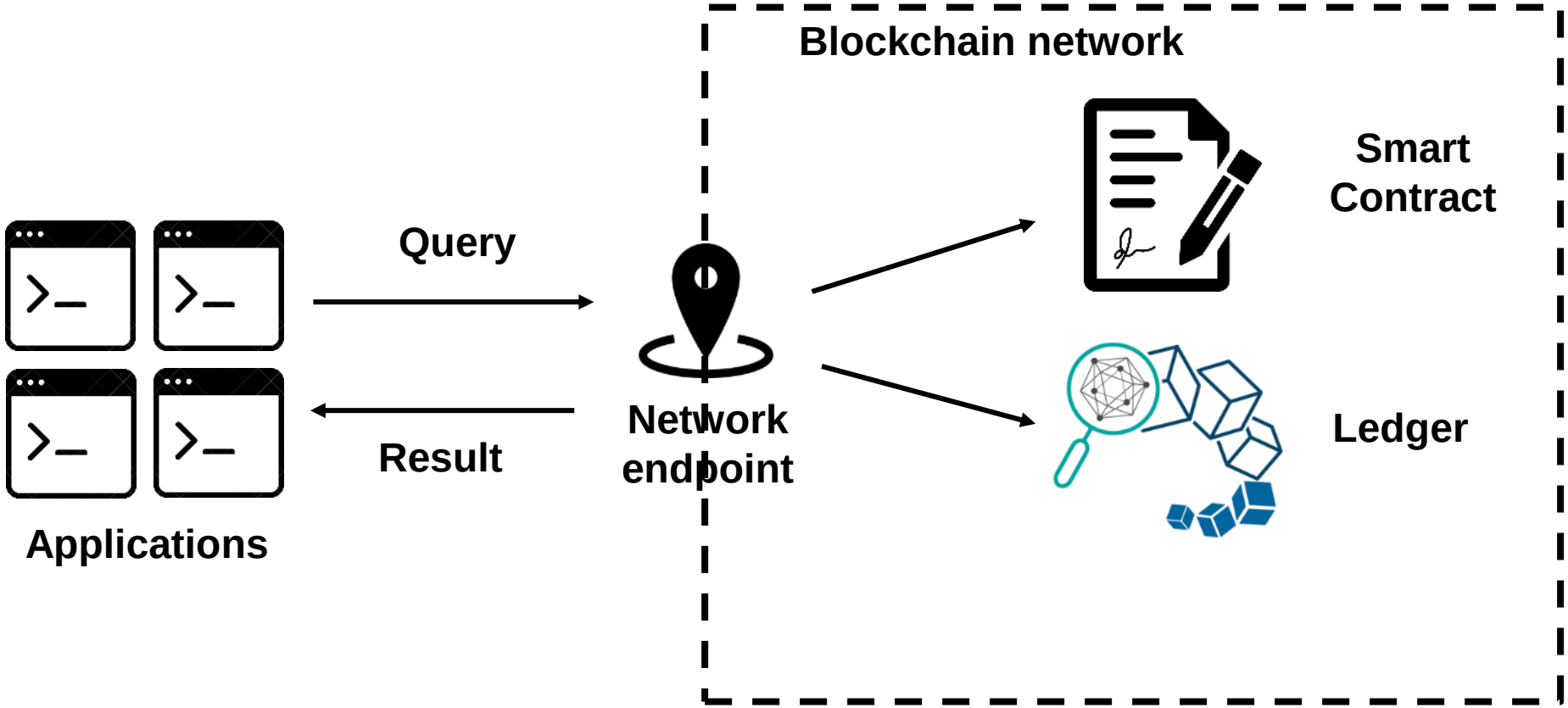
Challenge 4: Data Retrieval

- Two aspects to achieve low-latency data retrieval.
  - (1) Bootstrap searching history (e.g. cache searching result).
  - (2) Optimize how data is stored.



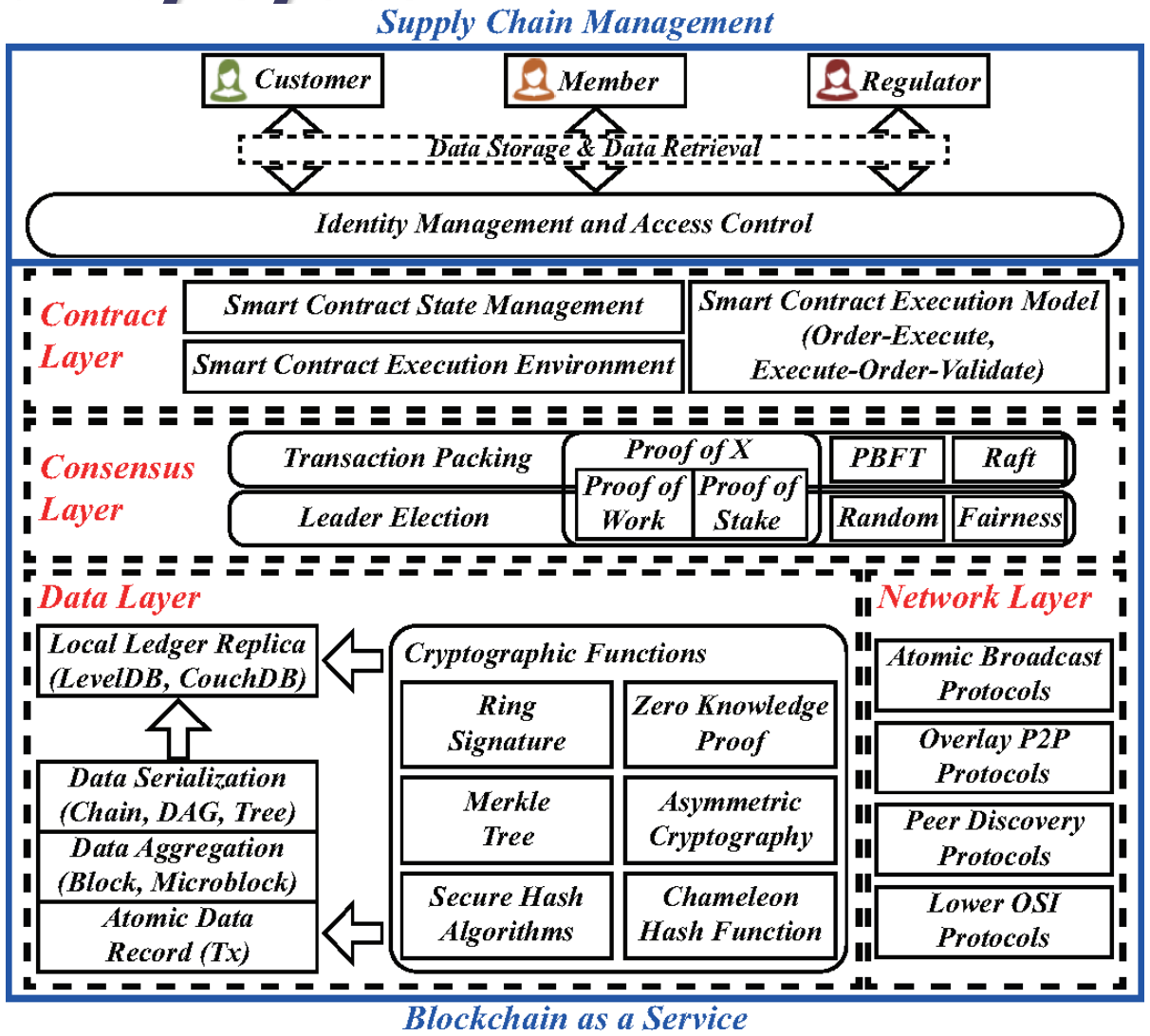
Challenge 4: Data Retrieval

- Reliable and Low-latency **for data retrieval**



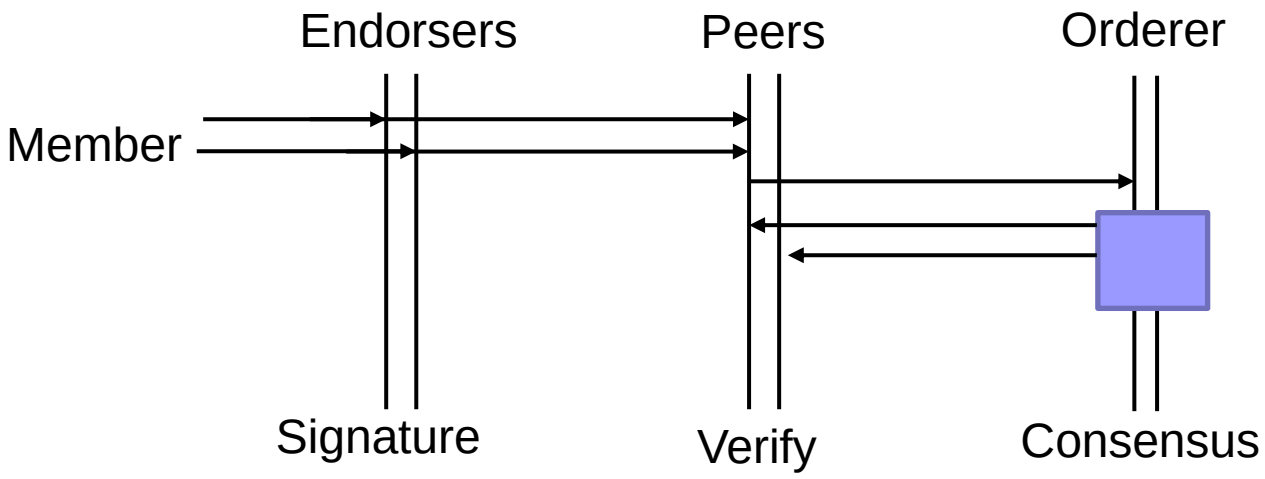
Framework of the blockchain-based food traceability system

- Blockchain as a Service
- Implementation
  - Hyperledger Fabric
  - Open-source
  - Permissioned Ledger
- Three kinds of identities
  - Customer
  - Member
  - Regulator



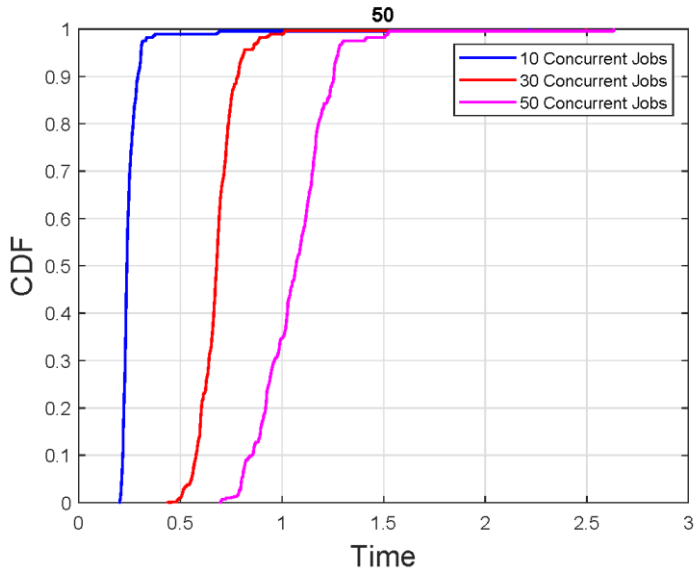
Experiment Setup

- System and Data Setup
  - 4 nodes: each running 4 docker containers
  - Each node: 2 for client peer, 1 for orderer and 1 for endorser

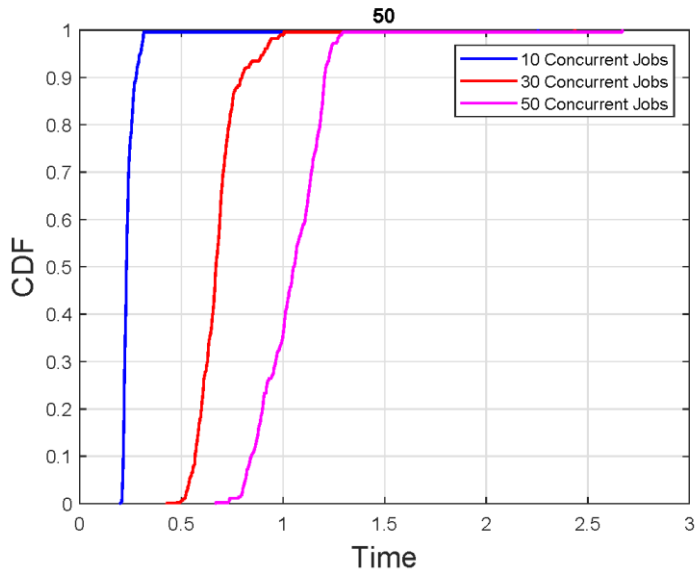


Experimental Result

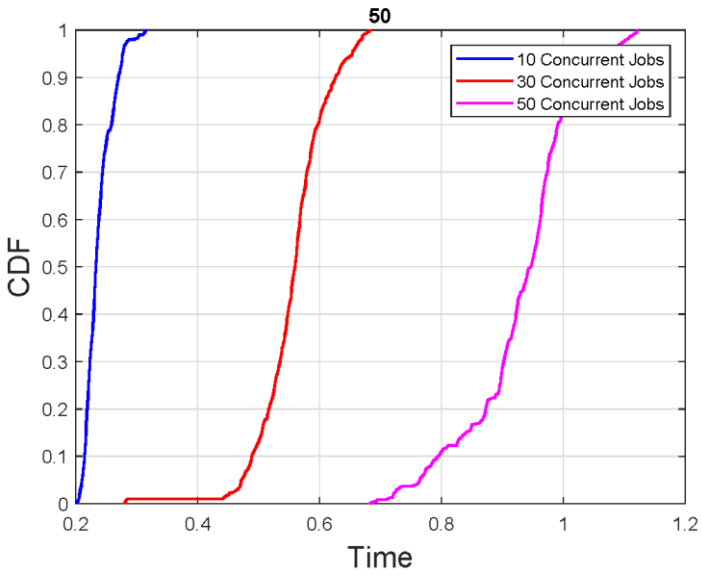
- System Performance Test
  - Invoke the chaincode with 10, 30, and 50 concurrent jobs at the same time



Member Registration Time



Transaction Uploading Time



Data Retrieving Time

Future Work

- Solution to 4 technical challenges
  - Scalability (Network & Storage)
  - Throughput
  - Fine-grained Access Control
  - Data Retrieval
- Real world supply chain deployment
  - Current system is using synthetic testing data with few nodes
  - Integrating real-world supply chain data with the current system and deploy the system on more federated nodes.

Acknowledgments

- This work is supported by Alibaba Innovative Research (AIR) Program by Alibaba (China) Co., Ltd. - H-ZG6N, Hong Kong RGC Research Impact Fund (RIF) - R5034-18, and Hong Kong RGC Collaborative Research Fund - CityU C1008-16G.

References

- K. Yoneyama and S. Kimura, “Verifiable and forward secure dynamic searchable symmetric encryption with storage efficiency,” Springer ICICS 2017
- L. Xu, L. Chen, Z. Gao, Y. Lu, and W. Shi, “Coc: Secure supply chain management system based on public ledger,” in 26th International Conference on Computer Communication and Networks, ICCCN 2017, Vancouver, BC, Canada, July 31 - Aug. 3, 2017, 2017, pp. 1–6.
- F. Tian, “A supply chain traceability system for food safety based on haccp, blockchain & internet of things,” in 2017 International Conference on Service Systems and Service Management. IEEE, 2017, pp. 1–6.
- D. D. F. Maesa, P. Mori, and L. Ricci, “Blockchain based access control,” in Distributed Applications and Interoperable Systems - 17th IFIP WG 6.1 International Conference, DAIS 2017, 2017, pp. 206–220.

Q&A