LedgerDB

LedgerDB : Alibaba’s Centralized Ledger Database

- Xinying(Derry) YANG

Alibaba 阿里巴巴 蚂蚁金服 ANT FINANCIAL

LedgerDB

Terminologies

- DLT (Decentralized Ledger Technology)
- CLT (Centralized Ledger Technology)
  - CLD (Centralized Ledger Database): LedgerDB, QLDB, Oracle BC Table, ProvenDB, etc.
- Immutability: Any piece of data, once committed into the system, cannot be modified by subsequent operations and becomes permanently available.
- Verifiability: The capability of validating specific data integrity and operation proofs.
- Auditability: The capability of observing a serial of user actions and operation trails based on predefined audit rules.
  - Internal audit: an internal user of the ledger can observe and verify the authenticity of all actions.
  - External audit: an external third-party entity can observe and verify the authenticity of all actions.

Alibaba 阿里巴巴 蚂蚁金服 ANT FINANCIAL

LedgerDB

Credibility for Traditional Database Applications

- Centralized DBMS
- Cloud (Distributed) DBMS
- Bigdata & No-SQL

$\neq$ Trustable Data Source

Here comes ledger technique

LedgerDB

DLT Dilemma

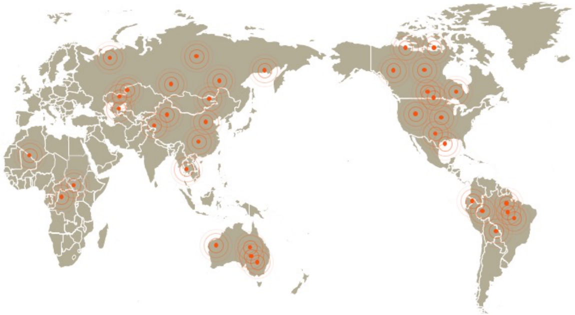
Permissionless blockchains: Bitcoin, Ethereum, etc.

Pros:

- Massive peers, widely spread, highly decentralized

Cons:

- Extremely low TPS (**7** for Bitcoin)



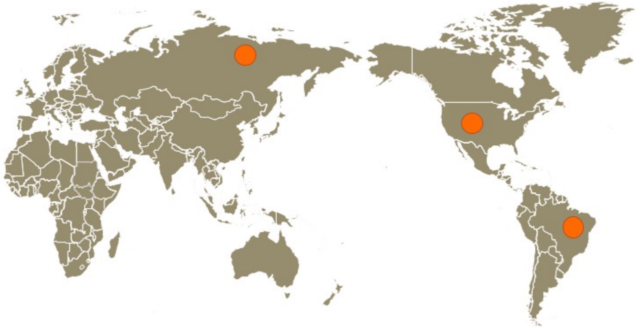
Permissioned blockchains: Fabric, Corda, Quorum, etc.

Pros:

- Improved TPS, still can not be compared with RDBMS or NoSQL

Cons:

- Few peers, consensus can be broken/manipulated by malicious nodes



Ease of use + **NoSQL performance** + **Blockchain credibility** = **?**

Alibaba 阿里巴巴 蚂蚁金服 ANT FINANCIAL

LedgerDB

Why **CLD** is important & valuable ?

- Motivations
  - Decentralization is not proved to be indispensable for permissioned blockchain.
  - Conventional permissioned blockchain and CLD systems:
    - Low performance, storage overhead, regulatory issues, limited external auditability
- Gartner Forecast
  - Gartner Strategic Vision 2019

    Strategic Planning Assumption

    By 2021, at least 20% of projects envisioned to run on permissioned blockchains will instead run on centralized, auditable ledgers.

  - Gartner Strategic Vision 2020

    By 2021, most permissioned blockchain uses will be replaced by ledger DBMS products.

Alibaba 阿里巴巴 蚂蚁金服 ANT FINANCIAL

Highlight and Comparison

- LedgerDB – a ledger database that provides tamper-evidence and non-repudiation features in a centralized manner (CLD), which realizes strong auditability, high performance, and data removal support.

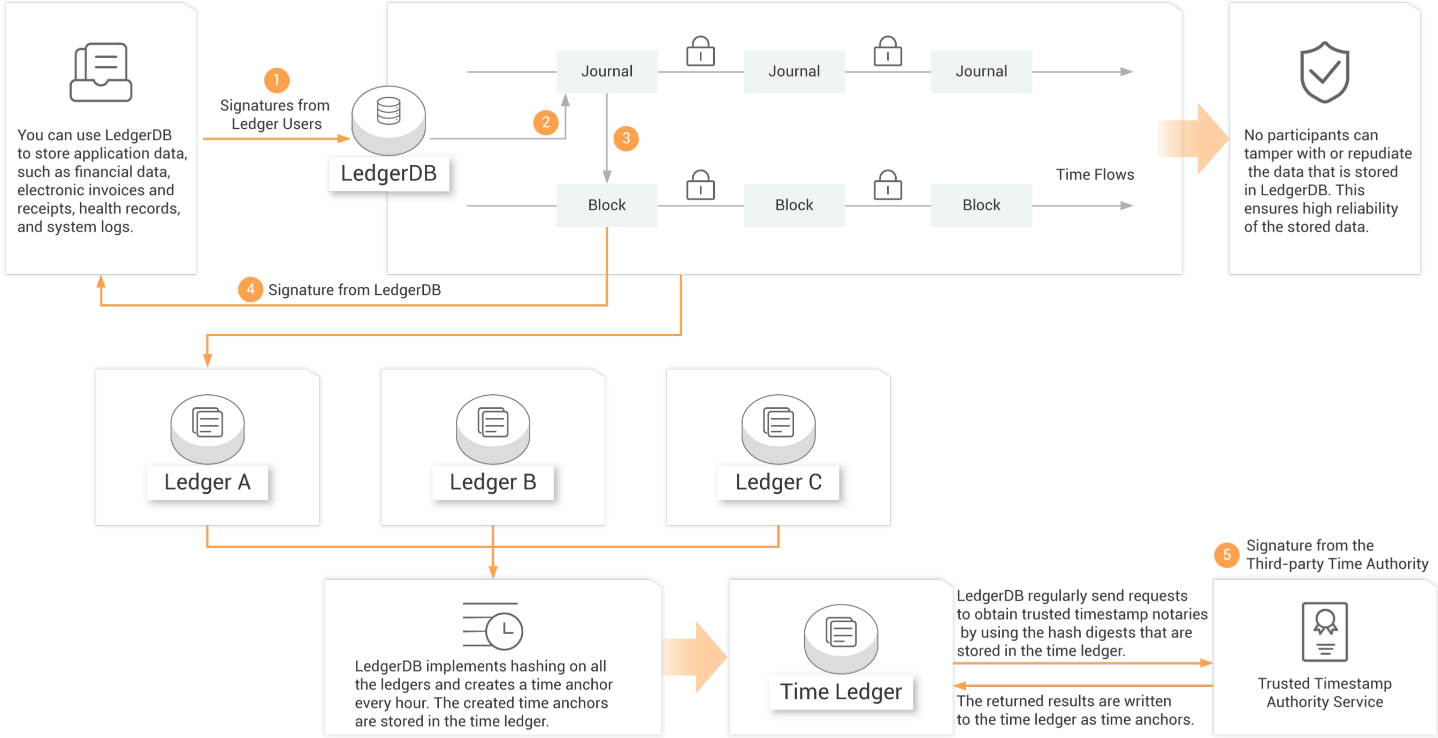
- Key comparisons between LedgerDB and other systems.

System	Throughput (max TPS)	Auditability				Removal		Non-Repudiation		Provenance
		external	third party	peg	capability	purge	occult	server-side	client-side	native clue
LedgerDB	100K+	✓	TSA	✓	strong	✓	✓	✓	✓	✓
QLDB [7]	1K+	✗	✗	✗	weak	✗	✗	✗	✗	✗
Hyperledger [6]	1K+	✗	✗	✗	weak	✗	✗	✓	✓	✗
ProvenDB [40]	10K+	✗	Bitcoin	✓	medium	✗	✓	✗	✗	✗
Factom [43]	10+	✓	Bitcoin	✓	strong	✗	✗	✓	✓	✗

Alibaba 阿里巴巴 蚂蚁金服 ANT FINANCIAL

LedgerDB

How it works

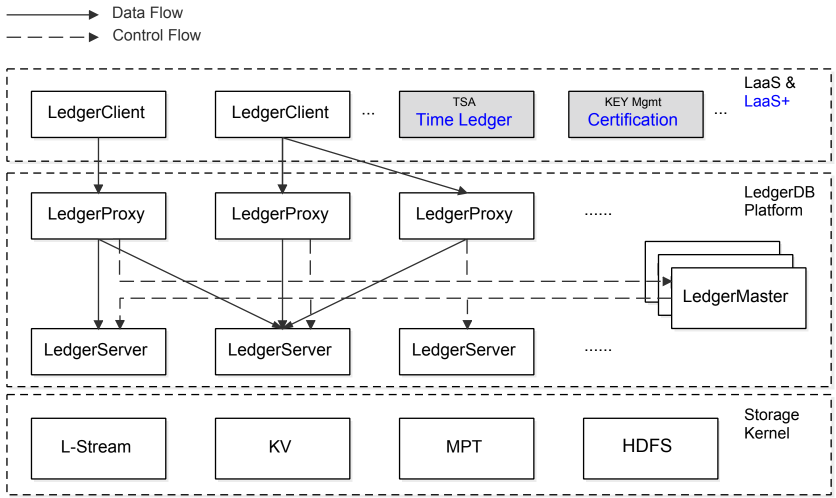


The time ledger is a public ledger.

LedgerDB

LedgerDB system architecture.

Ledger master - manage the runtime metadata of the entire cluster (e.g., status of servers and ledgers) and coordinate cluster-level events (e.g., load balance, failure recovery).
Ledger proxy - receive client requests and preprocesses, and then dispatch them to the corresponding ledger server.
Ledger server - complete the final processing of requests, and interact with underlying storage layer that stores ledger data.



LedgerDB Operators and APIs.

Append - append user transaction or system-generated transaction to ledger.
Retrieve - get qualified journals from ledger.
Verify - verify integrity and authenticity of returned journals from journal proofs.
Create - create a new ledger with initial roles and members.
Purge - remove obsolete journals from ledger.
Occult - hide journal(s) from ledger.
Recall - rollback a purge (within a limited time window).
Delete - removes entities in the system, such as a ledger, a role, a member, or a clue.

Operator	Method
Create	`Create(ledger_uri, enum, op_metadata)`

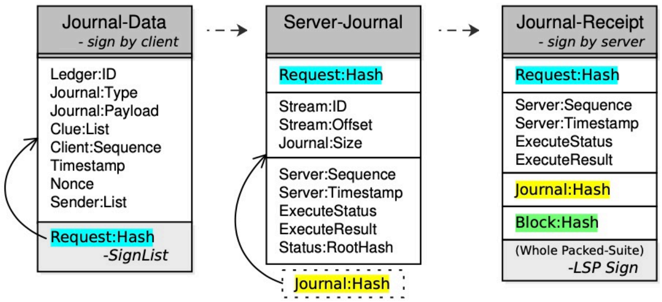
Append	`AppendTx(ledger_uri, tx_data, clue, set)`
	`SetTrustedAnchor(ledger_uri, jsn, level)`
	`GrantRole(ledger_uri, member_id, role)`
	`GrantTime(ledger_uri, timestamp, proof)`

Retrieve	`GetTx(ledger_uri, jsn)`
	`ListTx(ledger_uri, ini_jsn, limit, clue)`
	`GetTrustedAnchor(ledger_uri, jsn, level)`
	`GetLastGrantTime(ledger_uri, timestamp)`

Verify	`Verify(ledger_uri, jsn`|`clue, data, level)`
Purge	`Purge(ledger_uri, block)`
Occult	`Occult(ledger_uri, jsn`|`clue)`
Recall	`Recall(ledger_uri, purged_point)`
Delete	`Delete(ledger_uri, enum, op_metadata)`

LedgerDB

Journal Management

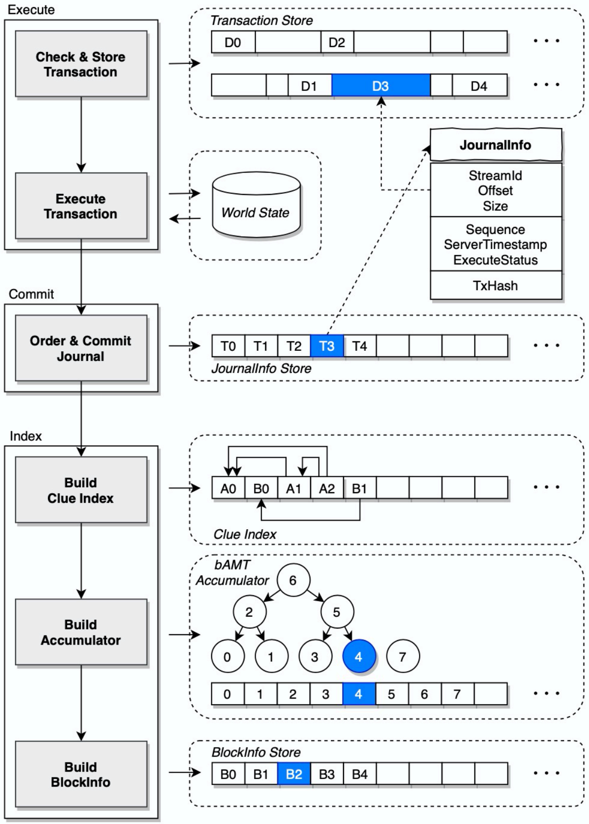


LedgerDB adopts an *execute- commit-index* transaction management approach:

① execute - a transaction first enters the execute phase based on its transaction type. It runs on ledger proxy for better scalability.

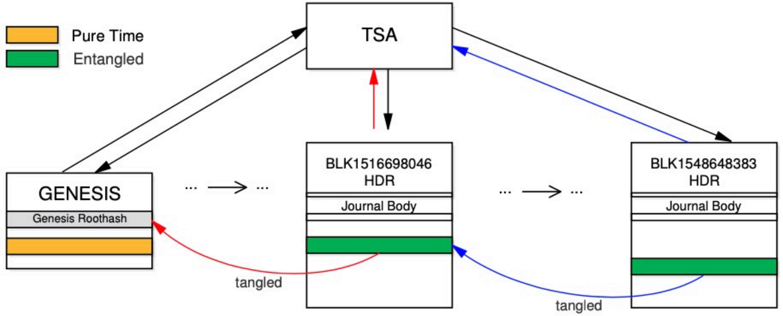
② commit - collect multiple executed transactions, arranges them in a global order (jsn), and persist them to the storage system. It runs on ledger server.

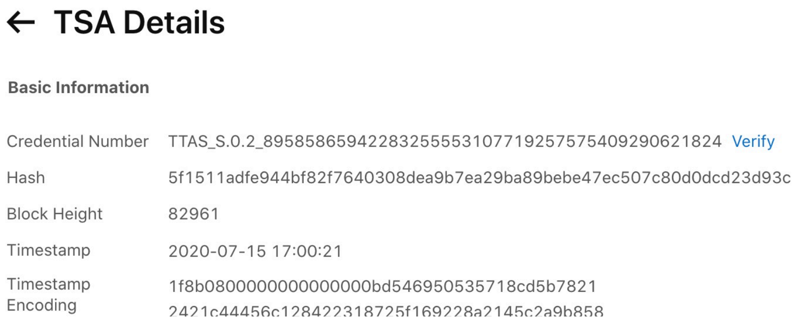
③ index - start on ledger server to build indexes for subsequent data retrieval and verification.



Alibaba 阿里巴巴 蚂蚁金服 ANT FINANCIAL

Two-way peg TSA notary journals

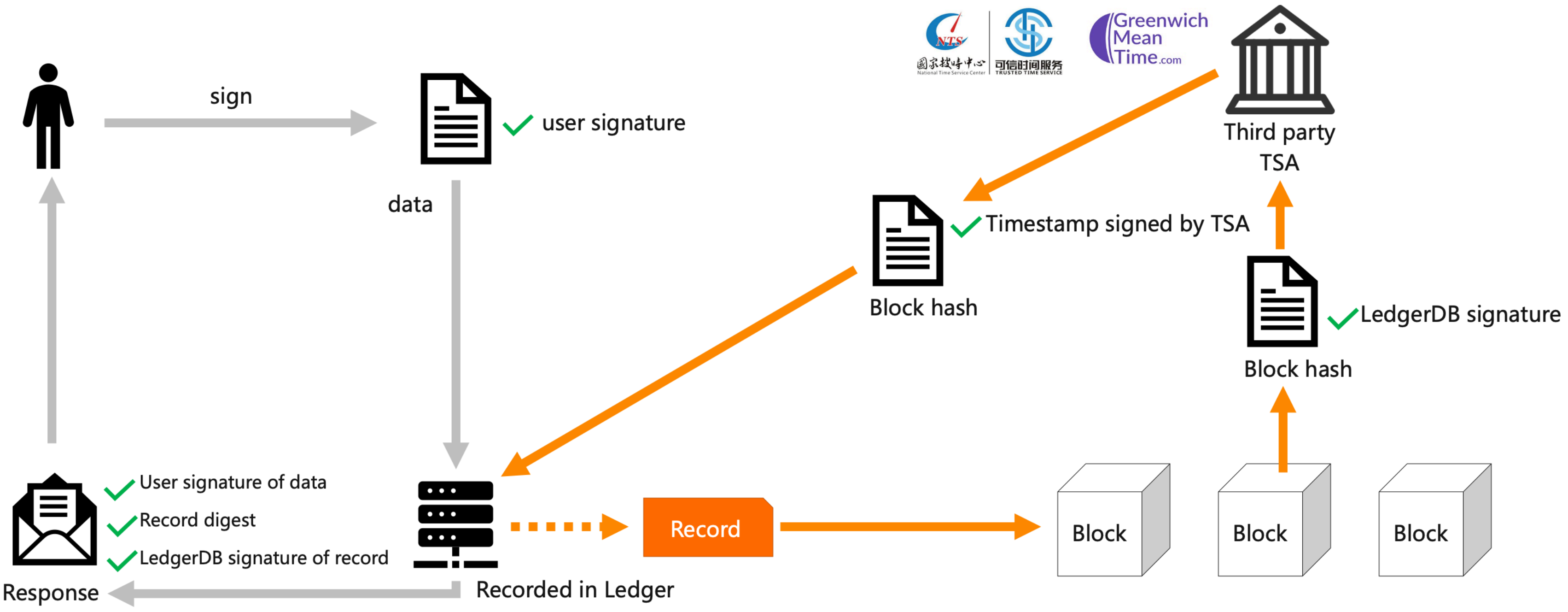




- A TSA journal contains a ledger snapshot (i.e., a ledger digest) and a timestamp, signed by TSA in entirety. These journals are mutually entangled between each other, which provide external auditability for timestamps.
- Two-way peg protocol: ① a ledger digest is first submitted and then signed by TSA;
  ② TSA journal is recorded back on ledger as a TSA journal.
- We offer T-Ledger service on Alibaba Cloud LaaS+ (Ledger-as-a-Service).

LedgerDB

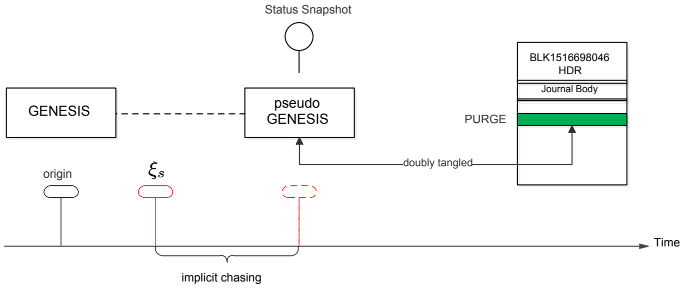
Credibility guarded by multipart signatures



Verifiable Data Removals

- Purge

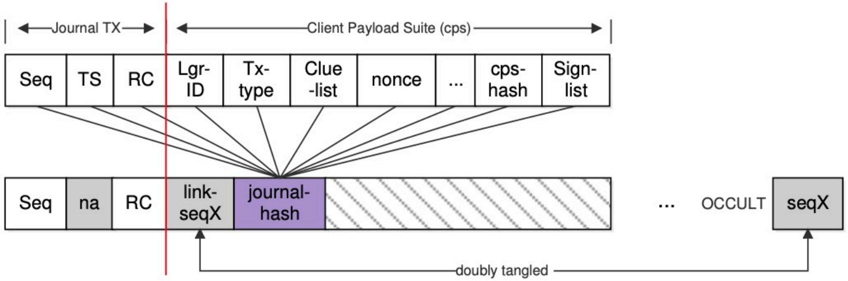
A purge operation deletes a set of contiguous (obsolete) journals starting from genesis to a designated jsn on ledger




```
DELETE FROM ledger_uri
 WHERE jsn < pur_jsn;
```

- Occult

An occult operation converts the original journal to a new one that only keeps its metadata, and retains its digest.

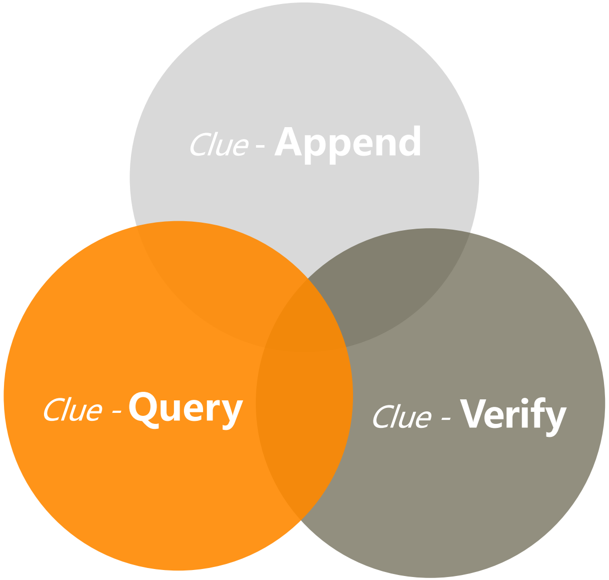


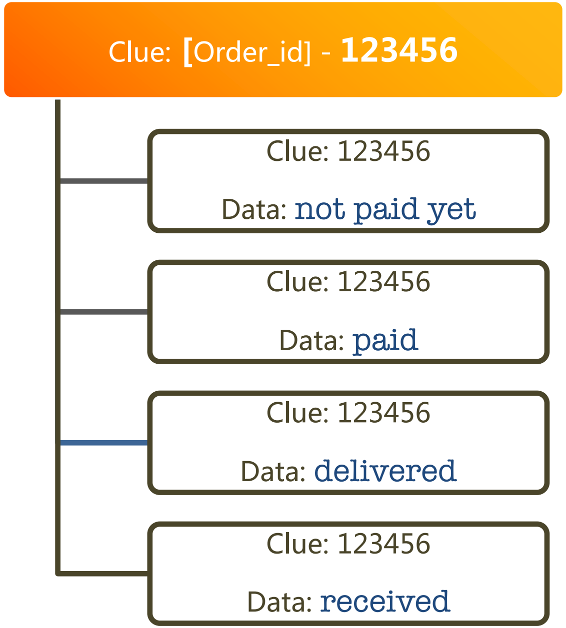

```
UPDATE ledger_uri
 SET TS = na, cps = CONCAT(
  seqX, journal_hash, blanks)
   WHERE jsn = Seq
      OR cid = des_cid;
```

Alibaba 阿里巴巴
蚂蚁金服 ANT FINANCIAL

Clue – Native lineage in LedgerDB

- A clue is a user-specified label (key) that carries on business logic for data lineage.
- Quick index is supported to fetch or verify through related events in chronological order.

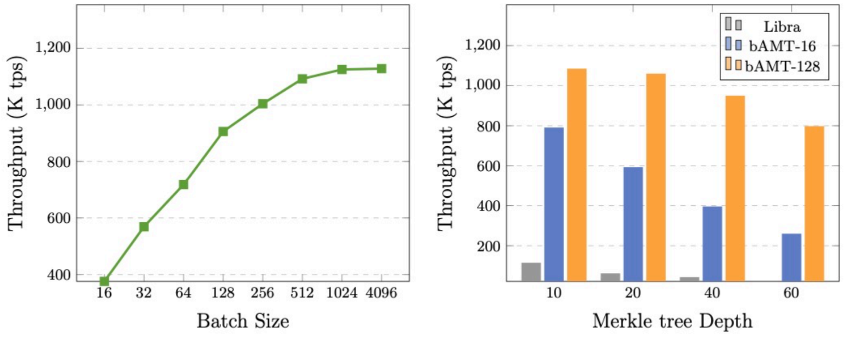




LedgerDB

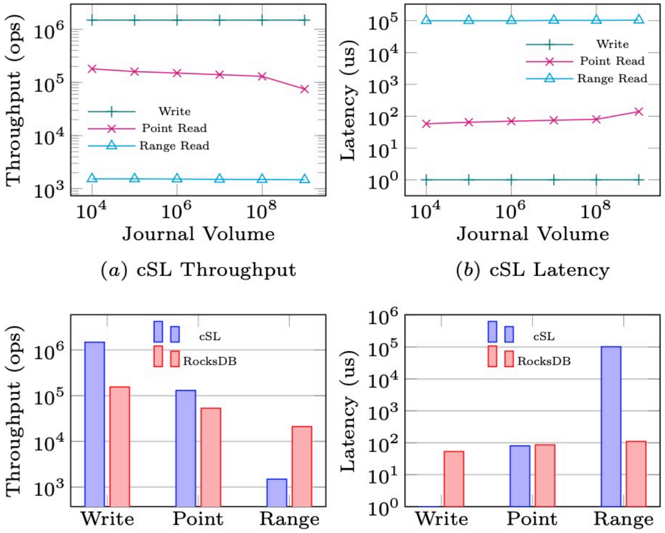
Evaluation – clue Skiplist (cSL) & batch accumulated Merkle-tree (bAMT)

bAMT vs. Libra accumulator



(a) bAMT root calculation (b) bAMT vs. Libra

cSL vs. RocksDB



(a) cSL Throughput (b) cSL Latency

(c) Throughput comparison (d) Latency comparison

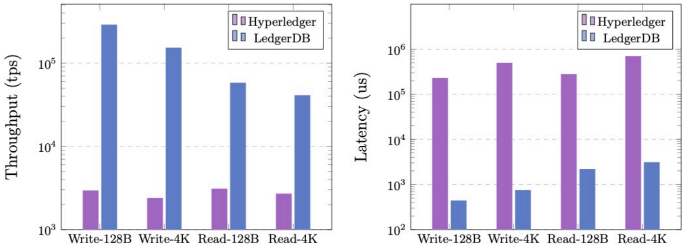
Alibaba 阿里巴巴 蚂蚁金服 ANT FINANCIAL

LedgerDB

Evaluation – performance and appl

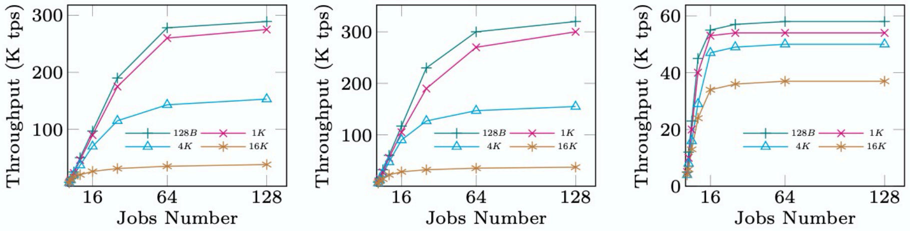
LedgerDB end-to-end performance

LedgerDB is 80X faster compared to Hyperledger Fabric in the same notarization application



(a) Throughput comparison

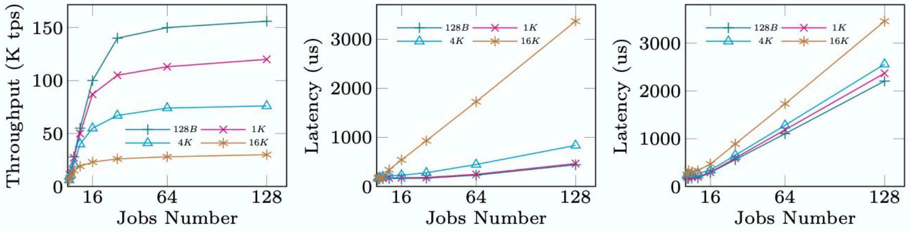
(b) Latency comparison



(a) Write

(b) Sequential Read

(c) Random Read



(d) Latest Random Read

(e) Write

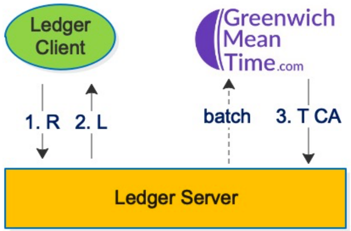
(f) Random Read

Alibaba 阿里巴巴

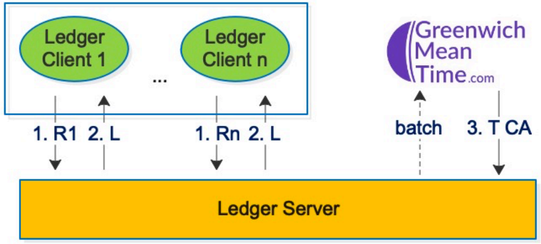
蚂蚁金服 ANT FINANCIAL

LedgerDB

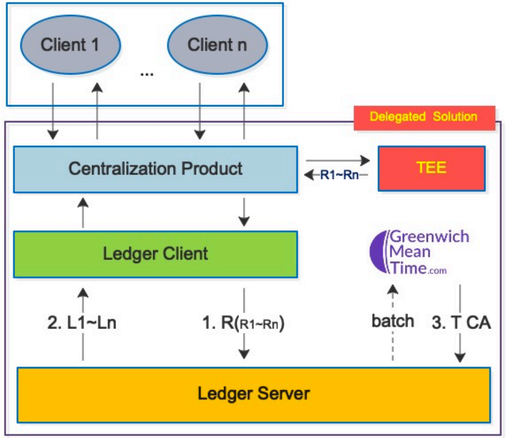
LedgerDB Solution Category



Mono Ledger



Federal Ledger

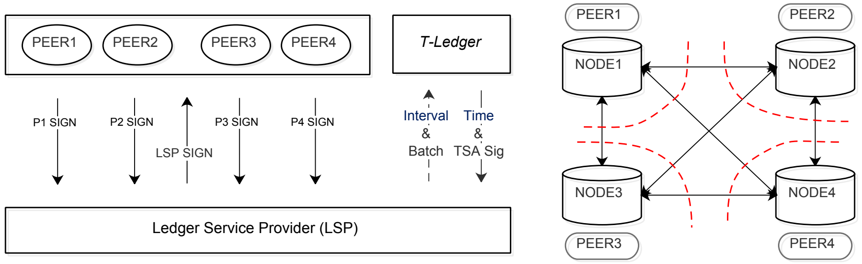


Delegated Ledger

LedgerDB

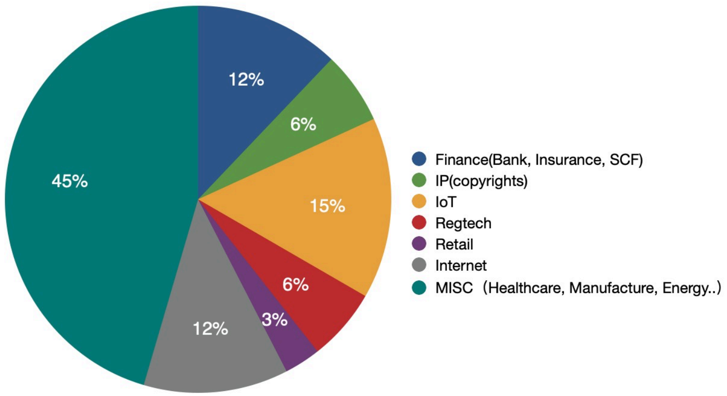
LedgerDB in Production

Federated ledger vs. permissioned blockchain



(a) (b)

LedgerDB customer use cases



Alibaba 阿里巴巴 蚂蚁金服 ANT FINANCIAL

LedgerDB

Decentralized vm-like exec is just an implementation, the soul of consensus in ledger technique is dancing with time and cryptographic theorem.

- LedgerDB

https://www.alibabacloud.com/product/ledgerdb

xinying.yang@alibaba-inc.com

Thanks!