A New Approach to the Development of Provenance Metadata Management Systems for Large Scientific Experiments

A. Demichev, A. Kryukov and **N. Prikhod'ko**

SINP MSU

Novgorod State University

Supported by RSF grant No. 18-11-00075

Provenance Metadata (PMD)

- Metadata describing data, provide context and are vital for accurate interpretation and use of data
- One of the most important types of metadata is provenance metadata (PMD)
  - tracking the stages at which data were obtained
  - ensuring their correct storage, reproduction and interpreting
  - ⇒ ensures the correctness of scientific results obtained on the basis of data
- The need for PMD is especially essential when large volume (big) data are jointly processed by several research teams

Types of storages: extremal cases

- Centralized
  - problems:
    - very expensive ⇒ funding ?
    - planning in advance the necessary storage capacity
- P2P-storage with special mechanisms of coding, fragmentation and distribution
  - problems:
    - to ensure a stable pool of resource providers,
    - before such a P2P-based storage can work stably, it requires significant technical, organizational and time costs in the absence of a result guarantee

Types of storages: intermediate solution

- organizations participating in a large project
  - integrate their local storage resources into a unified distributed pool
  - if necessary, rent in addition cloud storage resources, perhaps from multiple providers.
- may be particularly advantageous
  - if there is a need to store large amounts of data for a limited duration of a project
  - in a situation where the project brings together many organizationally unrelated participants
- ⇒ dynamically changing distributed environment

PMD MS Construction: Distributed Solution

- distributed environment ⇒ distributed registry for PMD
- we suggested to use the blockchain technology which provides
  - that no records were inserted into the registry in hindsight
  - no entries were changed in the registry
  - the registry has never been damaged or branched
  - monitoring and restoring the complete history of data processing and analysis

PMD MS Construction: Which Blockchain (1/2)

- type of the blockchains
  - permissionless blockchains, in which there are no restrictions on the transaction handlers
  - permissioned blockchains, in which transaction processing is performed by specified entities
- permissionless:
  - algorithms are based on
    - Proof-of-Work – highly resource-consuming, probability of reaching a consensus, which grows with time elapsing, ...
    - Proof-of-Stake – Nothing-at-Stake problem,…
  - suitable for open (public) networks of participants (Bitcoin, etc.)

PMD Projects Based on Permissionless Blockchains

- ProvChain, SmartProvenance: intended for a cloud storage
  - no DDS, no different administrative domains, no real consensus among the potentially conflicting parties
- Storj, Sia: intended for a P2P network of public storage resources
  - public blockchain - mainly for providing mutual settlements between suppliers and consumers of (P2P) resources
  - very restricted PMD facilities

PMD MS Construction: Which Blockchain (2/2)

- Permissioned:
  - there is a fixed number of trusted transaction/blockchain handlers
  - the handlers must come to a consensus about the content and the order of the recorded transactions
    - distributed consensus algorithm should be involved
  - form a more controlled and predictable environment than permissionless blockchains
  - suitable for networks with naturally existing trusted parties
    - our case: DMS, data owners,...

System state

- The state of the entire distributed storage = aggregated state of the set of files stored in it with their states at the moment
- The state of a data file is determined by PMD:
  - global ID + attributes, including:
    - local file name in a storage: fileName;
    - storage identifier: storageID;
    - creator identifier: creatorID;
    - owner identifier: ownerID
    - type: type=primary/secondary/replica
    - …

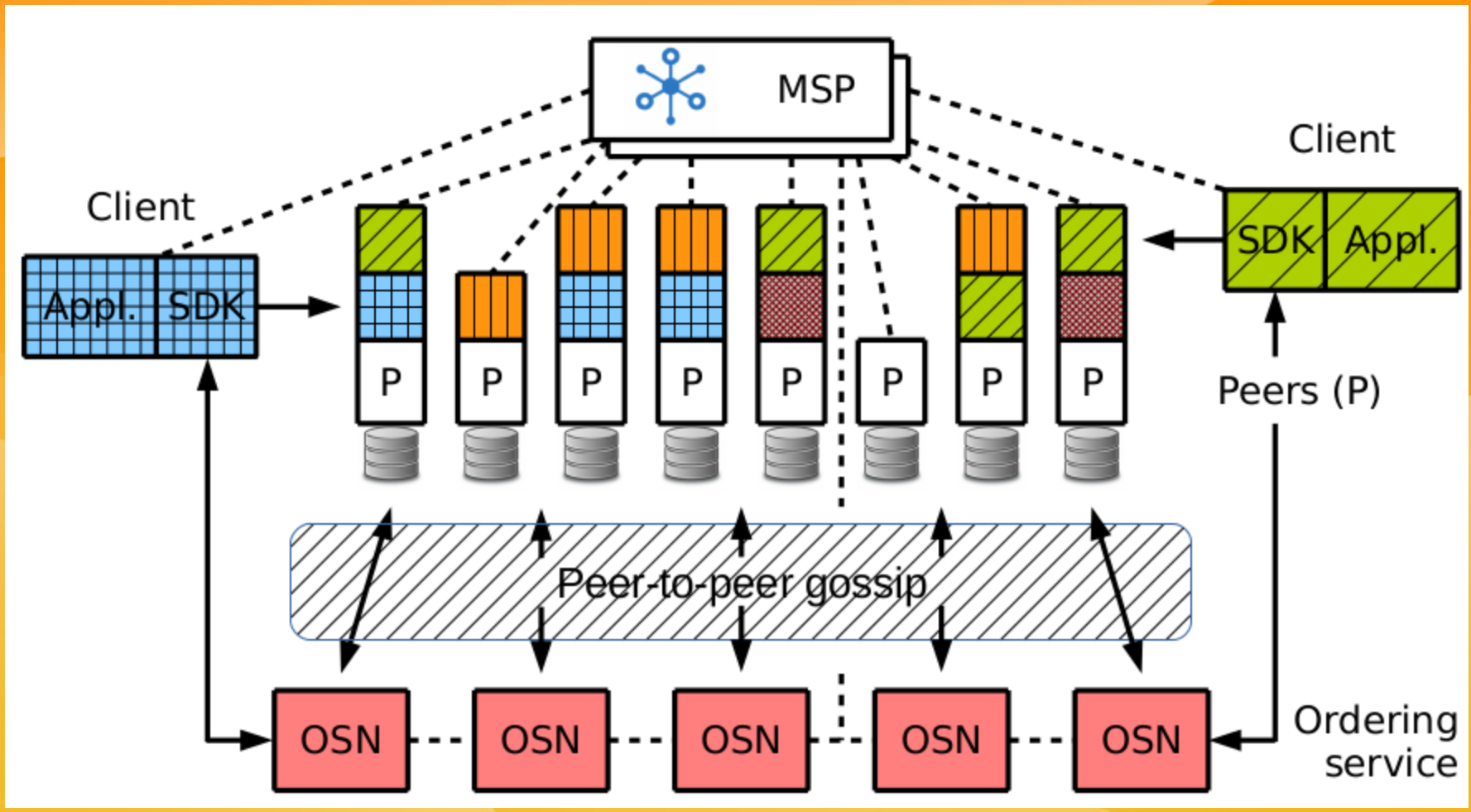
Basic transactions

- new file upload
- file download
- file deletion
- file copy
- copying a file to another repository
- transferring a file to another repository
  - each active transaction ⇒ update of some state attributes
    - for example, after the transaction "file download" the values of the keys change: "number of file downloads" and "users who downloaded the file".

HyperLedger Fabric (1/2)

- Analysis of existing platforms shows that the formulated problems most naturally can be solved on the basis of the
  - Hyperledger Fabric blockchain platform (HLF; www.hyperledger.org)
  - together with Hyperledger Composer (HLC; hyperledger.github.io/composer) = set of tools for simplified use of blockchains
- permissioned blockchains
  - transactions are processed by a certain list of trusted network members

HyperLedger Fabric (2/2)



From: *E. Androulaki et al. "Hyperledger Fabric: A Distributed Operating System for Permissioned Blockchains," in Proc 13th EuroSys Conf. 2018*

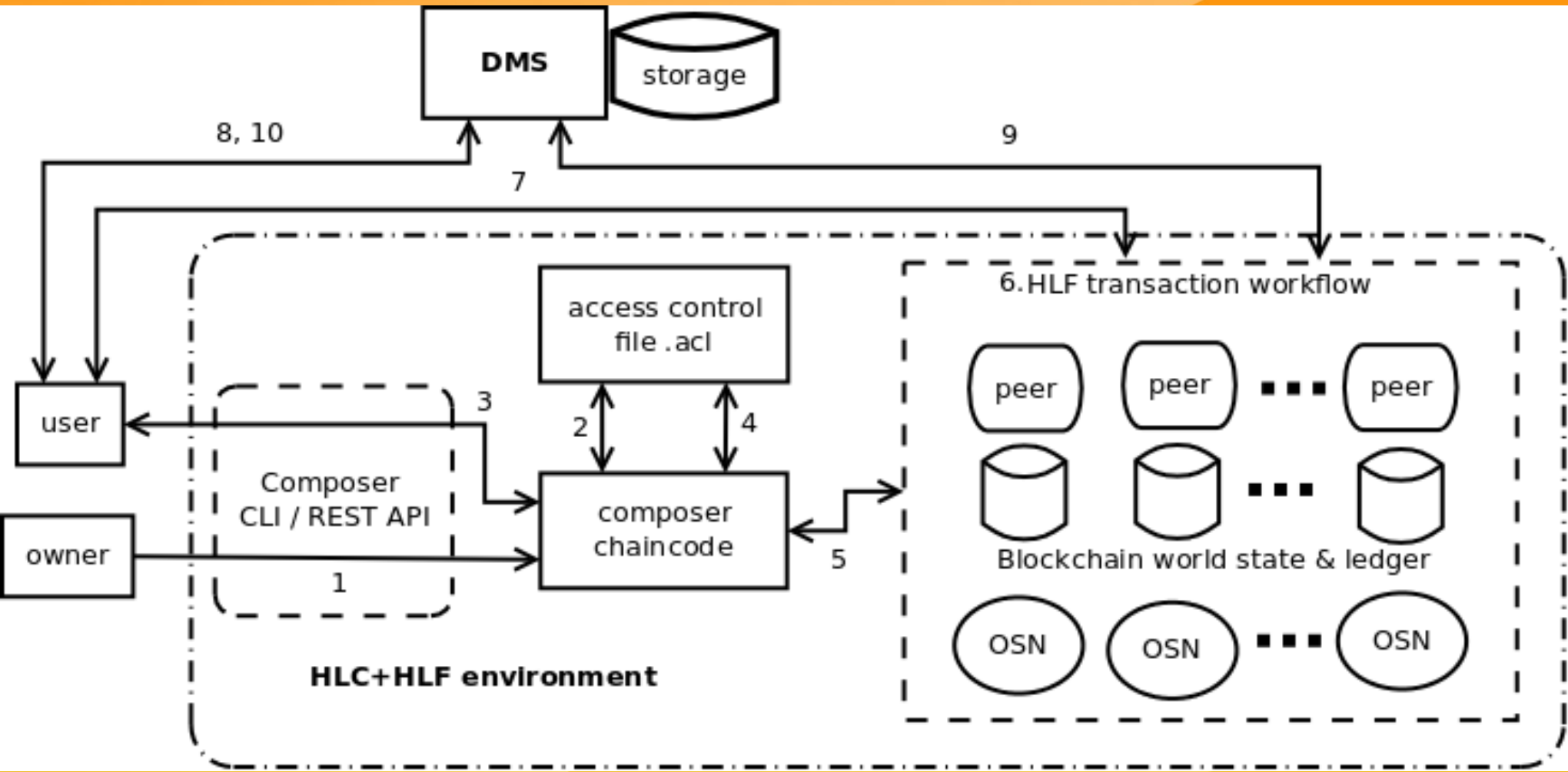
HyperLedger Fabric → ProvHL (1/2)

- ProvHL = Provenance HyperLedger
- operation of smart contracts (chaincodes)
  - adaptation of HLF for the business process of sharing storage resources
- provides a record of transactions & advanced query tools
- advanced means for managing access rights
  - access rights can be managed by network members within their competence

HyperLedger Fabric → ProvHL (2/2)

- thanks to its modular structure, it allows using different algorithms to reach consensus between business process participants
- has a developed built-in security system based on PKI

ProvHL operation



Simplified scheme for recording transactions with provenance metadata and managing data access rights based on HLF&C

Conclusion (1/2)

- we have suggested the new approach to the development of management system for PMD and data access rights
  - based on the integration of blockchain technology, smart contracts and metadata driven data management
- it is proposed to use
  - permissioned type of blockchain
  - Hyperledger blockchain platform, on the basis of which the ProvHL system is implemented.
    - fault-tolerant, safe and secure management system of provenance metadata, as well as access rights to data in distributed storages

Conclusion (2/2)

- At present, a testbed has been created on the basis of the SINP MSU
  - a preliminary version of the ProvHL prototype was deployed to implement the developed principles and refine the algorithms of the system
- In this preliminary version,
  - a trivial consensus algorithm is used, in which the transaction recording order is determined by a single server (centralized orderer Solo in the terminology of HLF).
  - in the future it is supposed to use full-fledged Byzantine fault tolerant consensus algorithms, in particular PBFT