

Blockchain for/with Information-Centric Network

李 睿棟

朝枝 仁

国立研究開発法人 情報通信研究機構

2019年12月26日

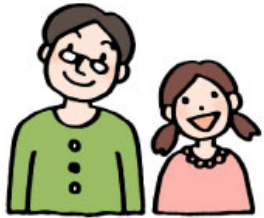
What will the future be?

Cyber Space

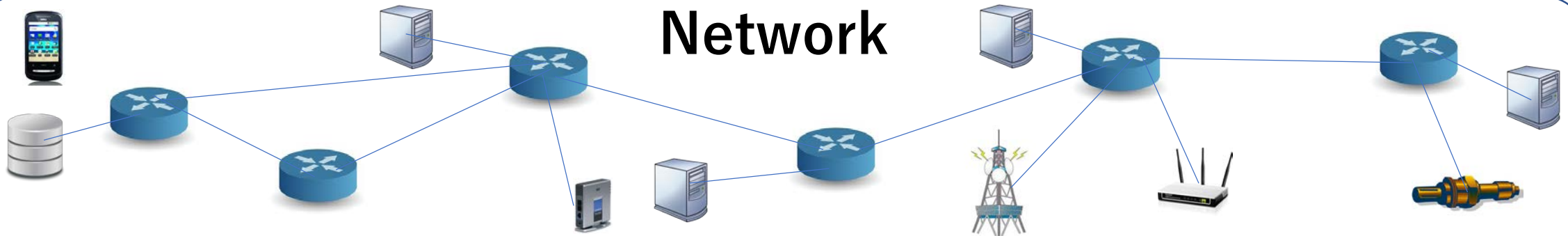
How to intelligently, efficiently, and securely process and provide the data to the persons who need them?

1. Decentralization (Move from centralized cloud/server to edges)
2. Ubiquitous data processing/caching/storage
3. Trust, privacy, auditability among unknown entities

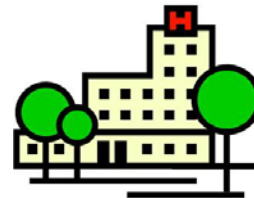
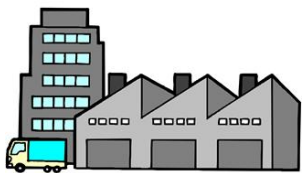
Data



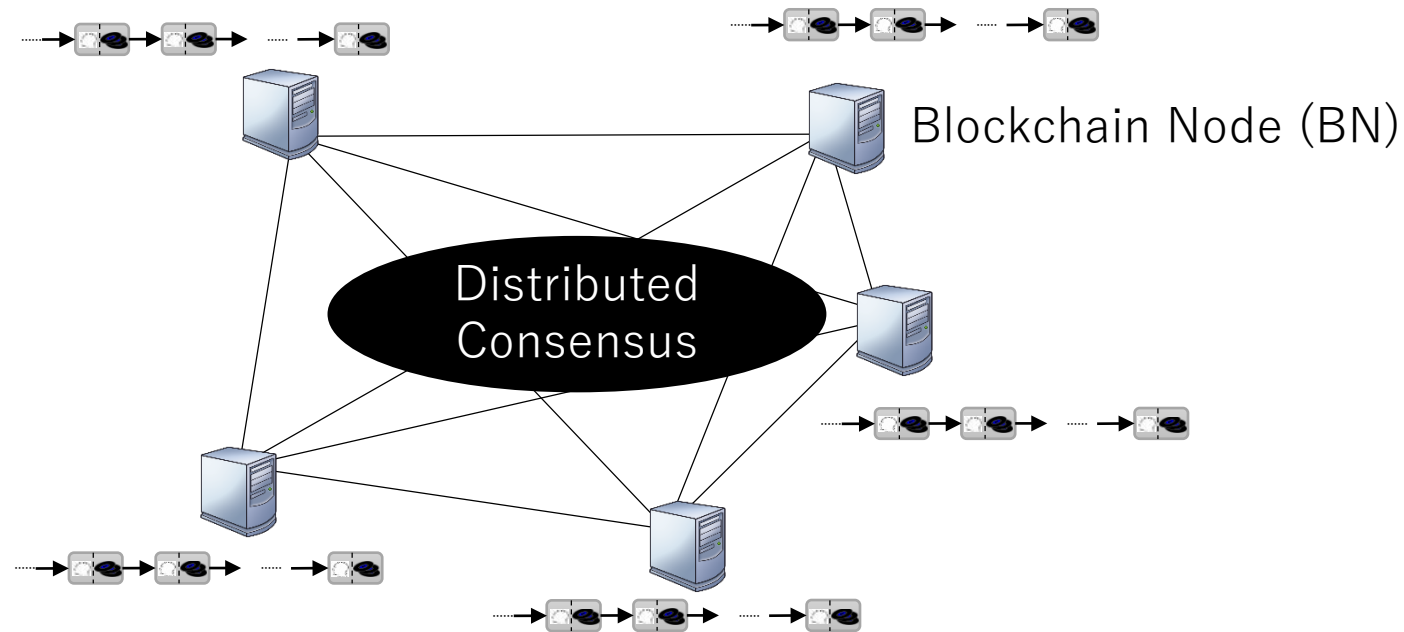
Network



Physical Space



Blockchain



- **Blockchain:** An open, distributed ledger and computing platform with transaction and block for recording, and smart contract for computing.
- **Transaction:** The signed data package that stores a message to be sent from an externally owned account.
- **Block:** A number of the transactions are aggregated into a block.
- **Smart contract:** the user-defined program deployed on the blockchain

Blockchain: Decentralized immutable data sharing with consensus

- **Immutability**

- **Data is immutable** because it must be modified in all nodes of the network to modify a piece of information, which is impossible

- **Decentralization**

- **No centralized authority** and all the information is shared among blockchain nodes

- **Enhanced security**

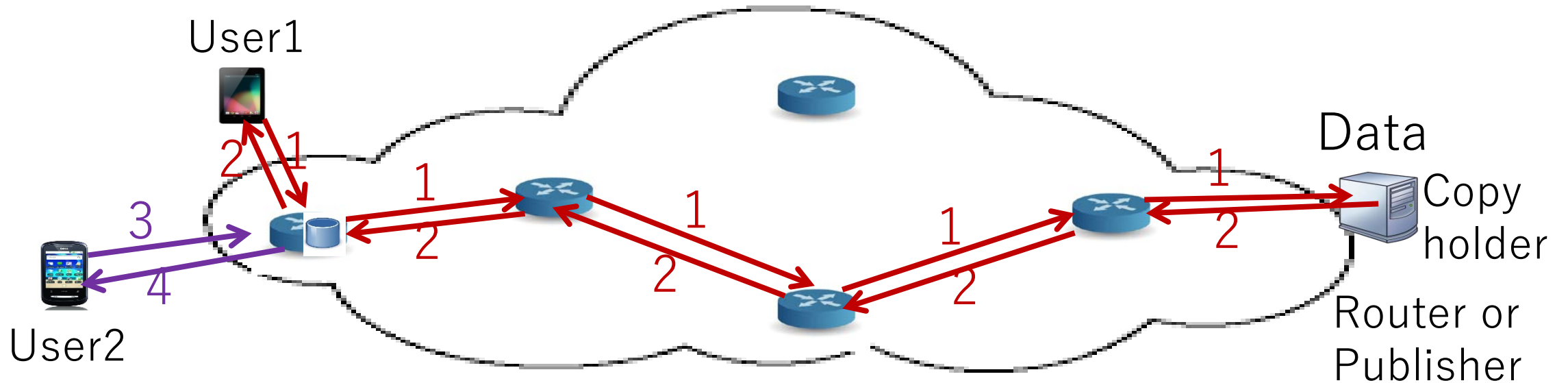
- Data can **be entirely secured** thanks to encryption with private and public keys.

- **Distributed ledger**

- Information in a blockchain is considered as **a shared file with consensus.**

Information-Centric Network (ICN)

Packet Types: Interest/Data

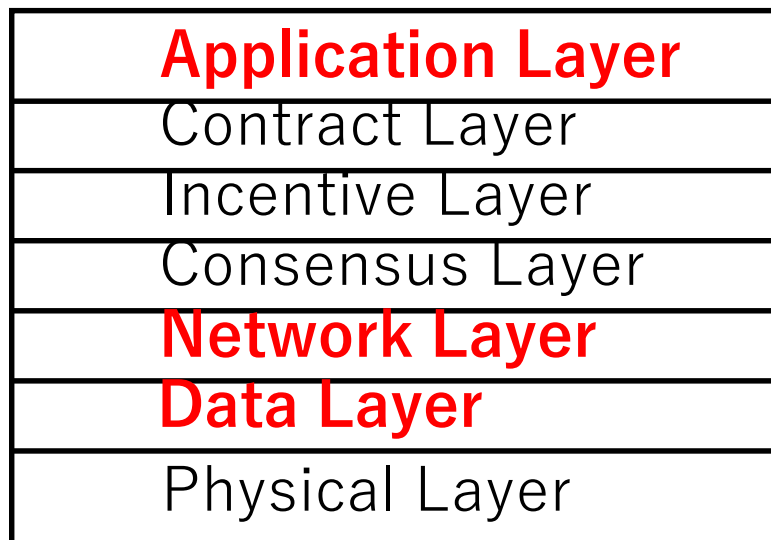


: ICN Router with caching capability

Typical ICN: Named Data Network (NDN)

Blockchain meets Information-Centric Network (ICN)

- **Blockchain:** Integration of technologies for decentralized immutable data sharing with consensus
 - Cryptographic technologies, Hash algorithm, Overlay network protocols, Consensus algorithm, Application service
- **ICN:** Ubiquitous (decentralized) named data caching for fast data retrieval



Blockchain



ICN

Blockchain meets Information-Centric Network (ICN)

- **Blockchain for ICN**

- Provide data immutability for ICN
- **Protect data with the consensus on the actions for data**, such as generation, publication, processing, usage, deletion
- Provide services for ICN, such as key management
- ...

- **Blockchain with ICN**

- **ICN for data (transactions and blocks) sharing**
- ICN for Blockchain nodes to interact with databases, such as InterPlanetary File System (IPFS)
- ...

- **Blockchain + ICN + ...**

- For Internet of Things services
- For security services
- ...

Blockchain for/with Information-Centric Network

- **Blockchain for ICN**

- Use blockchain to protect data lifecycle in ICN
- R. Li, and H. Asaeda, "A Blockchain-based Data Lifecycle Protection Framework for Information-Centric Network," IEEE Communications Magazine, vol. 57, no. 6, pp. 20-25, June 2019.

- **Blockchain with ICN**

- Use ICN to provide efficient transaction and block sharing for blockchain
- R. Li, and H. Asaeda, "DIBN: A Decentralized Information-Centric Blockchain Network," IEEE Global Communications Conference 2019 (GLOBECOM 2019), Hawaii, USA, Dec. 2019.

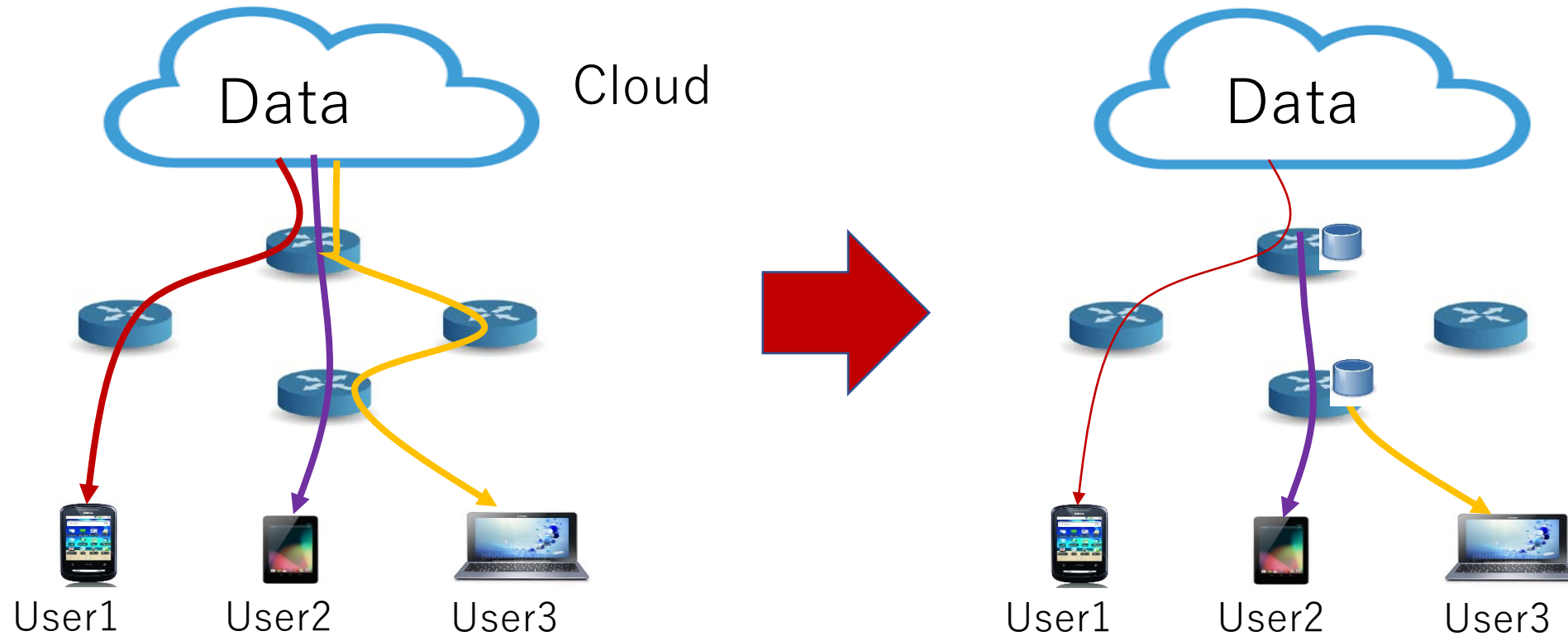
BDLP: A Blockchain-based Data Lifecycle Protection Framework for Information-Centric Network

IEEE Communications Magazine, 2019

Outline

- In-Network Data Caching
- Entities & Use Scenario
- Threats & Design Requirements
- Related Work
- Proposed BDLP
 - Key Idea
 - System Overview
 - Transactions, Smart Contracts
 - Data Retrieval Procedure
- Security Analysis
- Implementations

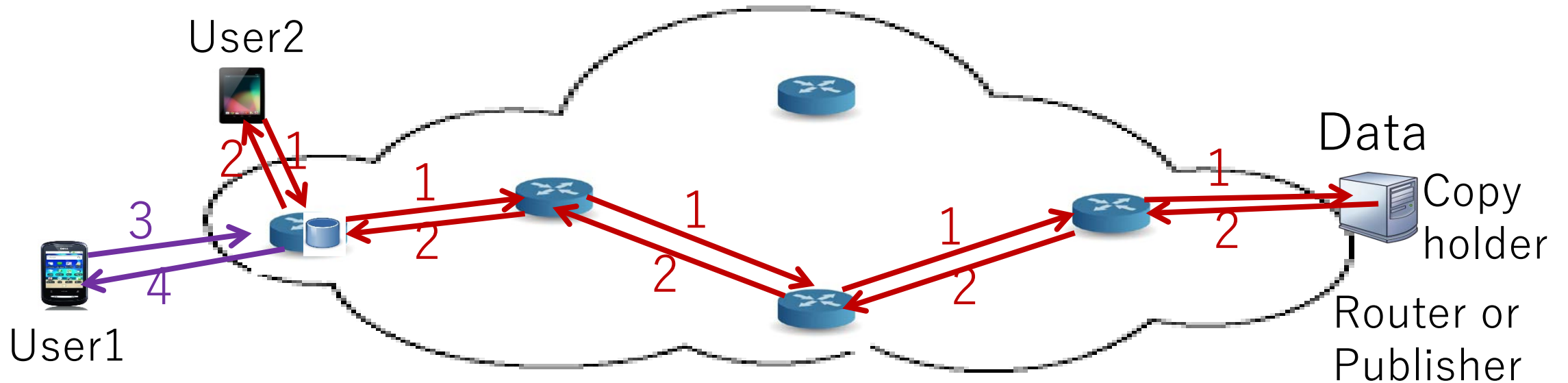
In-Network Data Caching



From centralized cloud/server-based data provision to distributed in-network caching-based data provision (Data can be cached at the intermediate nodes, such as routers, edge servers.)

Information-Centric Network (ICN)

Packet Types: Interest/Data



: ICN Router with caching capability

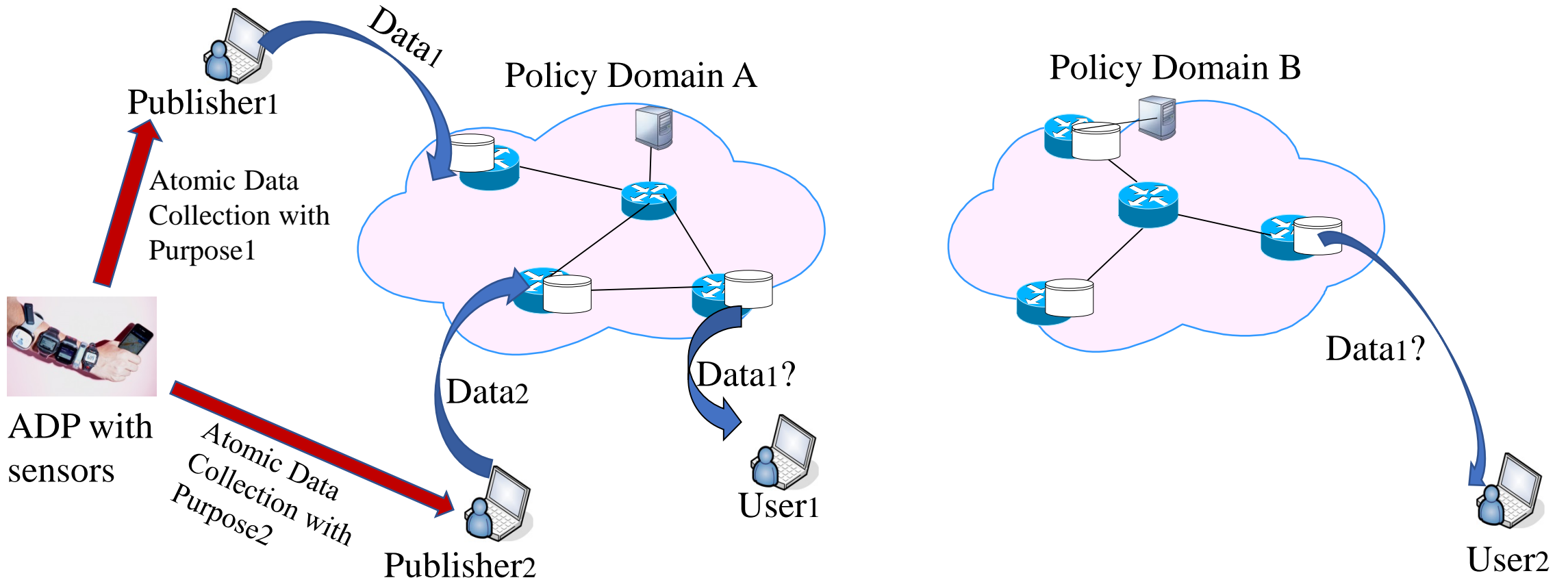
Typical ICN: Named Data Network (NDN)

Entities

- **Entities:**

- **Atomic Data Provider (ADP):** the entity that provides the original atomic data
- **Publisher:** the entity that collects original atomic data, forms and publishes data in network.
- **User:** the entity that retrieves data from network.
- **Blockchain Node (BN):** the entity in the blockchain system that works as the miner
- **Data Dam Blockchain Node (DDBN):** the entity in the blockchain that works as both a blockchain node and a data flow controller
- **Smart Contract (SC):** the user-defined program deployed on the blockchain to enable the automated processes to facilitate, execute and enforce certain restriction rules of negotiation and agreement.

Use Scenario



Use Scenario:

1. Publishers **collect atomic data** from atomic data providers (ADPs), such as sensors, actuators, and RFIDs.
2. With summarizing the atomic data, publishers **produce the data**.
3. **Data are cached** in ICN
4. Users **retrieve data** from the close copy holders (e.g., routers or publishers).

Data Lifecycle: Atomic data collection, data publication, caching, and retrieval.

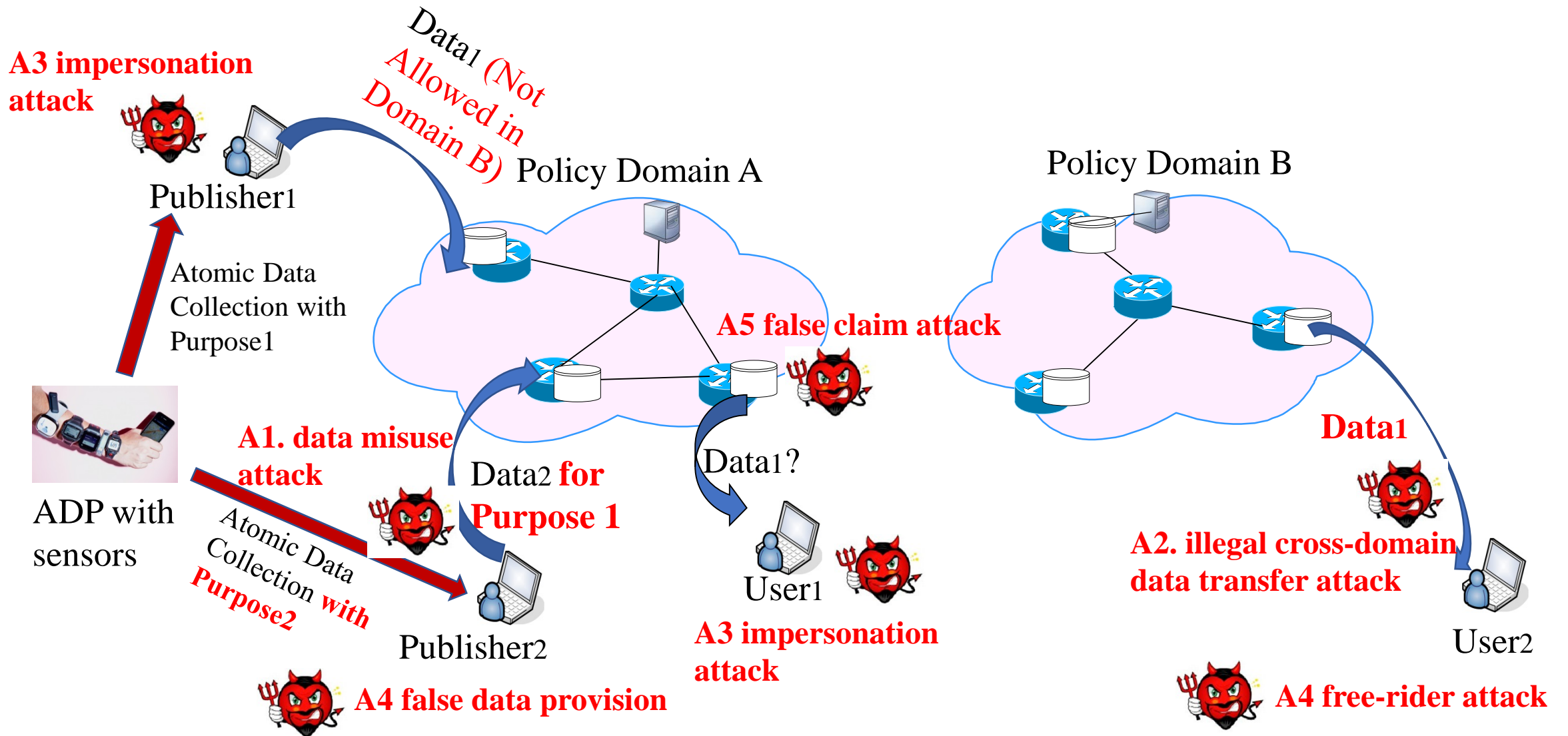
Data Protection: EU General Data protection regulation 2016/679

- <https://www.eugdpr.org/>
- 14 April 2016: Adoption by the European Parliament.
- 25 May 2018: Its provisions became directly applicable in all member states, two years after the regulations enter into force.
- This law defines the Data Protection regulations from the aspects of **accountability for EU citizens' data usage, the control on user rights on data movement, and security protections.**

Threats

- (A1) **data misuse attack**: the attacker deliberately **misuses the collected atomic data** with the purpose out of the one pre-negotiated with ADPs.
- (A2) **illegal cross-domain data transfer attack**: the attacker **illegally transfers data from one domain to another domain**, which is prohibited by the regulation.
- (A3) **impersonation attack**: the attacker **impersonates publisher to reply false data or impersonates user to request data**.
- (A4) **false data provision and free-rider attack**: the publisher **replying with illegal/out-of-expectation data** to the user or the user retrieves data without payment.
- (A5) **false claim attack**: the attacker makes **false claims on its operations on the data**.

Threats (Potential Attacks)



Design Requirements

- **Authentication**: to ensure that the **publisher and user can be proved to be as claimed**. (A3)
- **Accountability**: to **insure the execution of data-oriented operations** (i.e., who performs what operation on which data at which time). The main operations to be recorded include registration, atomic data collection, data publication, payment, and punishment. (A5)
- **Regulation compliance**: to guarantee that **the cached data is prohibited from being retrieved** in the policy domain(s), which is disallowed from accessing according to the regulations. For example, users outside the EU are not allowed to retrieve personal data in the EU. It also guarantees that **data is prohibited from being used for a purpose outside the pre-negotiation**. (A2) (A1)
- **Neutrality**: To guarantee that a **malicious publisher should be punished** and a **malicious user cannot retrieve data** without payment. (A4).

Related Work

- **ICN Security:**

- Researches on ICN security have **mostly focused on authentication, authorization, and access control**. However, most existing works can only protect one specific procedure in the data lifecycle, and still cannot protect the entire data lifecycle.
- **Lack of researches on neutrality** for ICN; both feedback on the retrieved data and assurance of the payment should be provided.

- **Blockchain:**

- **An open, distributed ledger and computing platform** with transaction and block for recording, and smart contract for computing.
- Blockchain can be classified into two categories: private blockchain, such as **Hyperledger Fabric**, and public blockchain, such as **bitcoin and Ethereum**.

- **Blockchain Applications:**

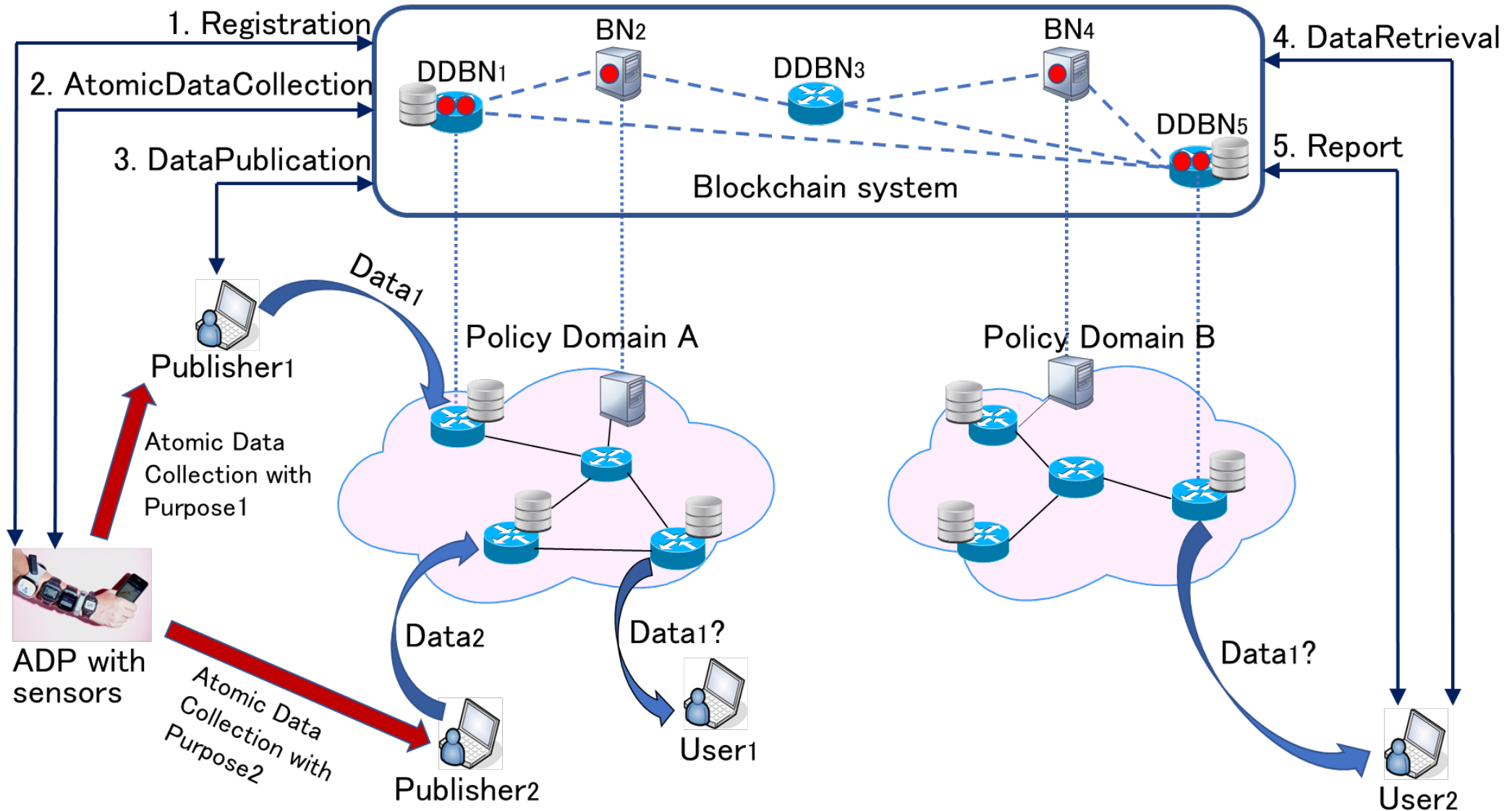
- The existing research on blockchain **addresses application areas differently from this work**, such as traditional security services for network.
- There is still **no existing work on data lifecycle protection in ICN with blockchain**.

Proposed BDLP: A Blockchain-based Data Lifecycle Protection Framework for ICN

- **Key Idea:**

- Exploit the transaction and smart contract in blockchain to provide a trustworthy and neutral environment for data provision and retrieval in ICN.
- Data dam blockchain node (DDBN) is designed to locally control registration and restrict data flow, besides the function of blockchain node.
- Five types of transactions (RegT, CollectT, PubT, PayT, and PunT) to achieve accountability
- Four types of smart contracts (PubSC, PaySC, AccSC, and RepSC) to achieve authentication, regulation compliance, and neutrality.

BDLP System Overview

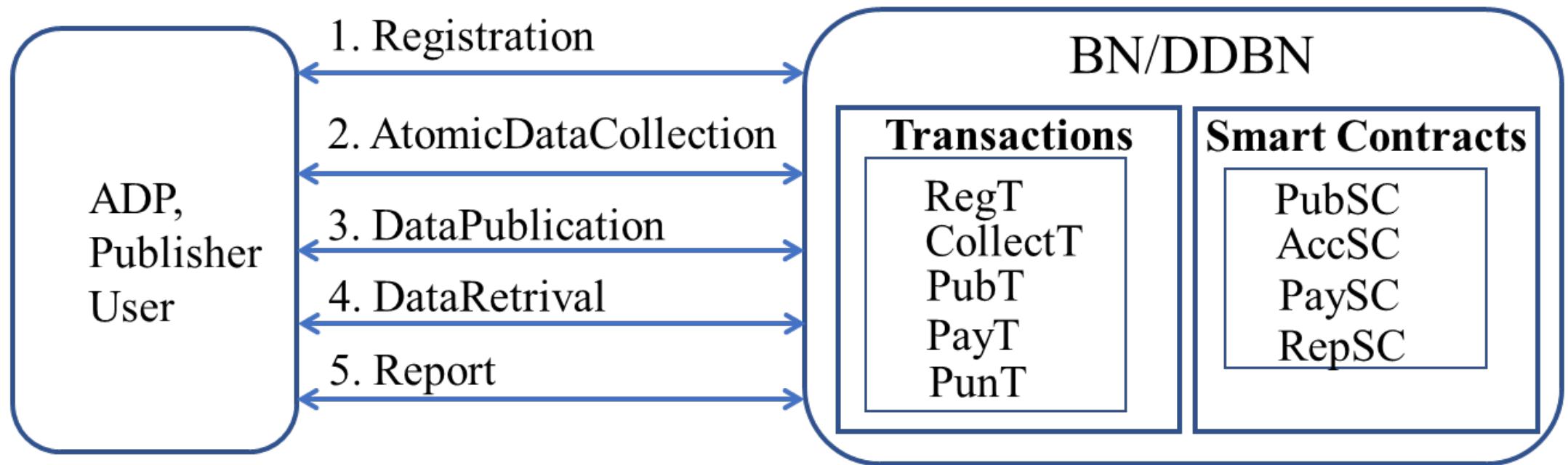


BN : Blockchain Node

DDBN : Data Dam Blockchain Node

● : Smart Contract

Procedures, Transactions, and Smart Contracts



ADP: Atomic Data Provider

BN: Blockchain Node

DDBN: Data Dam BN

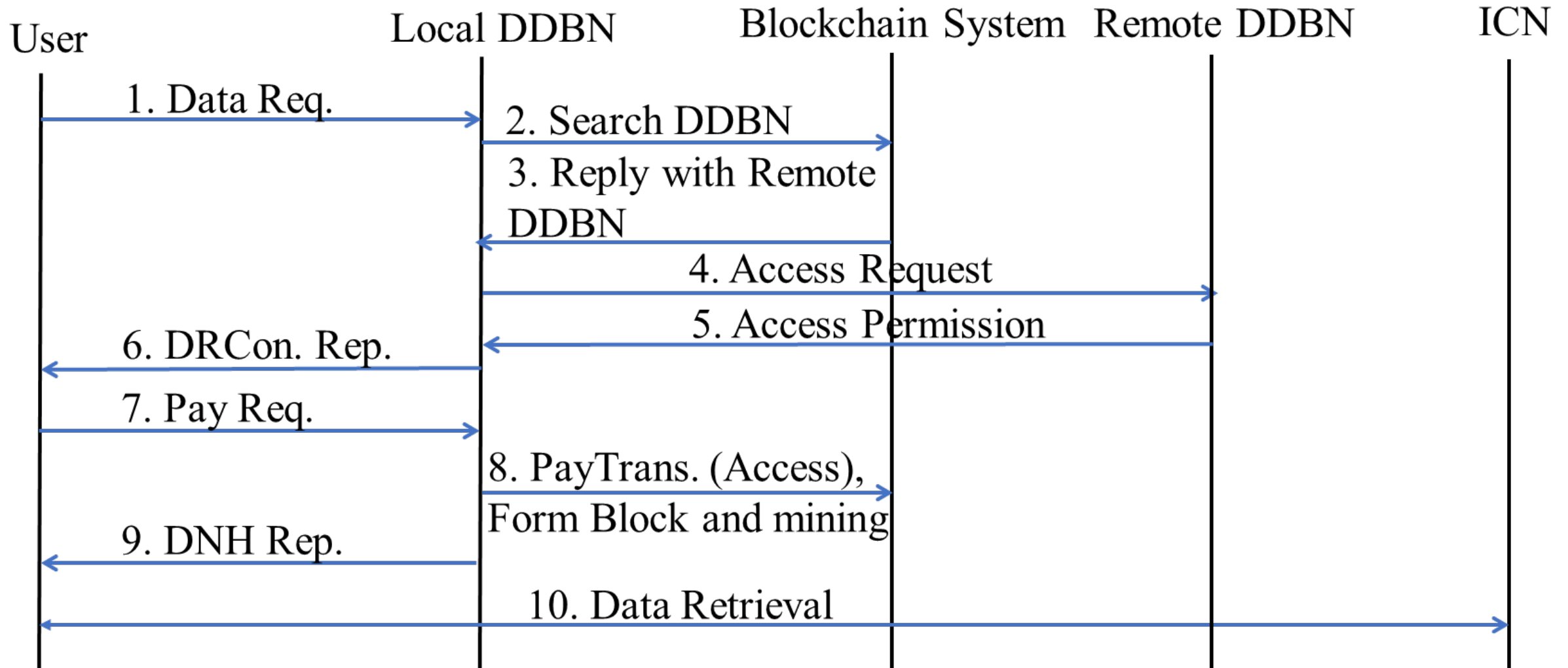
Transactions

- **RegT: Record the registration of node.**
 - $\langle H(\text{PubK}), \text{Registration time } t, \text{DDBN}, \text{Sig}_{\text{DDBN}} \rangle$.
- **CollectT: Record the atomic data collection behavior, approved by both the publisher and ADP.**
 - $\langle \text{ADP ID}, \text{Publisher ID}, \text{Atomic data name}, \text{Knowledge data name}, \text{Purpose}, \text{Sig}_{\text{ADP}}, \text{Sig}_{\text{Publisher}} \rangle$.
- **PubT: Record the data publication**
 - $\langle \text{Data name}, \text{Hash}(\text{Data}), \text{Domain name}, \text{DDBN ID}, \text{Price for retrieval}, \text{Purpose}, \text{Time } t, \text{Sig}_{\text{DDBN}} \rangle$.
- **PunT: Record punishment**
 - $\langle \text{Publisher ID}, \text{Data name}, \text{Contracted purpose}, \text{Evidence of current purpose}, \text{Rating}, \text{Time } t, \text{Sig}_{\text{ADP}} \rangle$
- **PayT: Record payment**
 - $\langle \text{payer}, \text{payee}, \text{number of digital coins}, \text{time}, \text{and signature of the payer} \rangle$

Smart Contracts

- **PubSC: process publication requests from publishers**
 - **Input:** \langle Publisher ID, Data name, Purpose, $H(\text{Data})$, Time t , A set of ADPs, Sig_P \rangle
 - **Process:** Check the purpose, regulation compliance, Generate Hash(dataname)
- **AccSC: process data retrieval requests from users**
 - **Input:** \langle User ID, Data name, Purpose, Time t , Sig_U \rangle
 - **Process:** Check the payment completion, regulation compliance, estimate payment amount
- **PaySC: process payment requests**
 - **Input:** \langle A, B, Coin, Time t , Evidence, Sig_A \rangle .
 - **Process:** Payment from A to B completion
- **RepSC: process a report**
 - **Input:** \langle A, B, Rating, Evidence, Original purpose, Sig_B \rangle
 - **Process:** Check the correctness of the report and perform the corresponding punishment

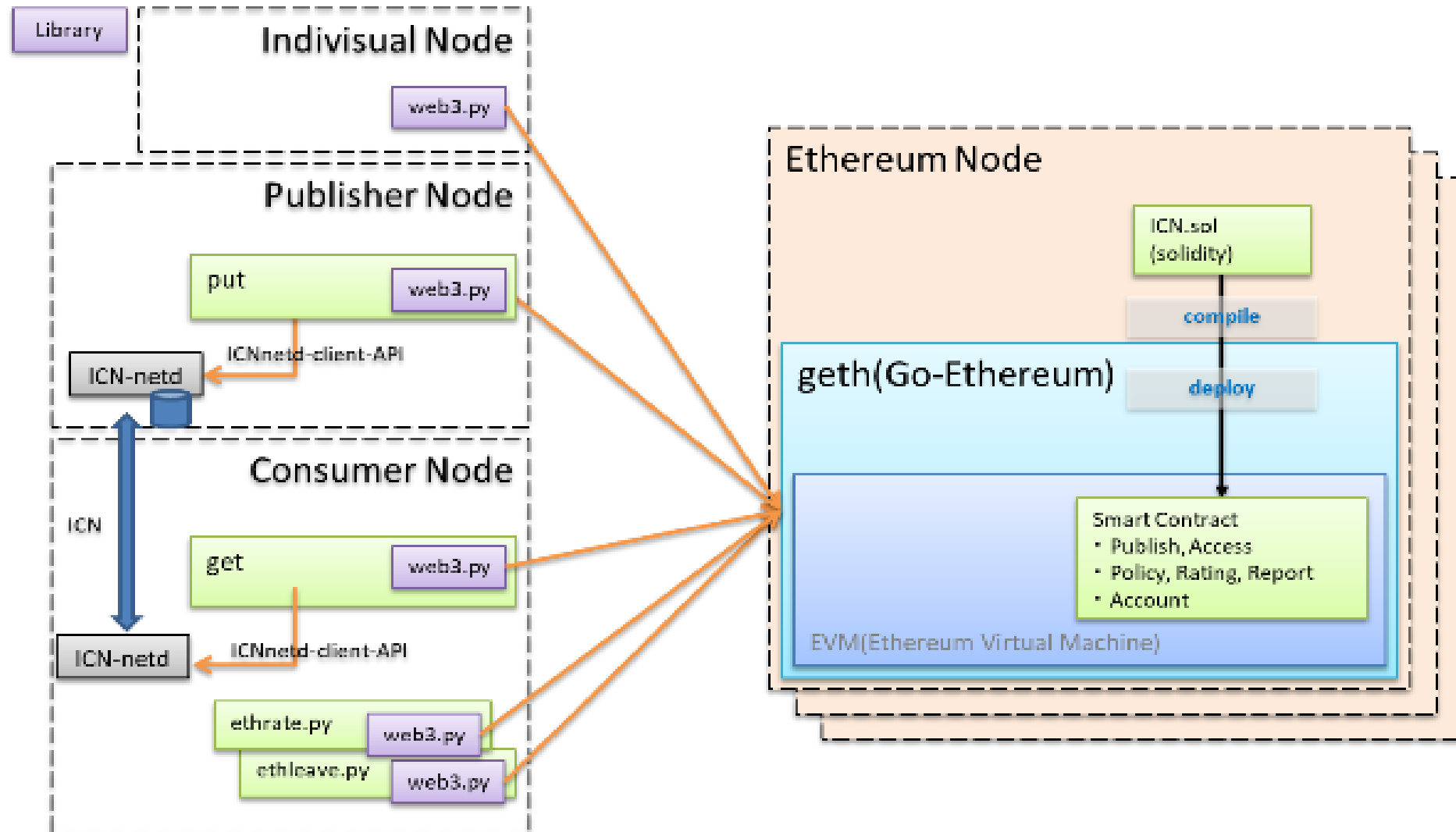
Data Retrieval Procedure in BDLP



Security Analysis

- **Authentication:** prevent **A3 impersonation attack** through registration procedure and RegT
- **Accountability:** inhibit **A5 false claim attack** through the immutability of transactions in blockchain.
 - The five types of transactions assure that no entity can deny what it did or assert something that it has not done.
- **Regulation compliance:** achieved through CollectT, PubT, PubSC, and AccSC
 - CollectT and PubT record the pre-negotiated purpose
 - PubSC and AccSC to enforce the regulation and policy and inhibit the **A2 illegal cross-domain data transfer** and **A1 data misuse attack**.
- **Neutrality:** provided without a trusted third party with PayT, PunT, PaySC, and RepSC
 - Enable publishers to obtain payment from users if providing correct data to inhibit **A4 free-rider problem**
 - Prevent **A4 false data provision** with report the misbehavior and punishment

Proof-of-Concept Implementations on Cefore



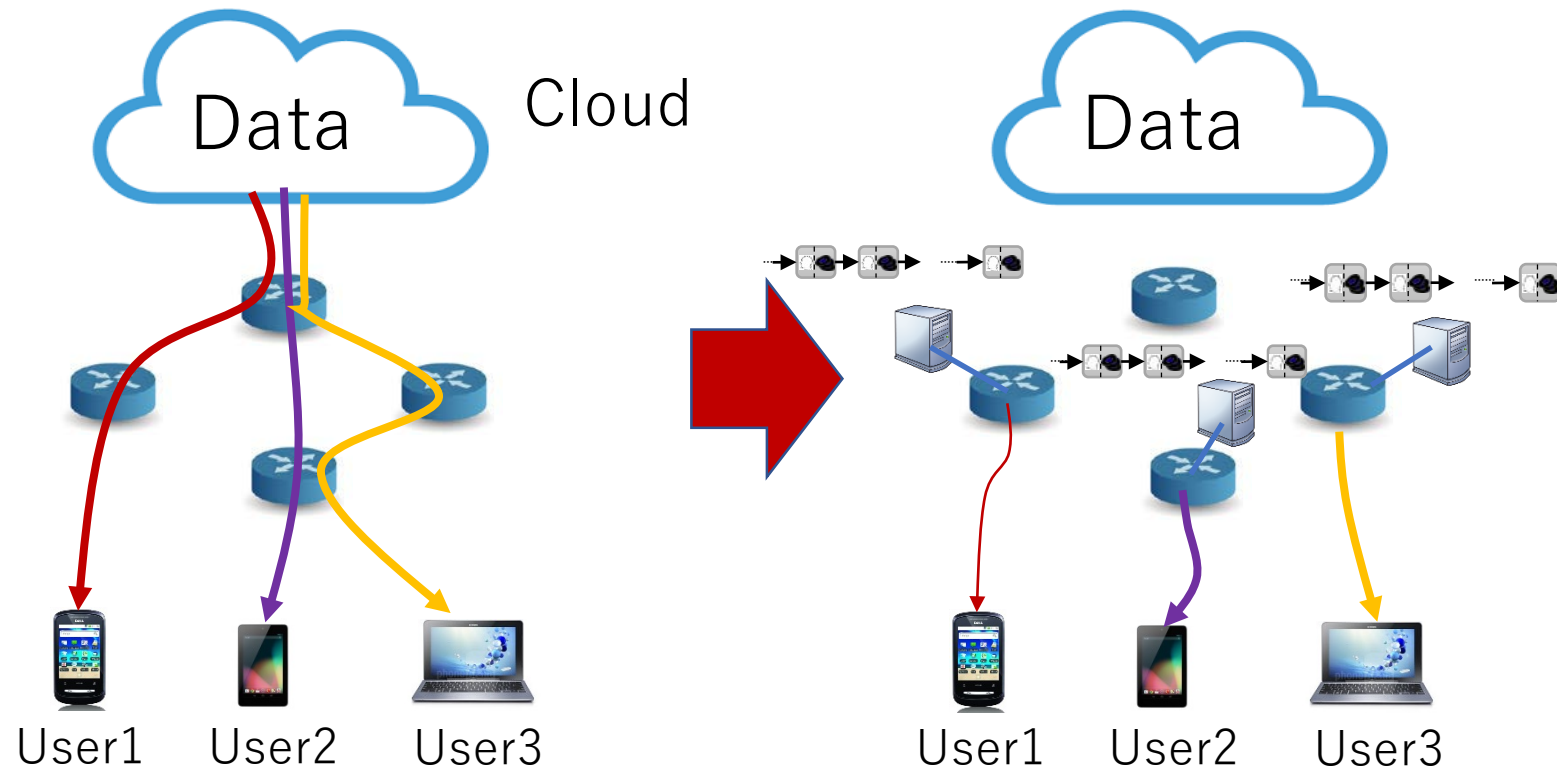
DIBN: A Decentralized Information-Centric Blockchain Network

Globecom 2019

Outline

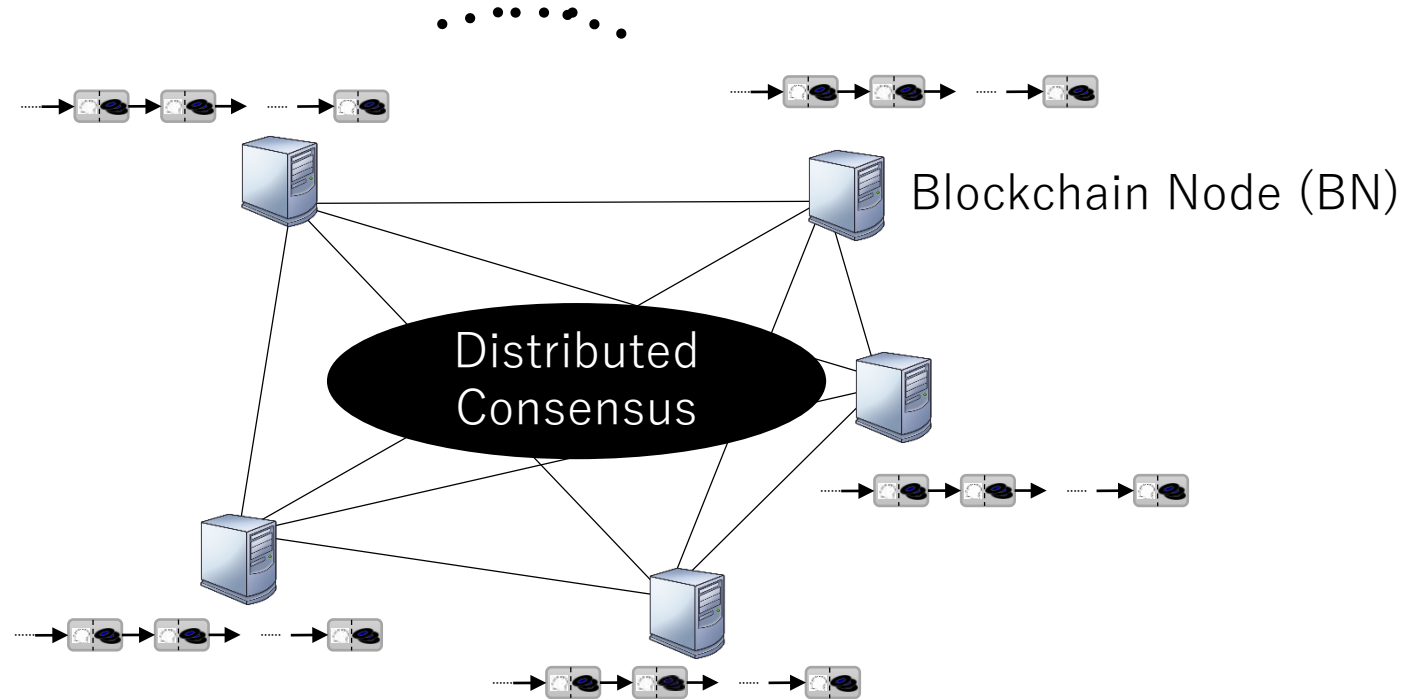
- Blockchain
- Blockchain Network Layer
- Related Work & Problem
- Proposed DIBN: Decentralized Information-Centric Blockchain Network
- Performance Analysis

Edge Data Security



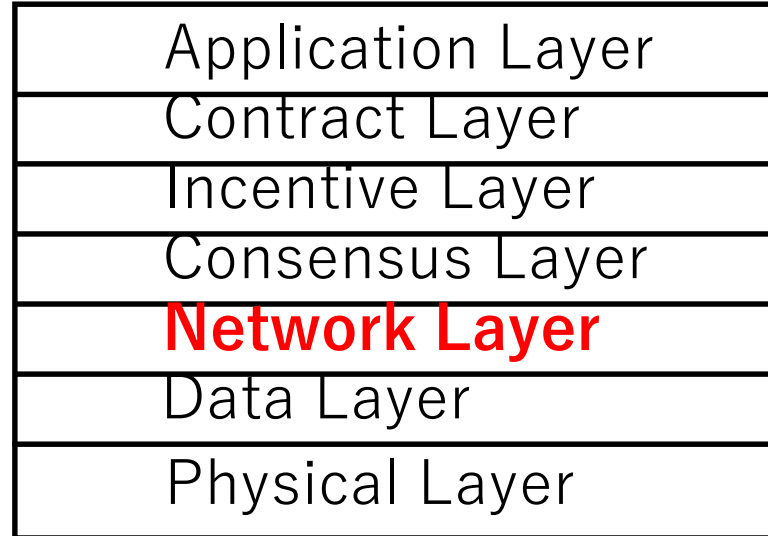
From centralized cloud/server-based data provision to distributed edge data provision

Blockchain



- **Blockchain:** An open, distributed ledger and computing platform with transaction and block for recording, and smart contract for computing.
- **Transaction:** The signed data package that stores a message to be sent from an externally owned account.
- **Block:** A number of the transactions are aggregated into a block.

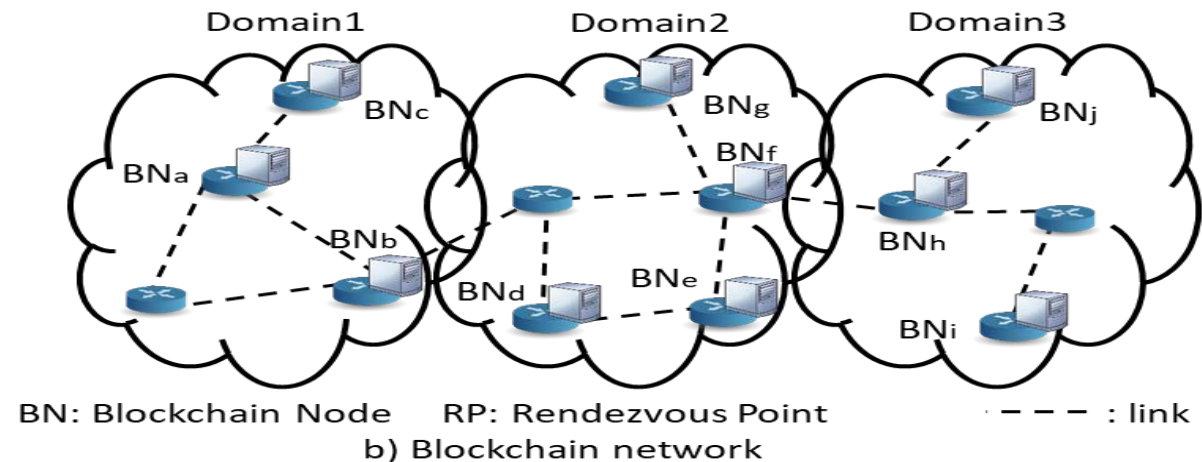
Blockchain Network Layer



- Dissemination of transactions and blocks to reach the consensus
 - Attachment strategy: Neighbor discovery and selection
 - Communication strategy: Data dissemination

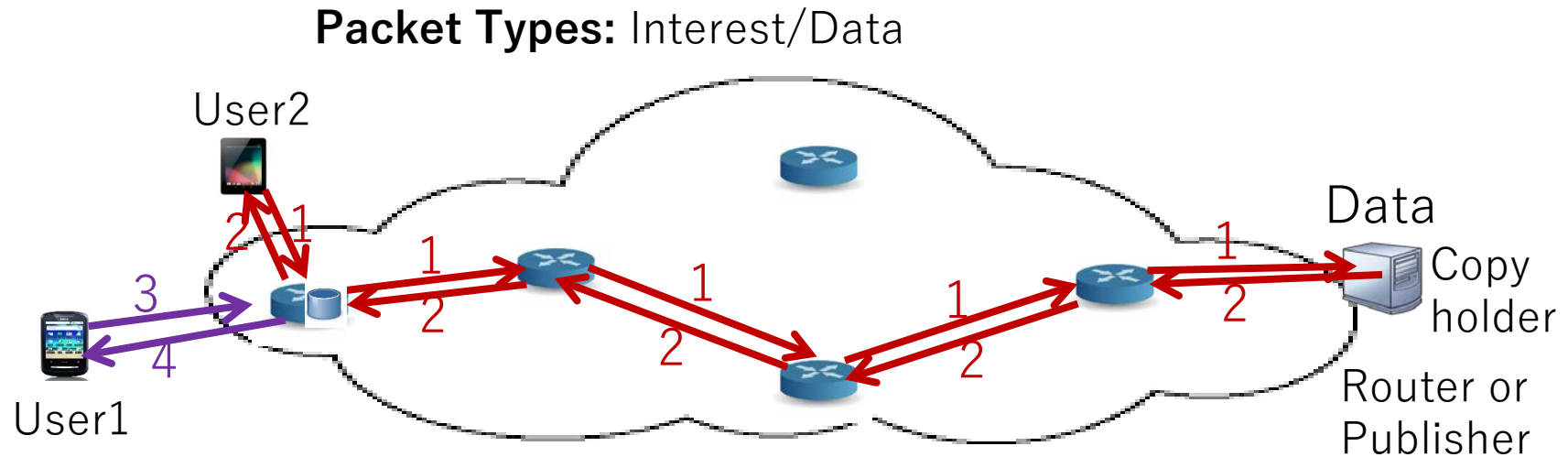
Related Work & Problems


- A blockchain network requires the **any-to-all data dissemination** mechanism
 - Any BN can broadcast the transactions or blocks to all other BNs.
- **Application layer multicasting (ALM)** protocols are used based on the peer-to-peer (P2P) overlay network technology, especially the unstructured overlay.
 - **Mesh-based approach**, such as Narada
 - **Tree-based**, such as NICE



- **Problems:**
 - Its performance suffers from the well-known **mismatch problem between the overlay traffic flows and the underlying physical network topology**, causing a large volume of redundant traffic.
 - A special node, **rendezvous point**, is necessary for the group management.

Information-Centric Network (ICN)



: ICN Router with caching capability
Typical ICN: Named Data Network (NDN)

Proposed DIBN: Decentralized Information-Centric Blockchain Network

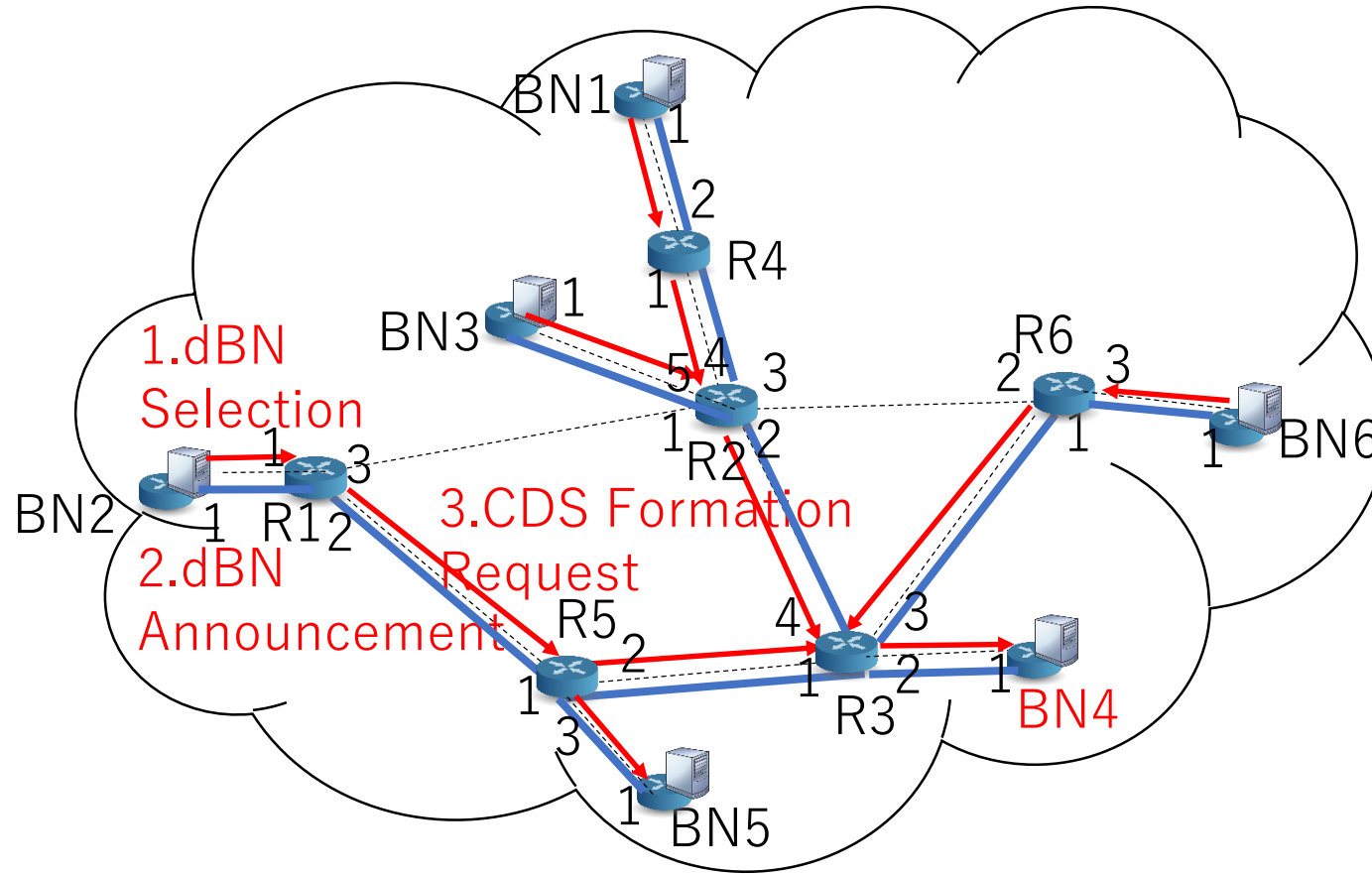
- **Key Idea:**

- To release such traffic concentration, transactions and blocks are named and organized based on categories, and category communication structures (CDSes) are constructed as the communication primitive in a distributed manner.
- A bi-directional data dissemination tree is constructed for each category with inclusion of all Blockchain Nodes (BNs), where one High-Performance Blockchain Node (HPBN) is selected to be the root, dBN, for this tree with considerations on load balancing.
- Over the CDSs, any-to-all communication strategy is realized, where any BN can efficiently disseminate the data to all other BNs aligning the traffic with the underlying network.

Naming

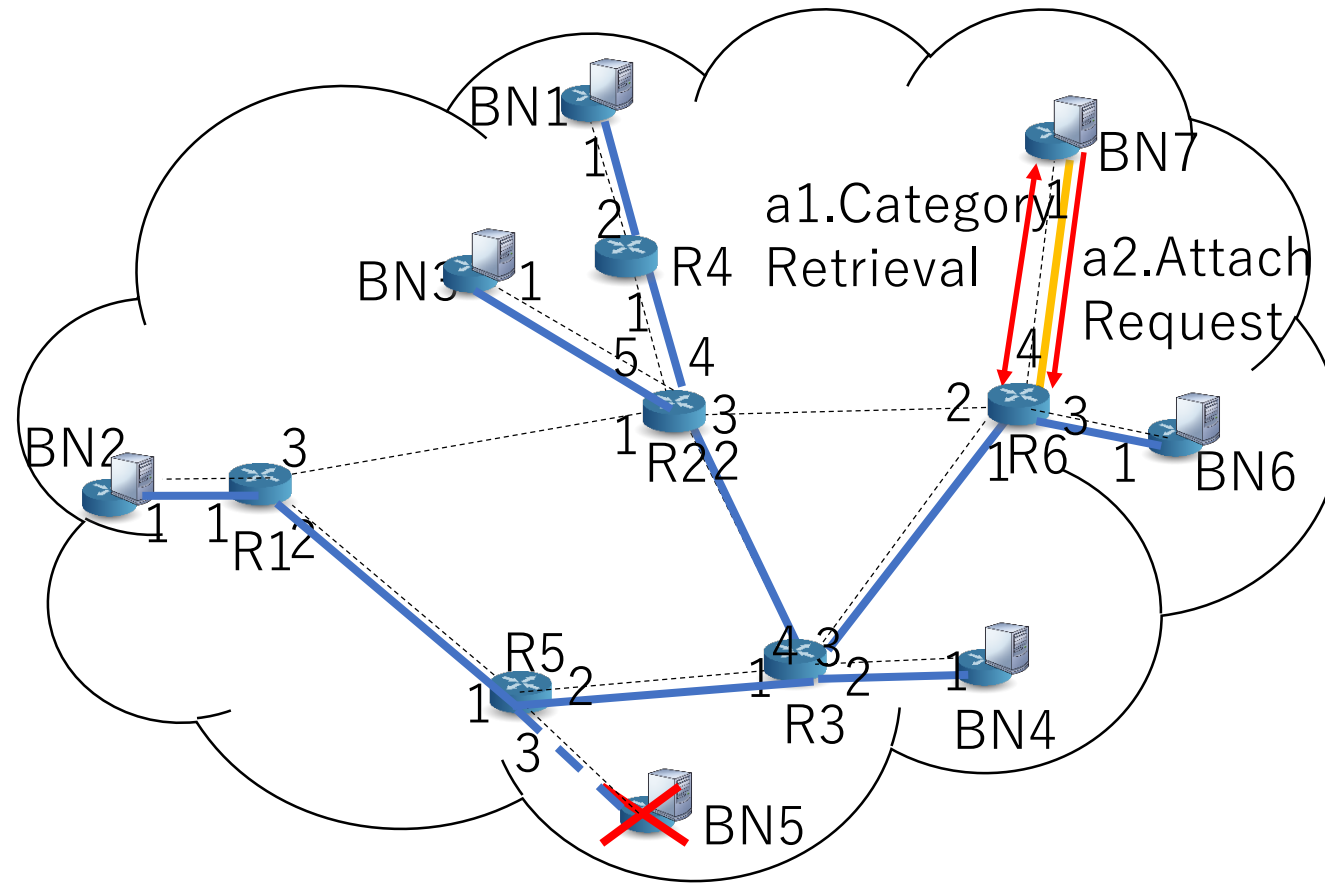
- **For transactions**, a category is afforded **with a semantically meaningful name**, such as rating on data, to share a set of specific evidences. <Category Name|Transaction Name| DIBN domain Name>
- **For blocks**, a category can be simply **labeled with a meaningless name** to decentralize the traffic with considerations on load balancing. <Category Name| Block Name| DIBN domain Name>

Category Dissemination Structure (CDS) Formation



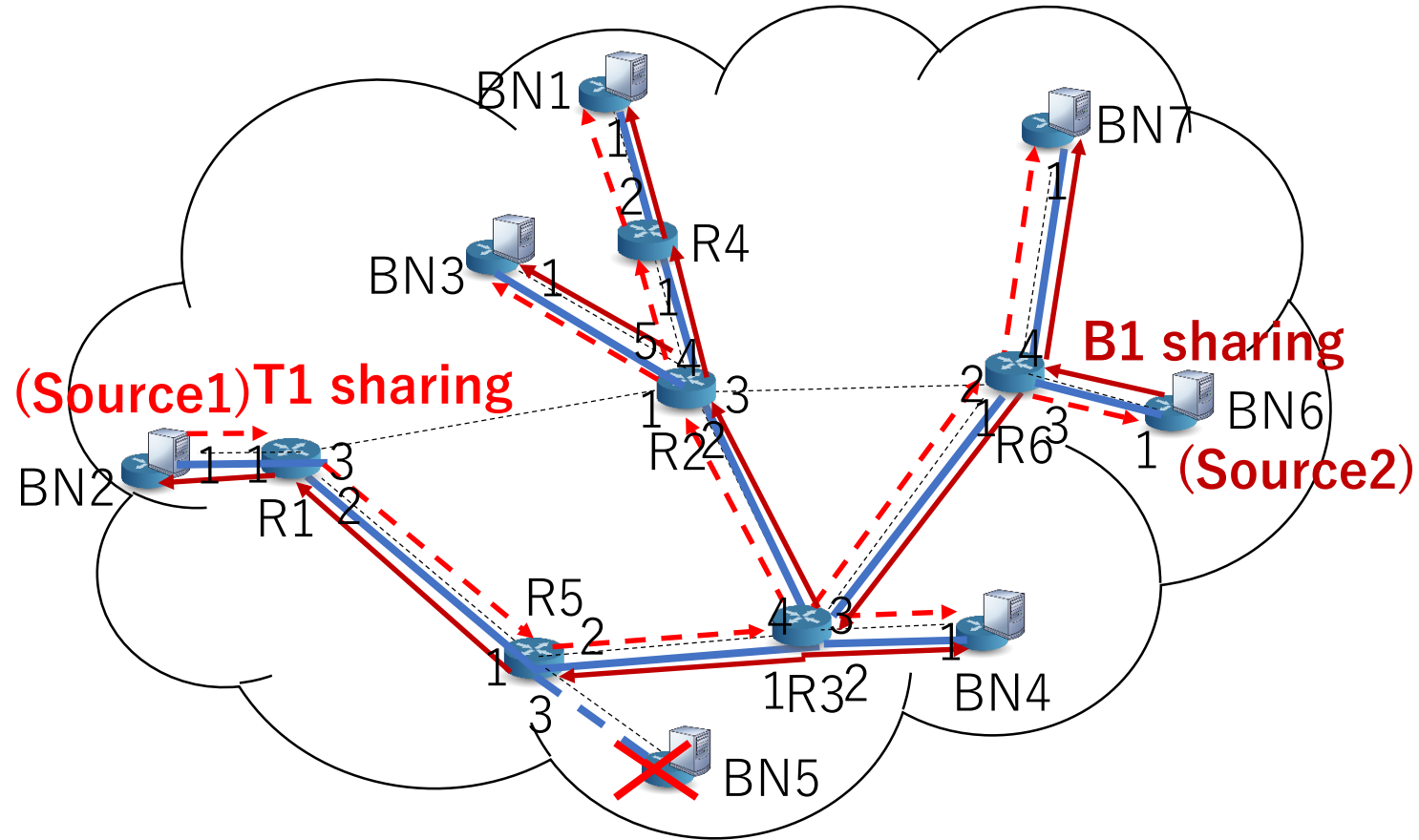
BN2 initiates one category and creates the communication structure for it

Category Dissemination Structure (CDS) Formation: Attachment Strategy



BN7 intends to attach to the blockchain network

Any-to-All Communication Strategy



BN2 announces the transactions T1

BN6 announces the block B1

Performance Analysis

- With the existing ALM in blockchain, the data, transactions and blocks, are disseminated over the unstructured P2P overlay network.
- With DIBN, they are disseminated with named transactions and blocks and communication structure.
- We compare the proposed DIBN with the typical ALMs, Narada and NICE.
- Metric:
 - Average path length (L): the average number of physical hops for data to be sent from the source BN to a given BN, which is the typical metric to examine the efficiency of ALM

Performance Analysis

- Average path length for the proposed DIBN

$$\begin{aligned}L(BN_s, BN_d) &\sim O(a \times b \times \log(K)) \quad (1 \leq a \leq 2) \\ &\sim O(\log(K))\end{aligned}$$

- Average path length for Narada

$$L_{Narada}(BN_s, BN_d) \sim O(M \times \log(K))$$

- Average path length for NICE

$$L_{NICE}(BN_s, BN_d) \sim O(\log(M) \times \log(K))$$

K: Total number of nodes in the network

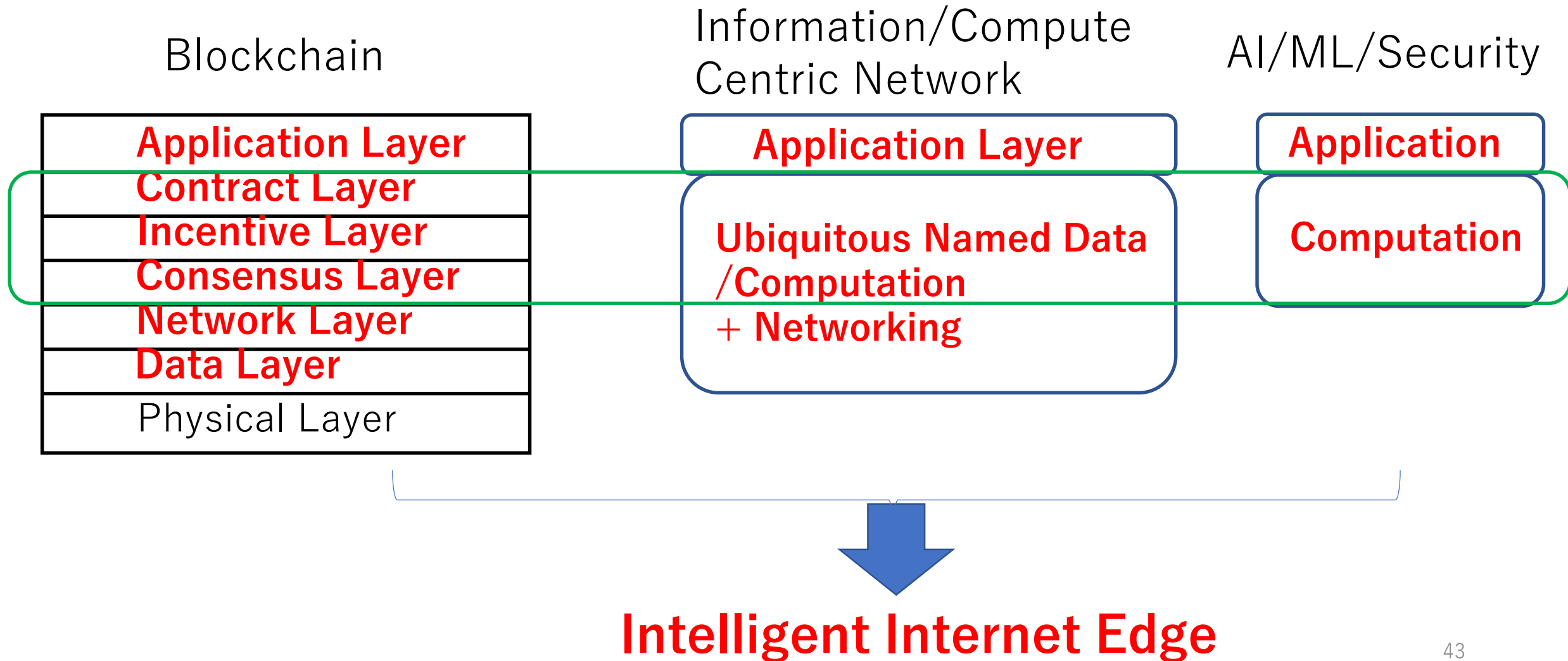
M: Total number of blockchain nodes

Conclusions

- We proposed **BDLP, A Blockchain-based Data Lifecycle Protection Framework for ICN.**
- We proposed a **DIBN, A Decentralized Information-Centric Blockchain Network.**

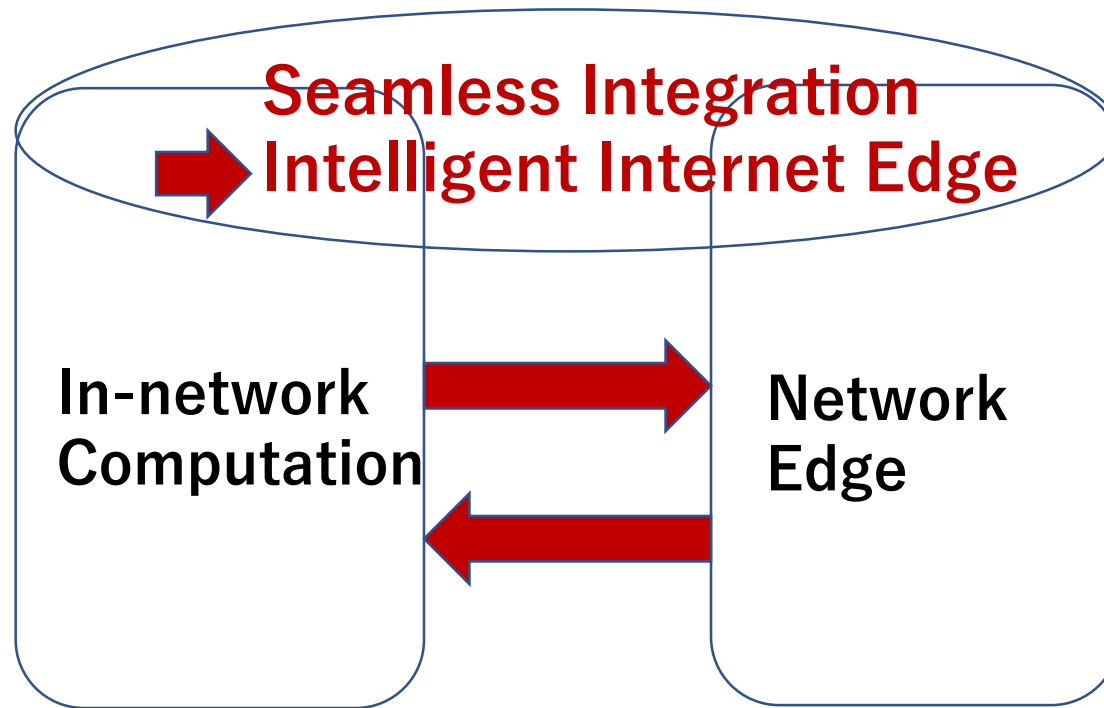
Future Work

- ICN-> Compute-Centric Network



IEEE SIG on Intelligent Internet Edge

- <https://github.com/IEEEITCIE/Home/wiki>



Thank you!