Concept of Decentralized Access Control for Open Network of Autonomous Data Providers

Łukasz Opioła | Michał Wrzeszcz | Łukasz Dutka | Renata Słota | Jacek Kitowski

ACK CYFRONET AGH, Kraków, Poland

AGH University of Science and Technology, Kraków, Poland

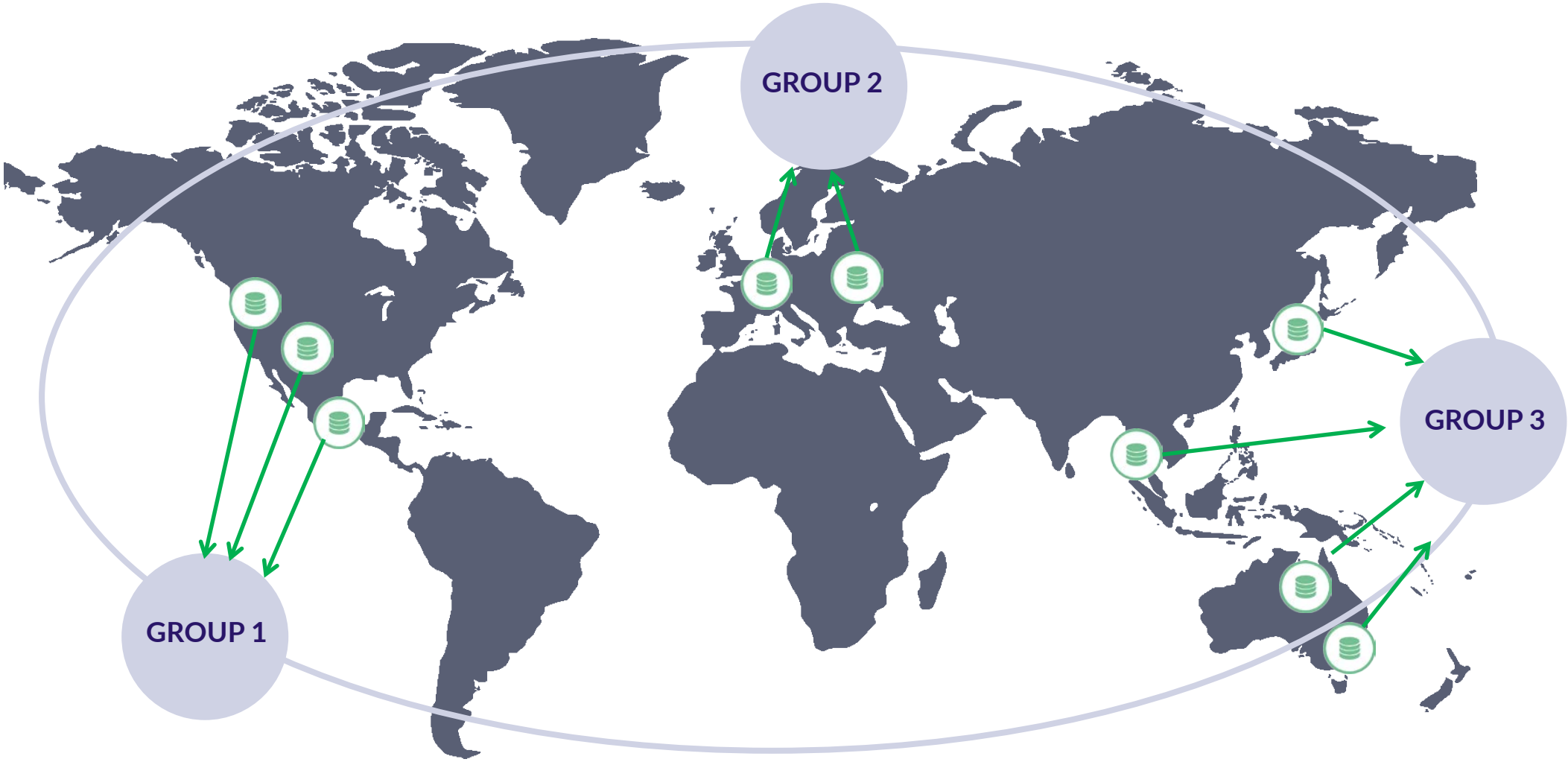
Faculty of Computer Science, Electronics and Telecommunications

Department of Computer Science

Agenda

1. Data access requirements of modern science
2. Challenges of global data access
3. Our vision of global data access
4. **Proposed concept of decentralized data access control**
5. Conclusions

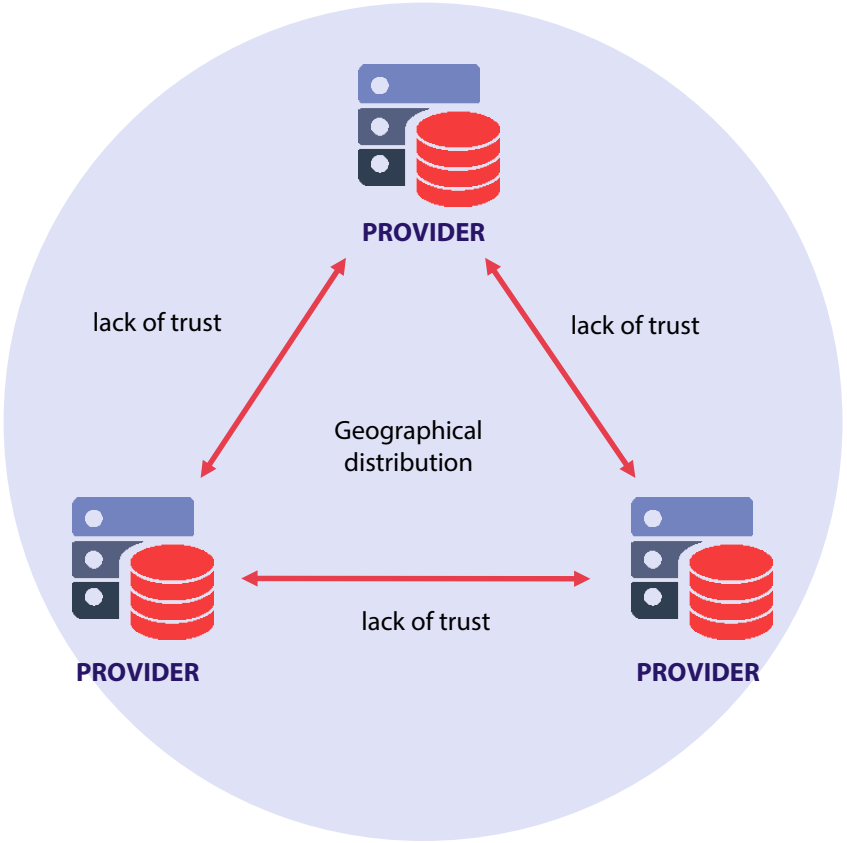
DATA ACCESS REQUIREMENTS OF MODERN SCIENCE



GLOBAL DATA ACCESS

- Effective access to large data sets
  - Data sets are distributed among many institutions
  - Data is stored on heterogeneous storage systems
  - Complete download before processing is not always possible due to sheer data size
- Transparent data access
- Cross-border collaboration
- Secure and easy data sharing

CHALLENGES OF GLOBAL DATA ACCESS



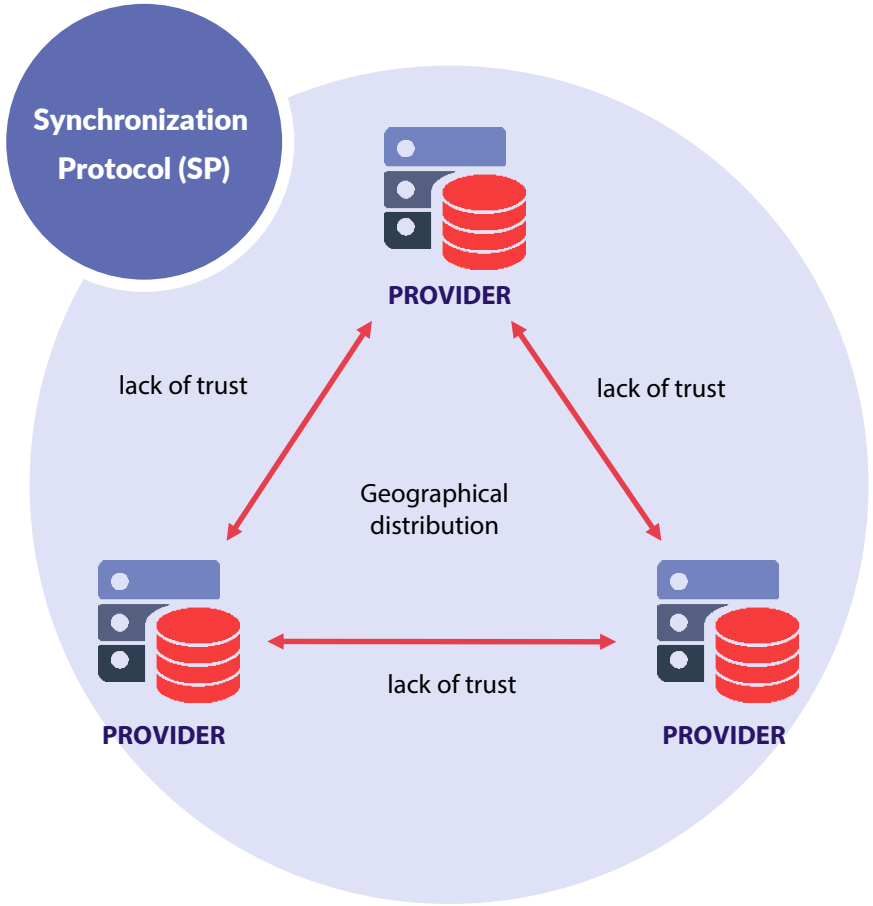
Goal: a network of cooperating providers

- Autonomy must be retained
- Inherent lack of trust
- Openness and discovery on demand

Our focus: decentralized data access control

- Metadata exchange between providers
- Metadata consistency across the system
- Secure metadata exchange
- Authority delegation

CHALLENGES OF GLOBAL DATA ACCESS



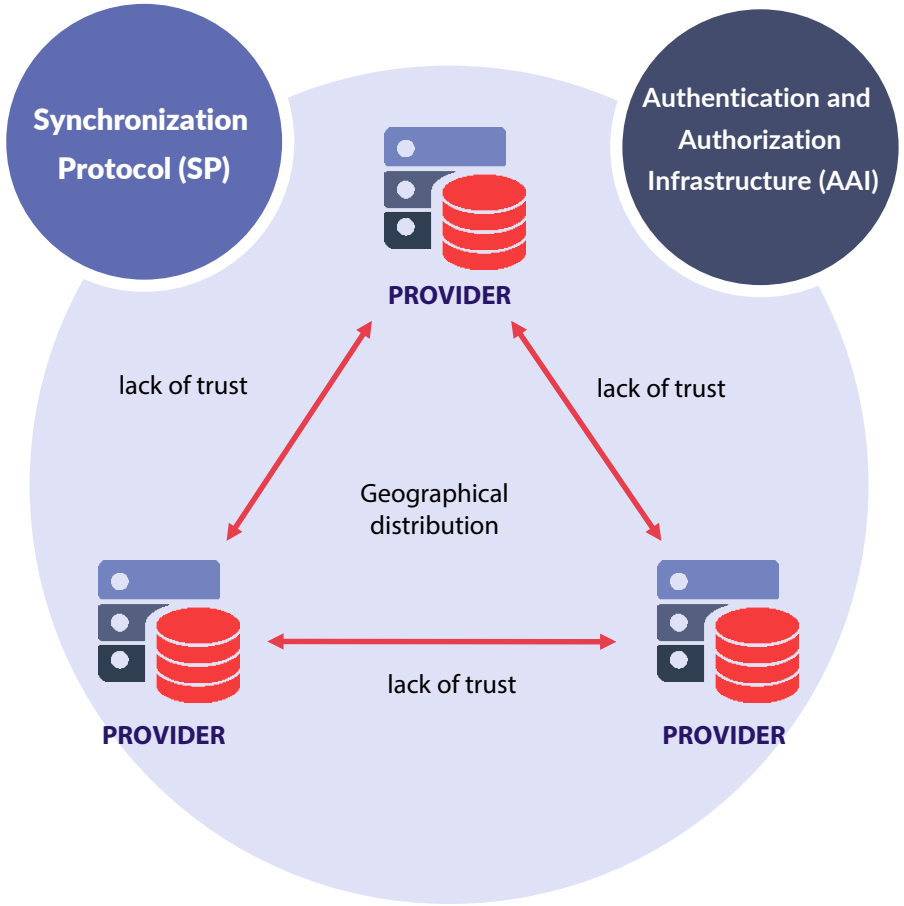
Goal: a network of cooperating providers

- Autonomy must be retained
- Inherent lack of trust
- Openness and discovery on demand

Our focus: decentralized data access control

- Metadata exchange between providers
- Metadata consistency across the system
- Secure metadata exchange
- Authority delegation

CHALLENGES OF GLOBAL DATA ACCESS



Goal: a network of cooperating providers

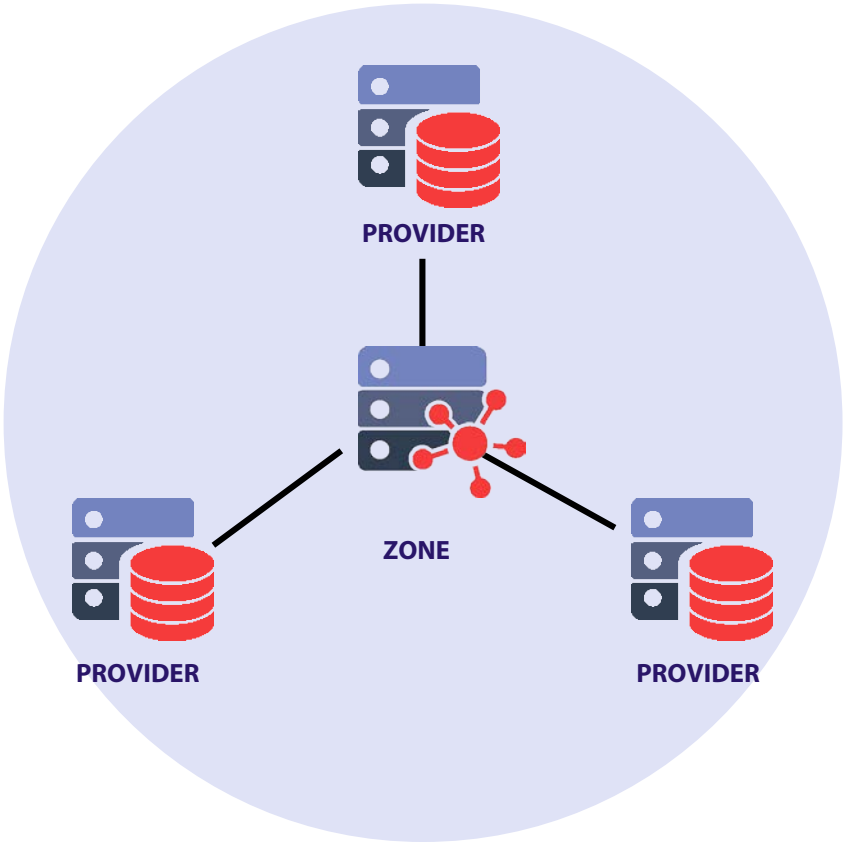
- Autonomy must be retained
- Inherent lack of trust
- Openness and discovery on demand

Our focus: decentralized data access control

- Metadata exchange between providers
- Metadata consistency across the system
- Secure metadata exchange
- Authority delegation

PROPOSED CONCEPT OF DECENTRALIZED DATA ACCESS CONTROL

1) Introducing zones



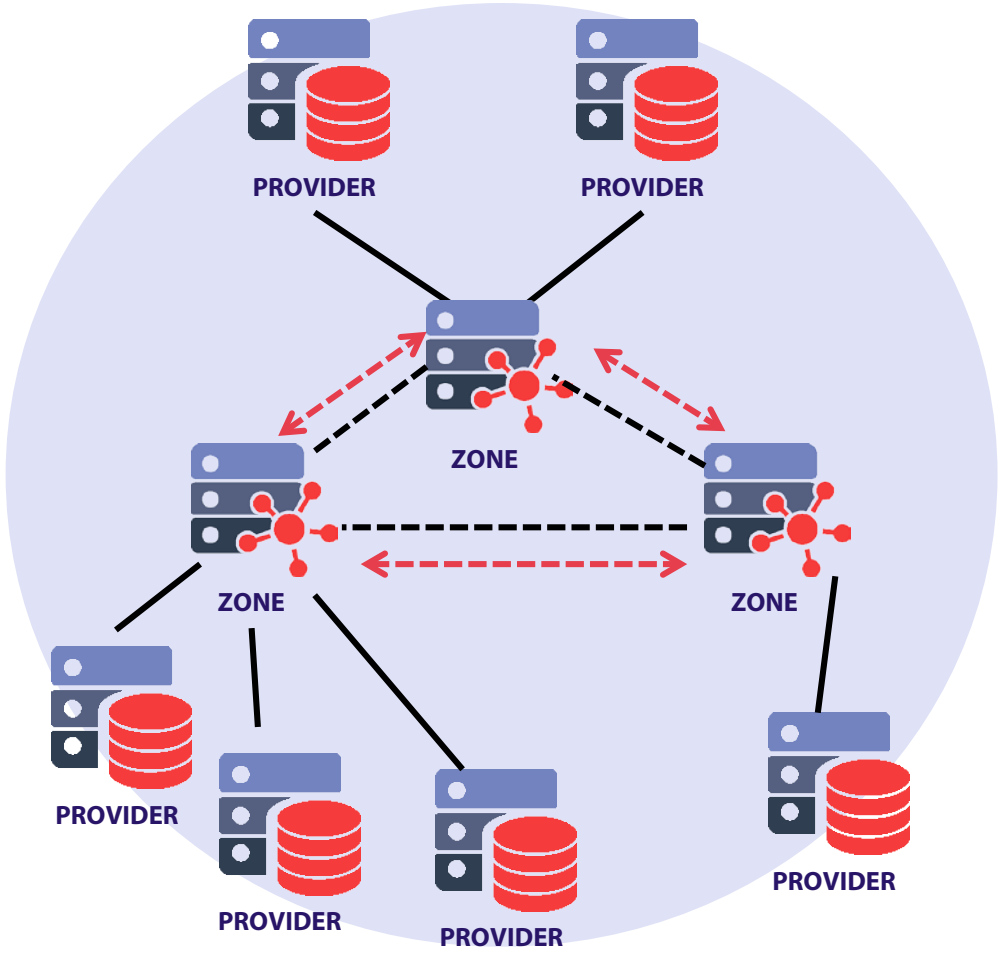
The purpose of a zone:

- Organizes providers into a cooperating group
- Serves as AuthN & AuthZ center (AAI)
- Holds the metadata of users and data sets
- Provides metadata synchronization API (SP)
- Entry point to the system

Note: every provider trusts its chosen zone

PROPOSED CONCEPT OF DECENTRALIZED DATA ACCESS CONTROL

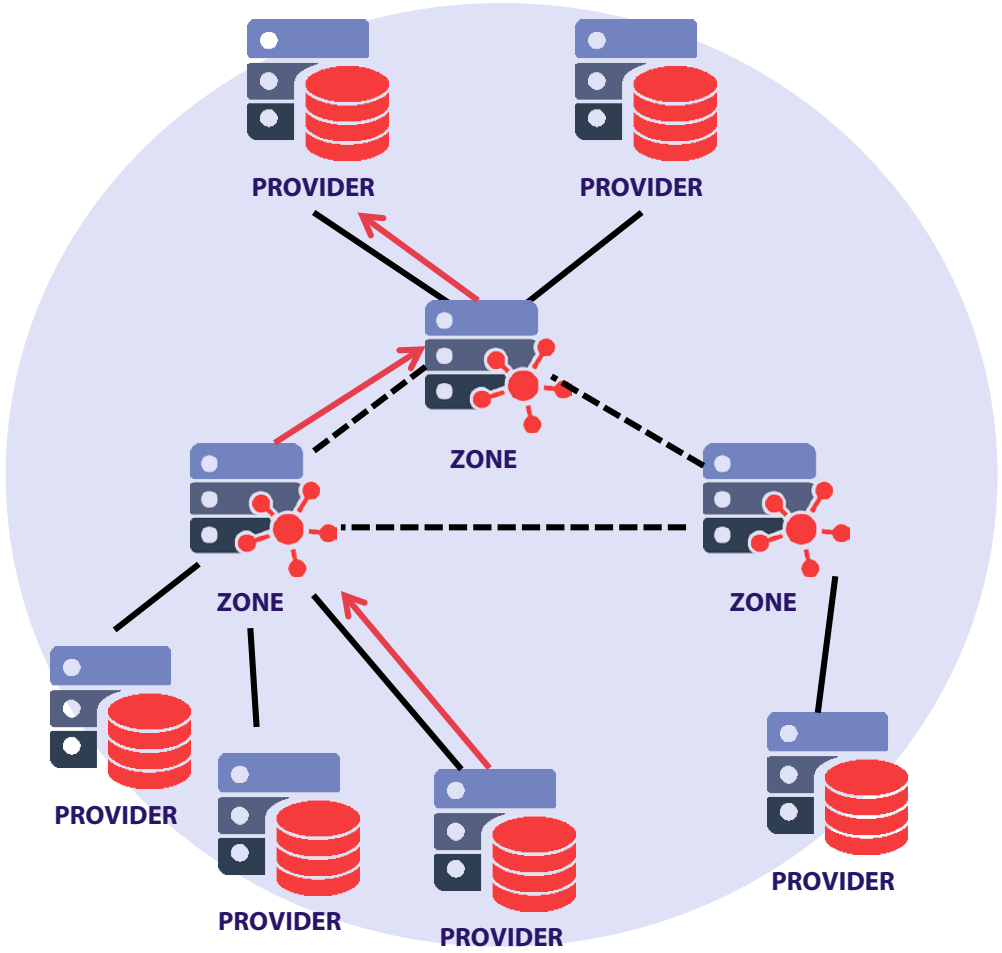
2) Cross-zone cooperation



- Zones can cooperate with each other
- Discovery on demand as required
- The network is open and scalable
- Each zone is an authoritative AAI and SP center for its providers
- Decentralization – P2P network of zones

PROPOSED CONCEPT OF DECENTRALIZED DATA ACCESS CONTROL

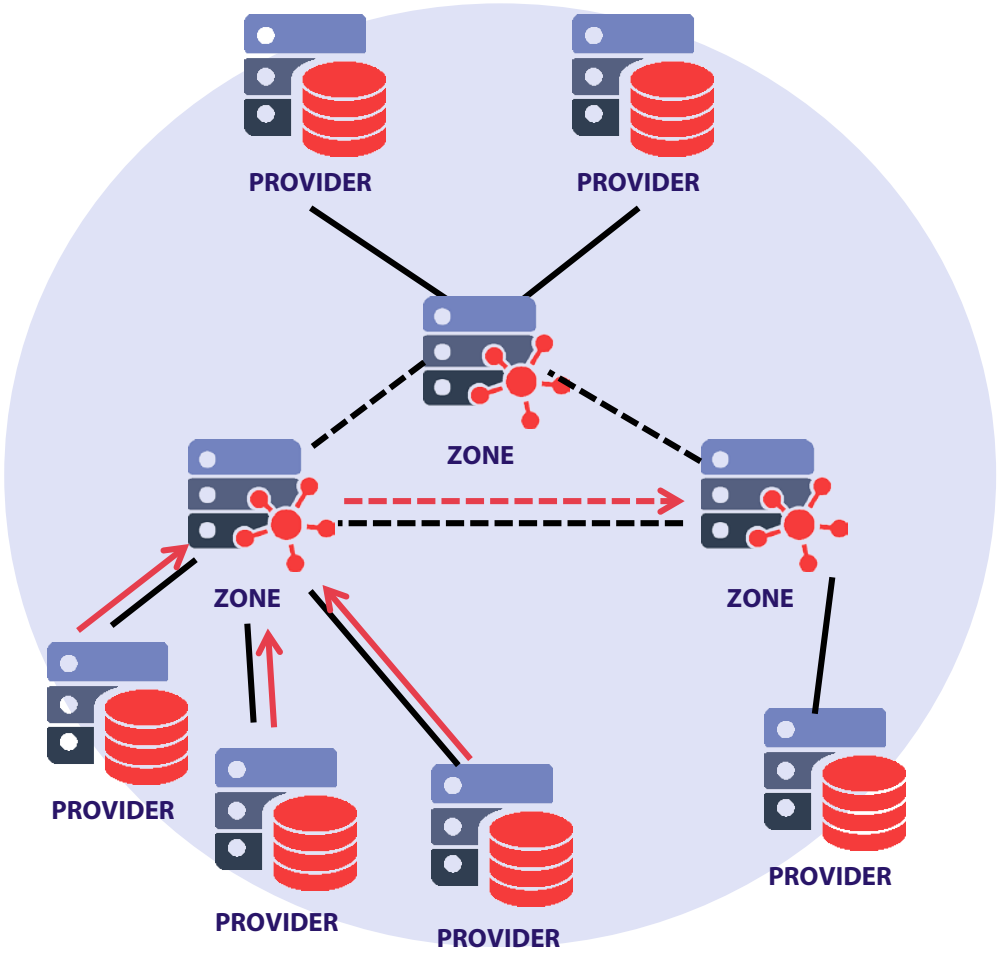
3) Authentication and Authorization Infrastructure (AAI)



- Based on macaroon tokens (by Google)
- Each zone verifies tokens that it issued
- Tokens can be passed around for delegation and confined as required
- Providers rely on their zone to securely connect to others
- Users can sign in to other zones using OpenID Connect (single, global account)

PROPOSED CONCEPT OF DECENTRALIZED DATA ACCESS CONTROL

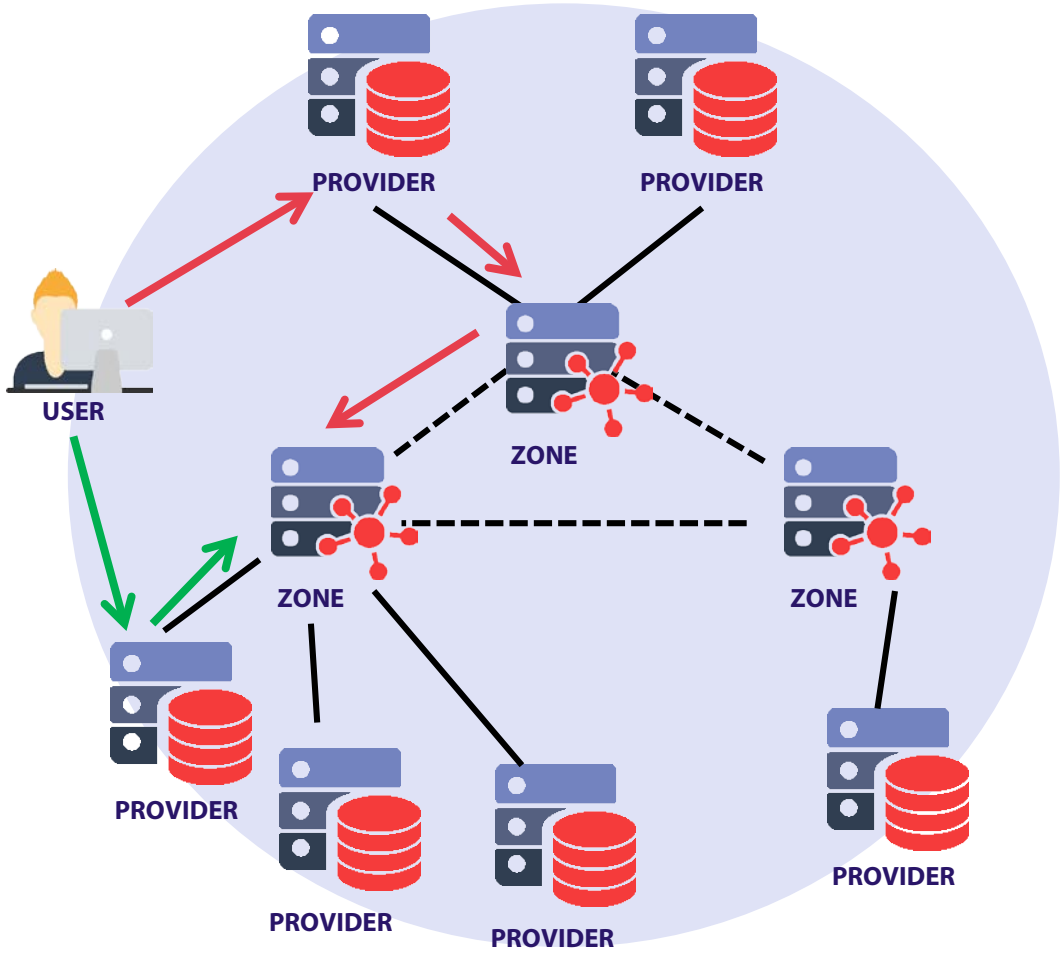
4) Synchronization Protocol



- Each zone is authoritative for given metadata subset
- Other zones act as clients in synchronization
- All SP connections are based on authorization tokens

PROPOSED CONCEPT OF DECENTRALIZED DATA ACCESS CONTROL

5) Users as carriers of trust



- Users choose their zones and providers
- Asking for services is an implicit act of trust
- User presents authorization (token) to chosen provider, which contacts its zone:
  - **User's home zone** – authorization is verified locally
  - **Another zone** – the token is delegated to issuer zone for verification

USE CASE – CROSS-ZONE DATASET SHARING

USER A

USER B

ZONE A

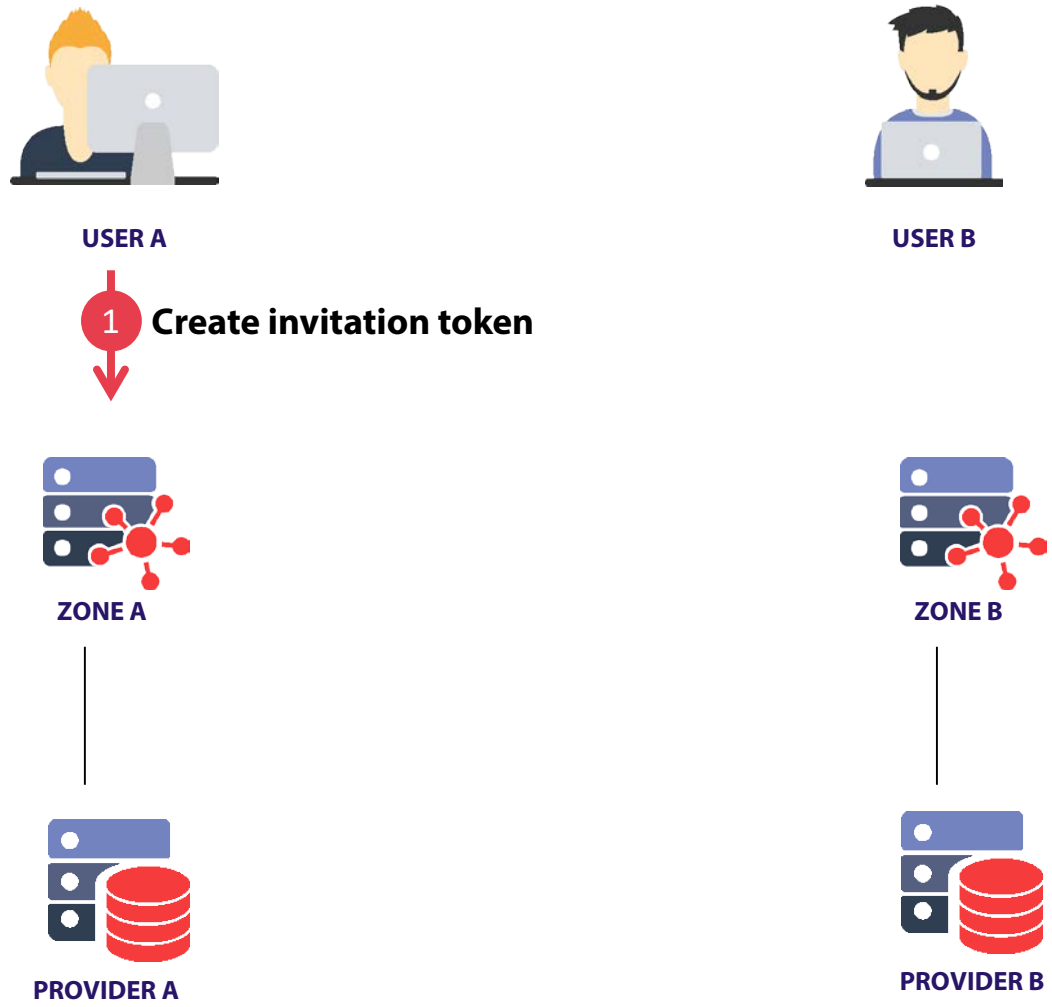
ZONE B

DATASET

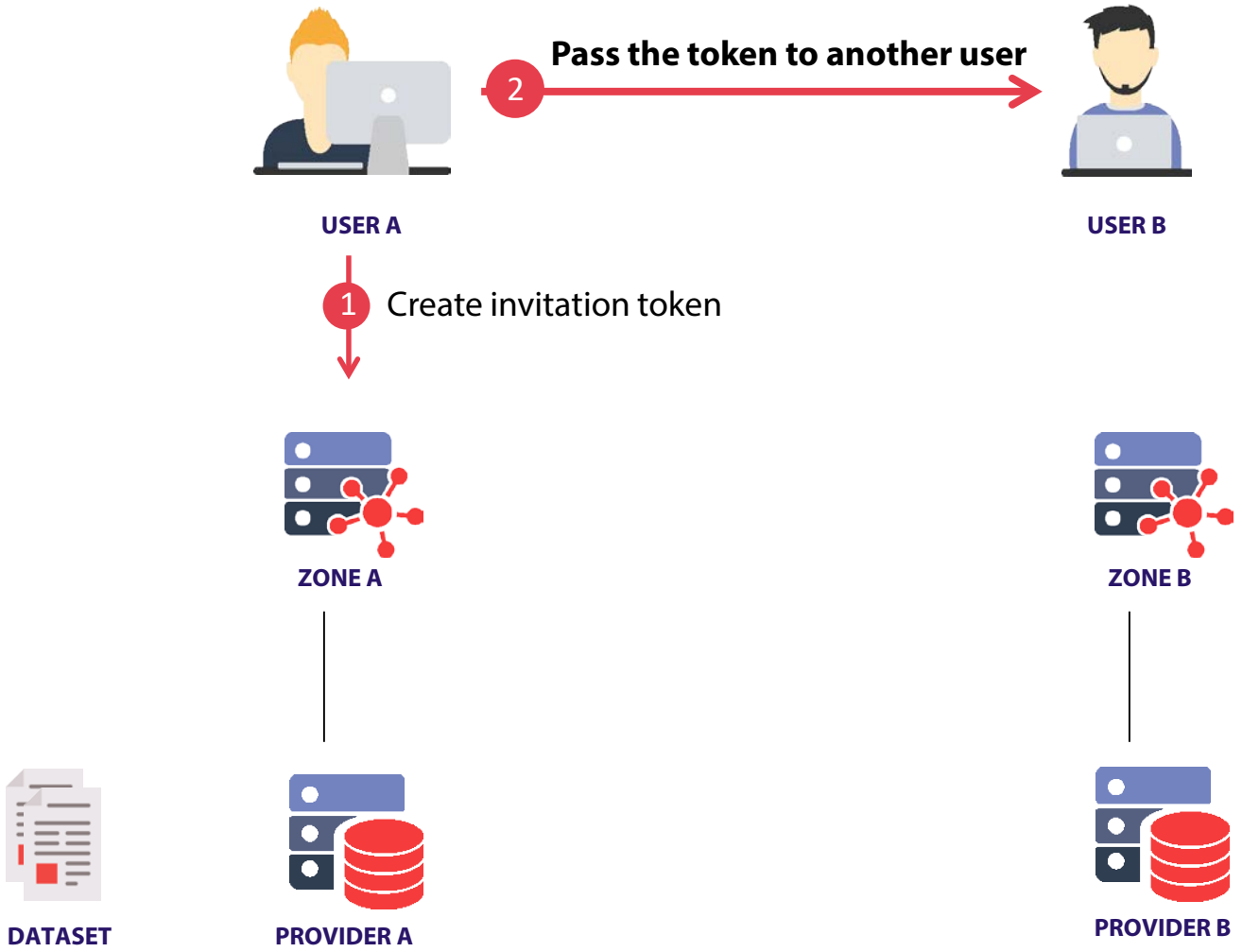
PROVIDER A

PROVIDER B

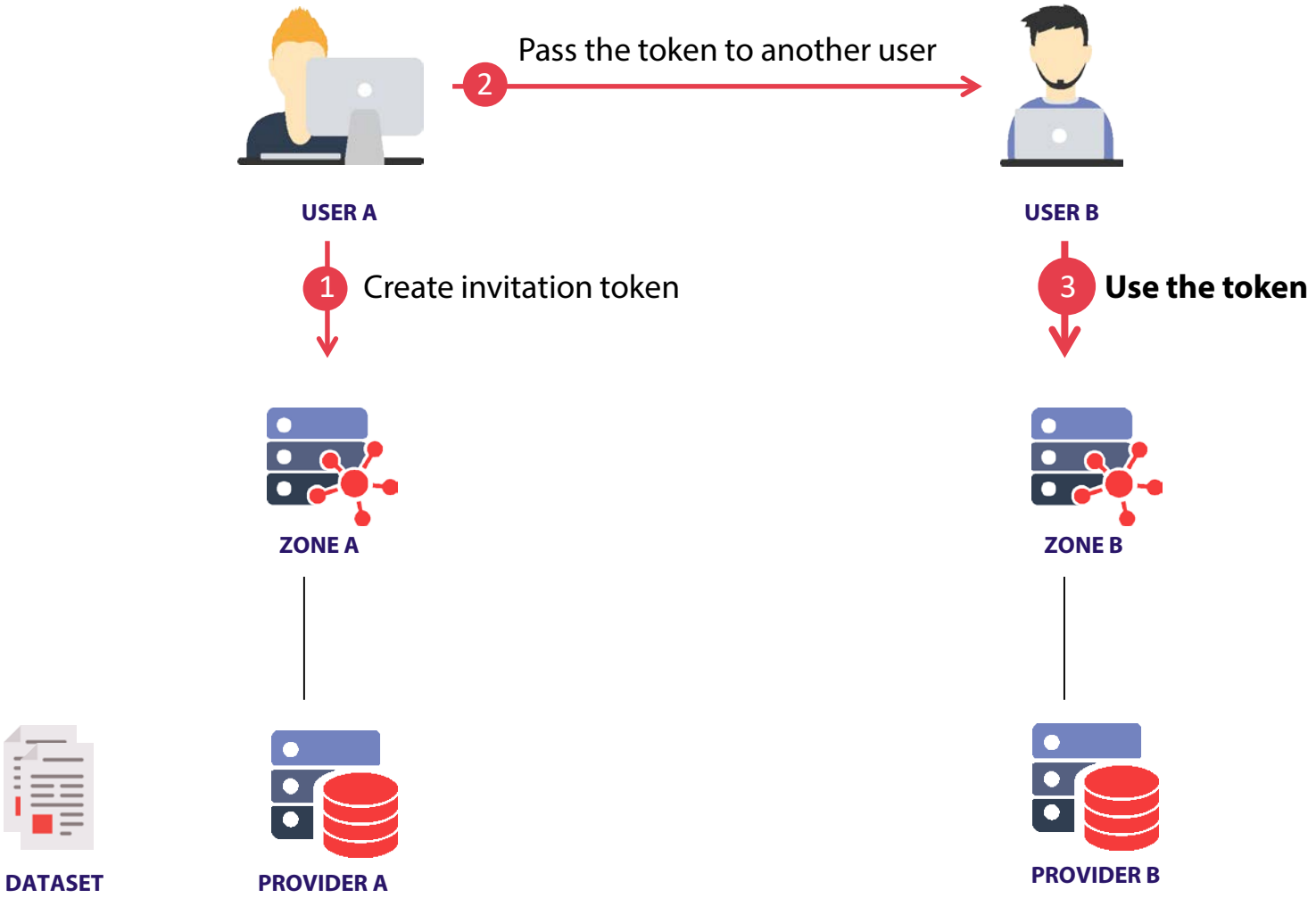
USE CASE – CROSS-ZONE DATASET SHARING



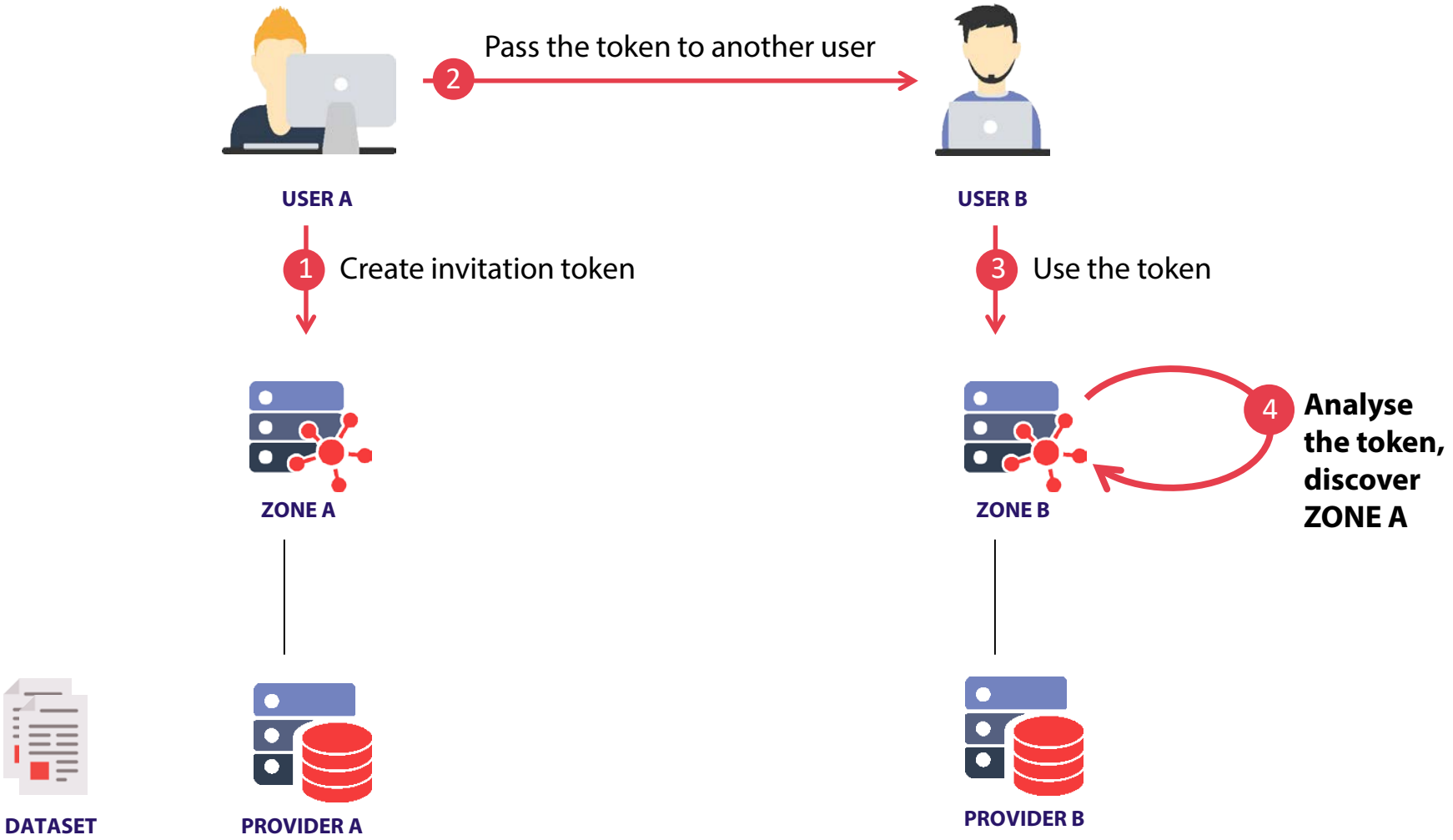
USE CASE – CROSS-ZONE DATASET SHARING



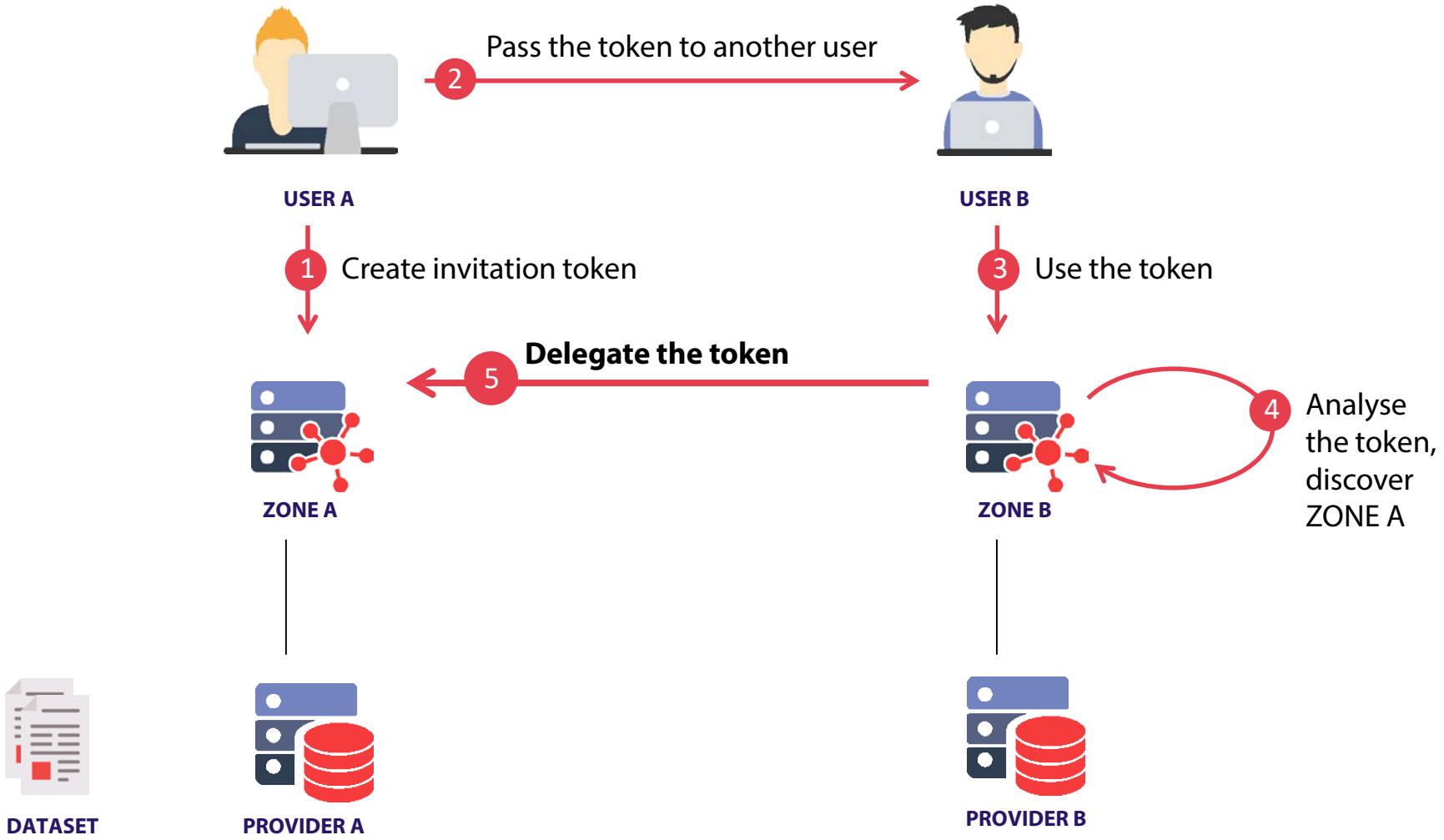
USE CASE – CROSS-ZONE DATASET SHARING



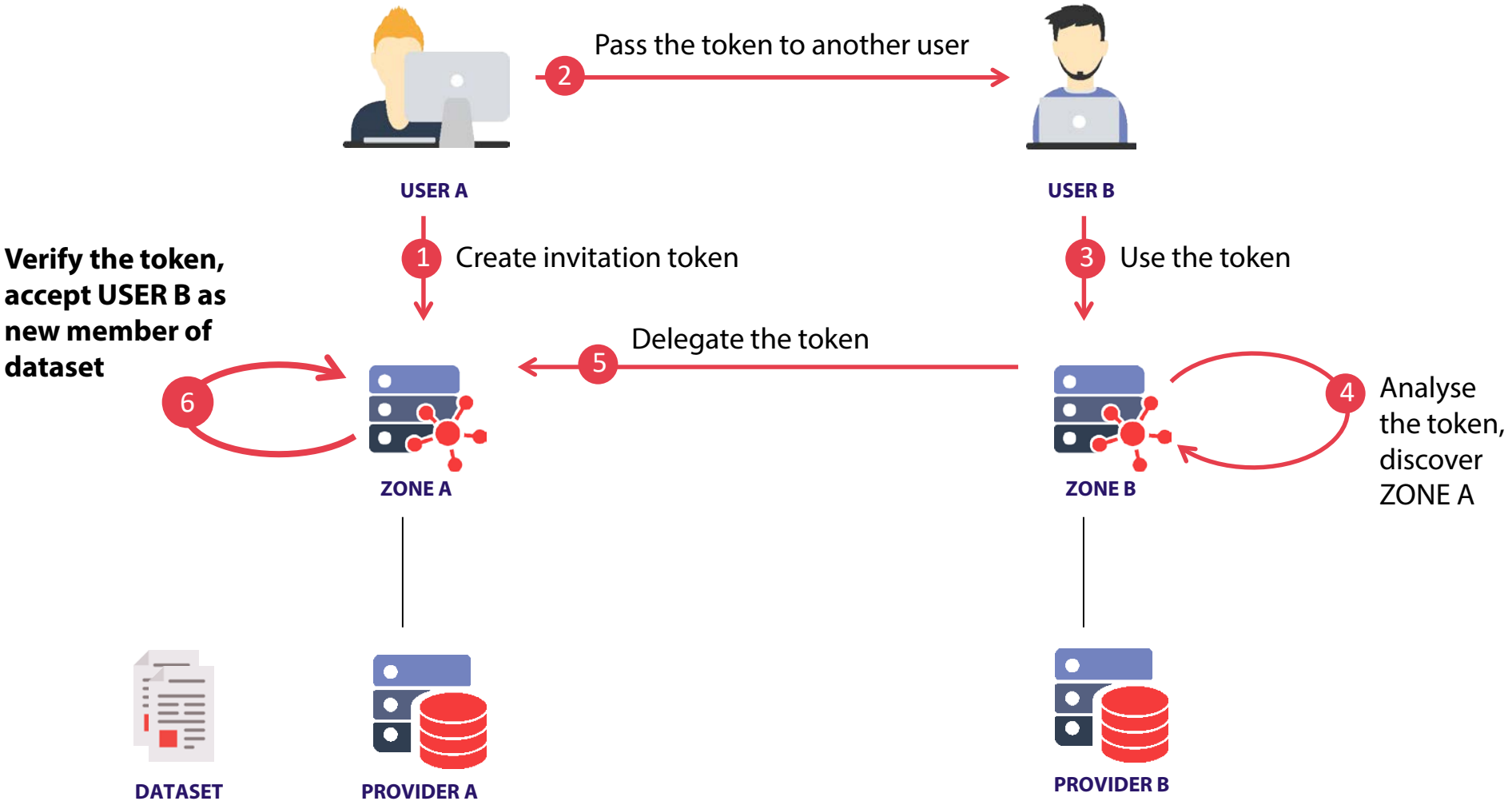
USE CASE – CROSS-ZONE DATASET SHARING



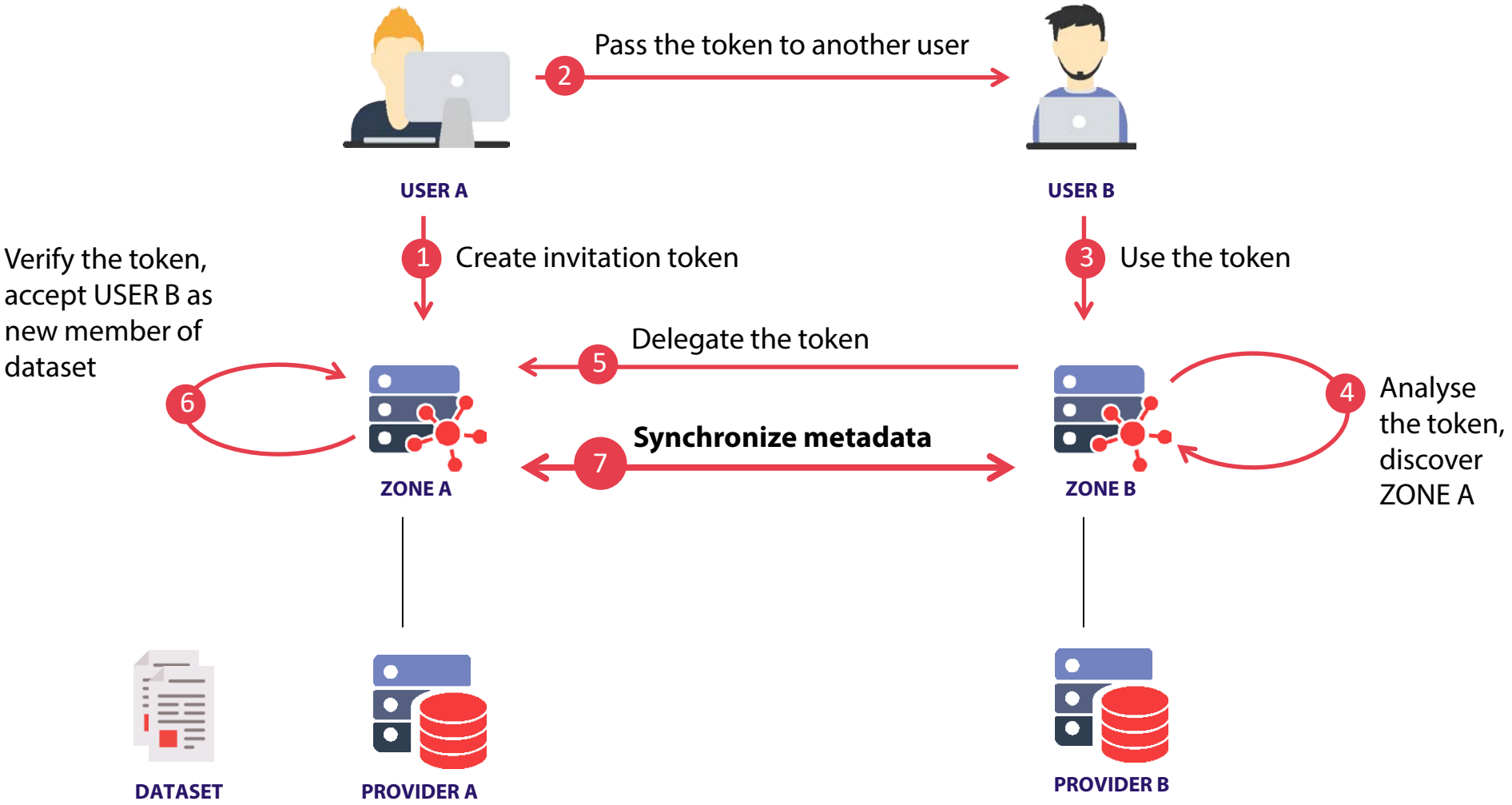
USE CASE – CROSS-ZONE DATASET SHARING



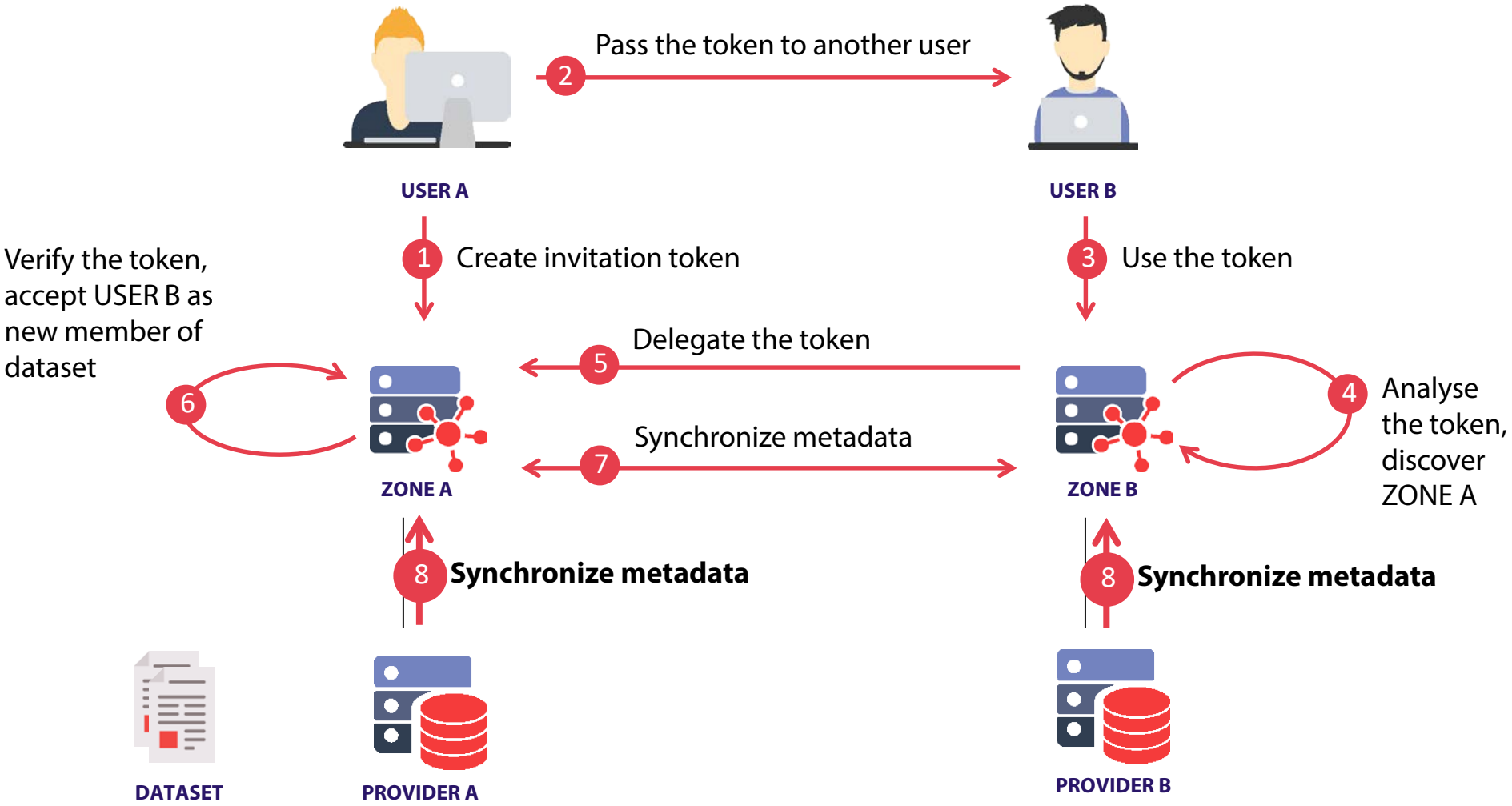
USE CASE – CROSS-ZONE DATASET SHARING



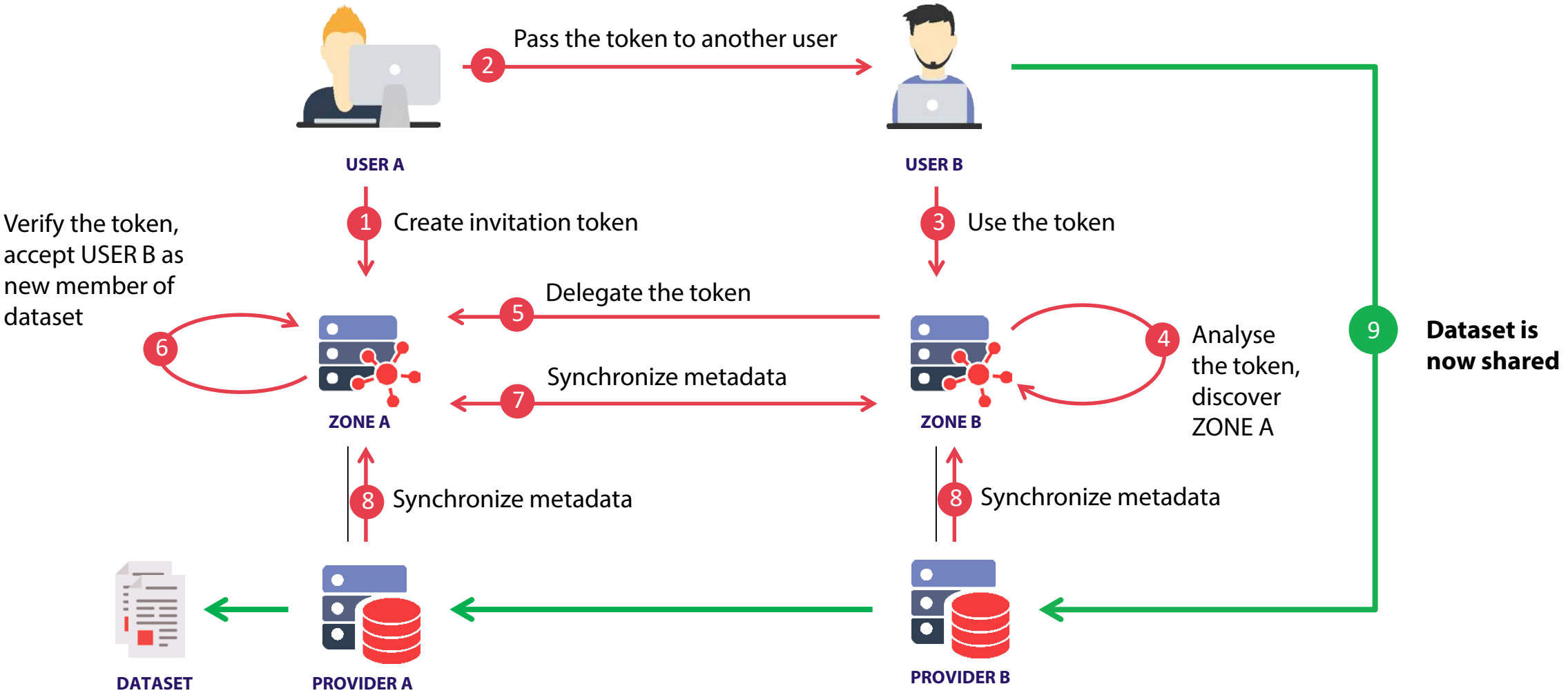
USE CASE – CROSS-ZONE DATASET SHARING



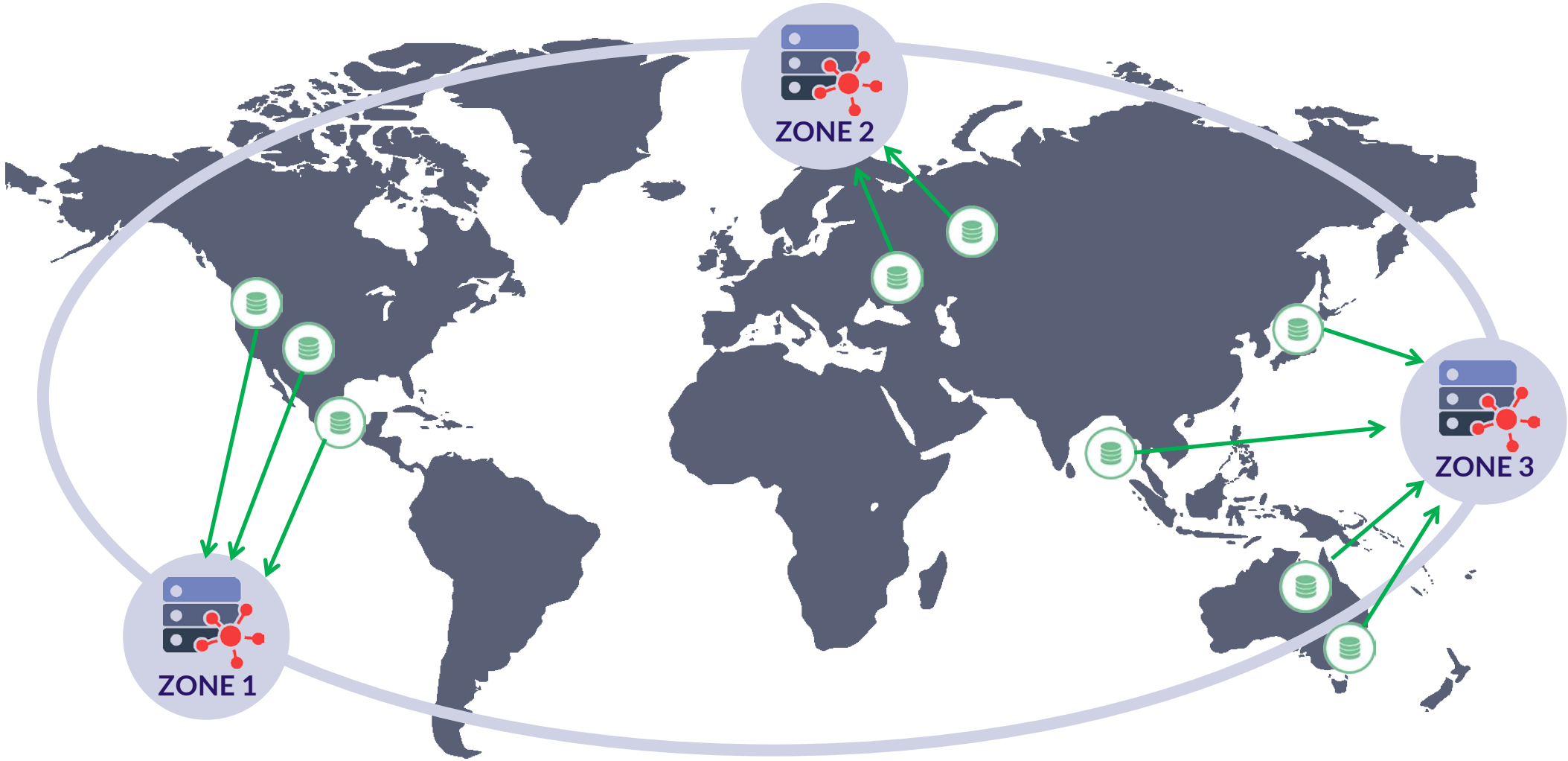
USE CASE – CROSS-ZONE DATASET SHARING



USE CASE – CROSS-ZONE DATASET SHARING



OUR VISION OF GLOBAL DATA ACCESS



Conclusions

- Global data access can be achieved by creating an open network of data providers
- We propose a concept of **decentralized data access control** for such network
- Proposed concept is being implemented in Onedata, a distributed virtual file system

- ✔ Zone server acting as an entry point as well as SP and AAI center
- ✔ Synchronization Protocol for single zone scope
- ✔ Macaroon based AAI for single zone scope
- ✔ Data provider server with cooperation mechanisms
- 🚧 Cross Zone cooperation support
- 🚧 SP and AAI supporting global, cross zone scale

Onezone

Oneprovider

(candidate)

THANK YOU