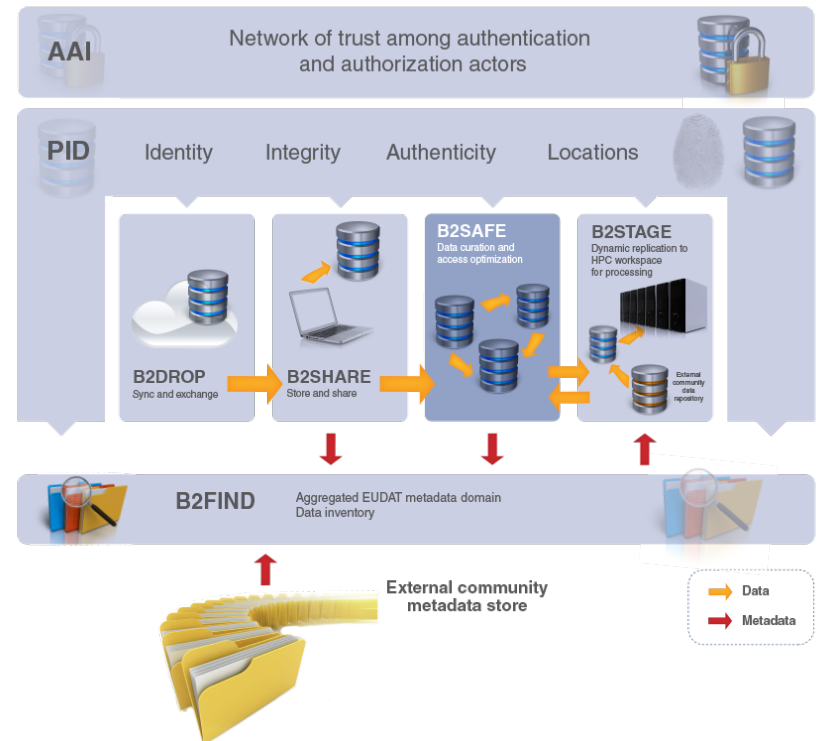


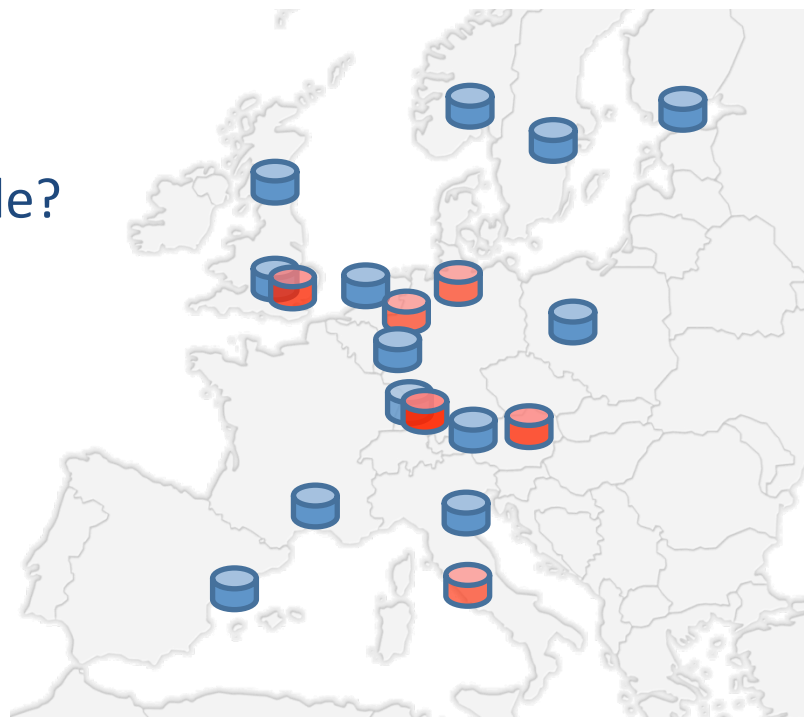


- Core data management service, rule and **policy-driven** data replication **across domains**
- Provide an **abstraction layer** which **virtualises large-scale data resources**
- Guard against data loss in long-term **archiving and preservation**
- Optimise access** for users from different regions
- Bring data **closer to powerful computers** for compute-intensive analysis



# Just scp - where is the problem?

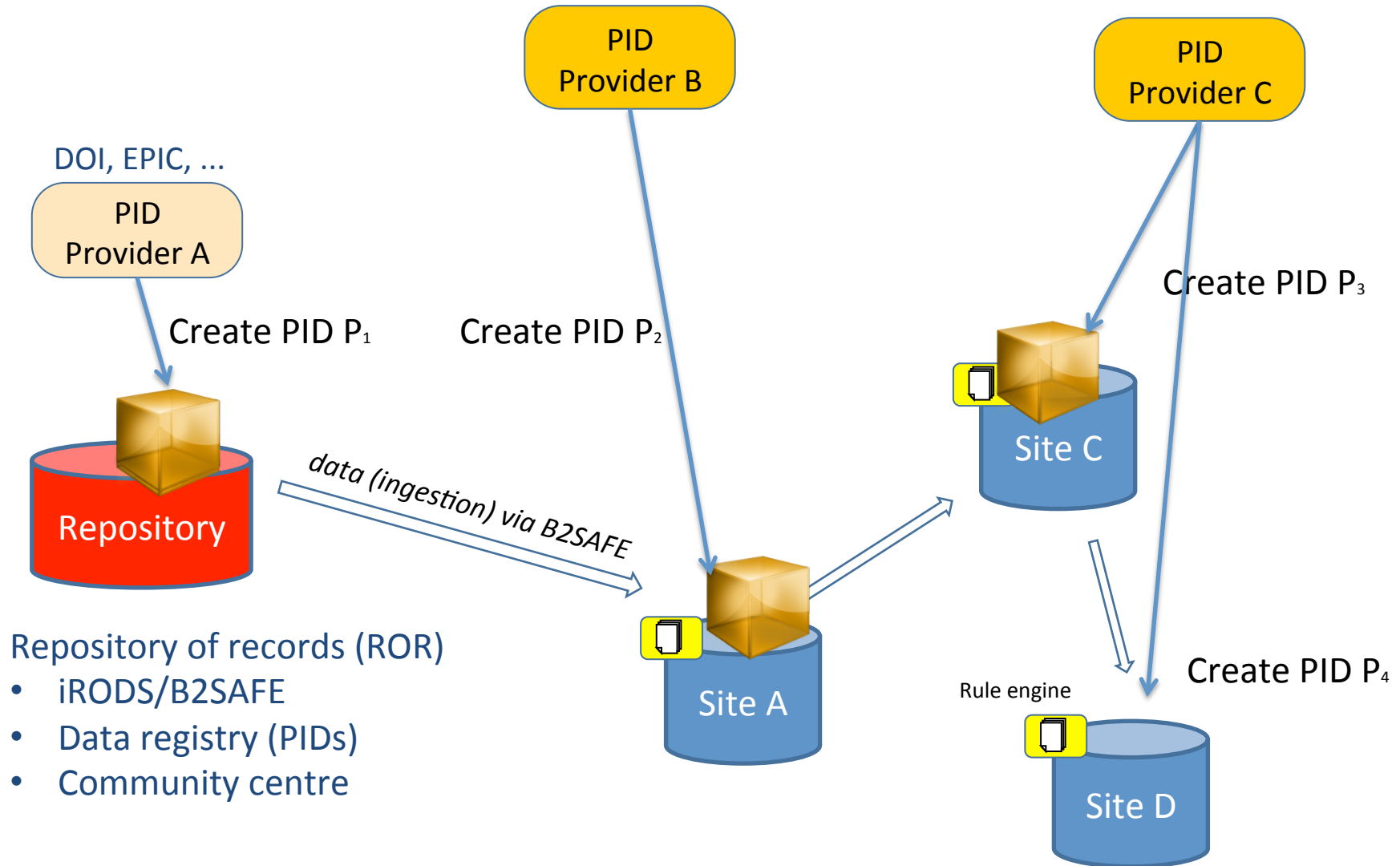
- How do I know in 20 years **where replicas are located**?
- How are the replicas related to each other, which is the original file?
- How to **ensure data integrity** across different sites?



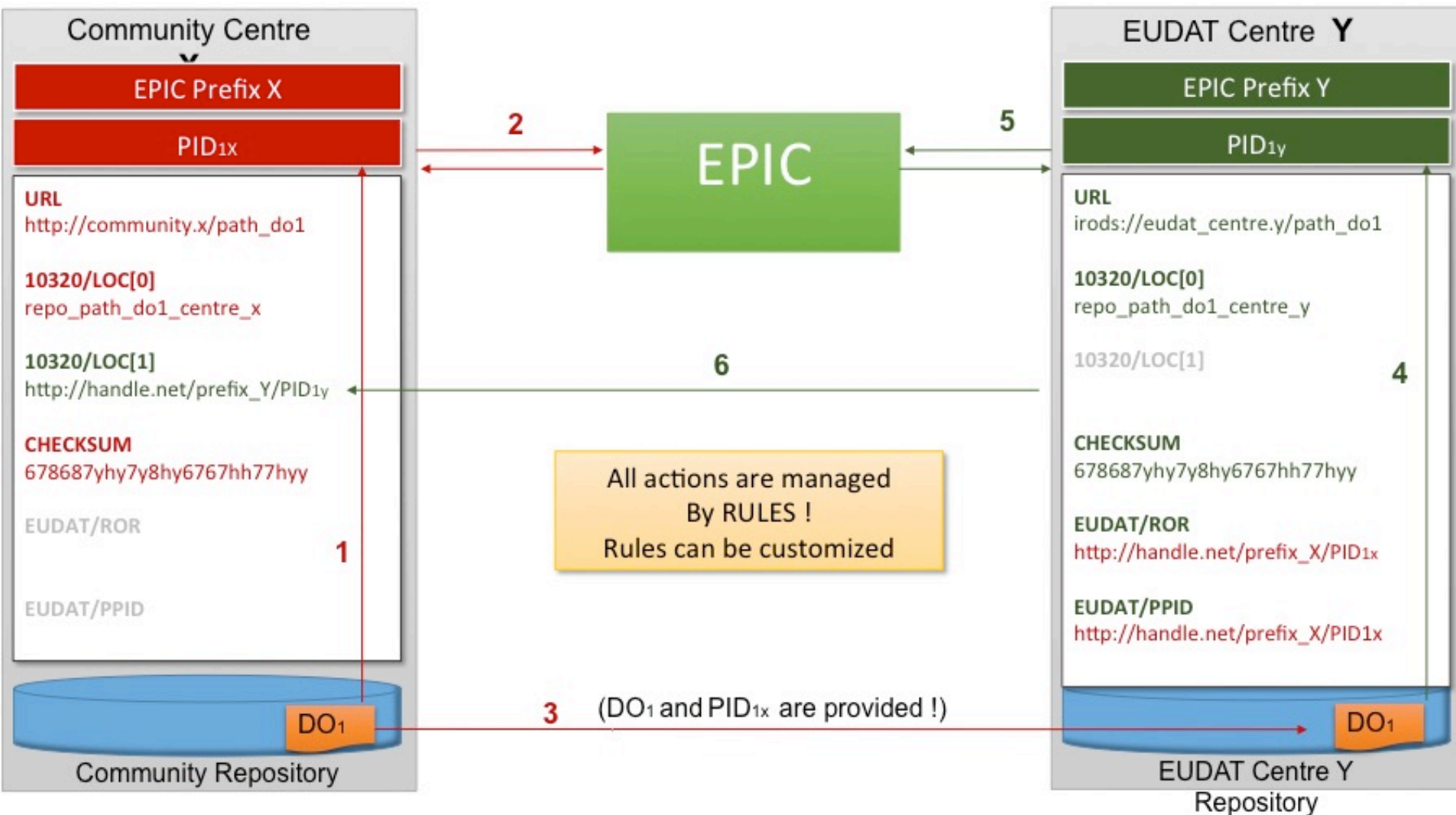
# What do we need?

- **Managing data replication and replicas:**
    - Steer data movements across domains automatically
    - Replicate data automatically upon change to all other sites
    - Send data to specific sites for reprocessing
- **Policy-based** data replication and management across sites
- We need more than a simple copy mechanism
    - **scalable data registry**
    - **mechanism for integrity checks**
    - **managing changes of location of data**

# B2SAFE replication - Joining



# Linking of replicas



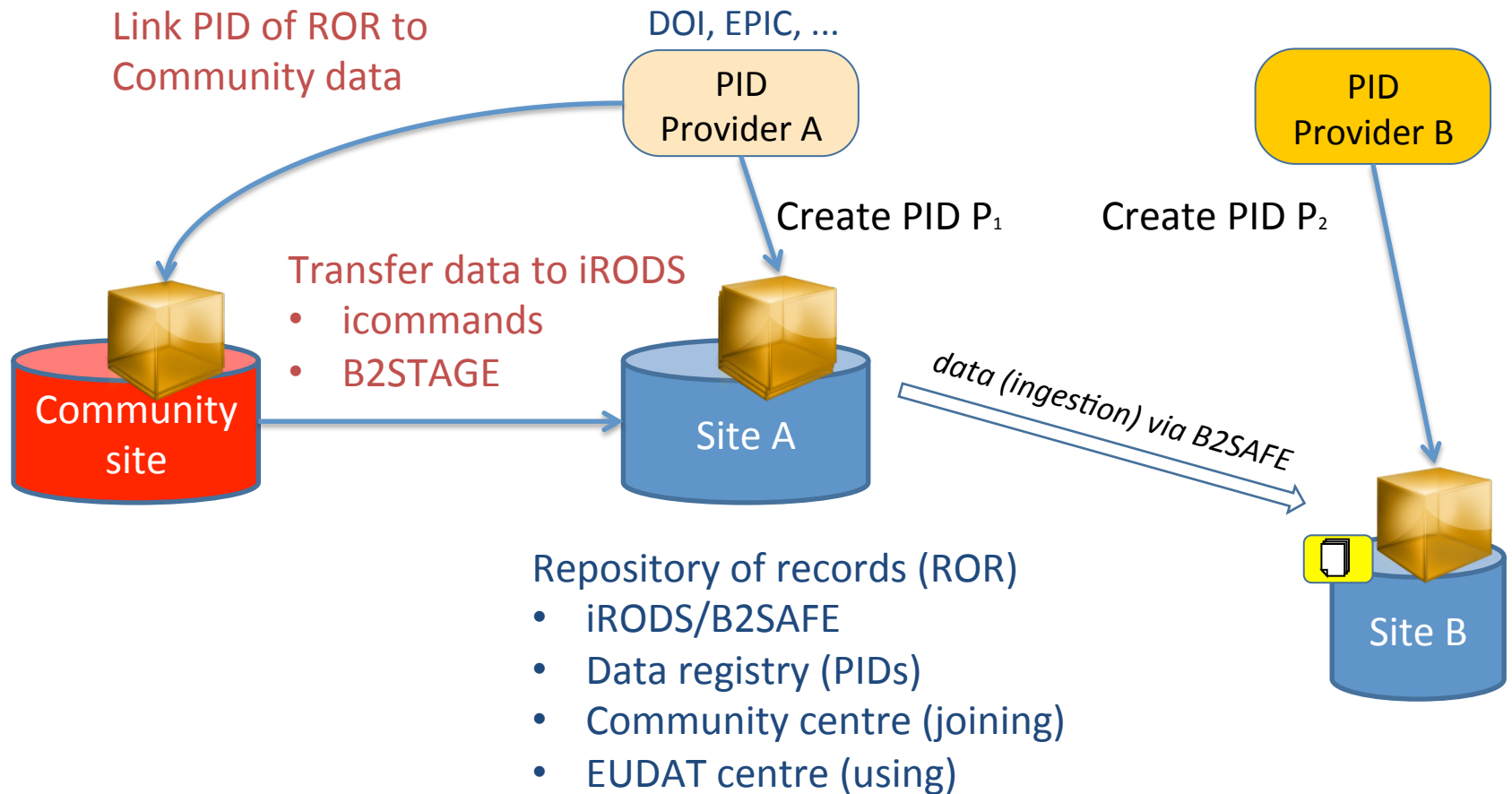
## Example: PID entries for replicas

- PID resolver <http://epic3.storage.surfsara.nl:8001/>

841/7e420ee0-94e3-11e5-bbbe-040091643b95, tick “no redirect”

idx	type	parsed data
1	URL	irods://145.100.59.149:1247/aliceZone/home/b2safeuser#bobZone/testData/put1.txt
2	10320/LOC	<locations><location href="irods://145.100.59.149:1247/aliceZone/home/b2safeuser#bobZone/testData/put1.txt" id="0"/></locations>
3	CHECKSUM	sha2:of/w/++56s5yMMJOUHMfCpHGL5zv3+dxlcL2BxJd/64=
4	EUDAT/ROR	841/841ce83c-94e1-11e5-9438-040091643b9f
5	EUDAT/PPID	841/841ce83c-94e1-11e5-9438-040091643b9f

# B2SAFE replication - Using



# B2SAFE – technologies

- Based on iRODS
- Employs PIDs (handle system) to track data
- Provides iRODS rules
  - Creating checksums
  - Creating PIDs
  - Updating PID entries
  - Replication



# Outline Hands-on

