



A Pan-European Research Storage Infrastructure

Its timeline and relevance for international scientific collaborations

Bo Bai, DeiC

Kalle Happonen, CSC

GN5-2 WP4

Problem statement – for the storage of research data



Storage needs for research data are growing



Datasets are increasing in volume, complexity, and retention requirements



Compliance has become a necessity



Rising user expectations



Sovereignty has regained its importance

Stated Goals for the Pan-European Object Storage Infrastructure

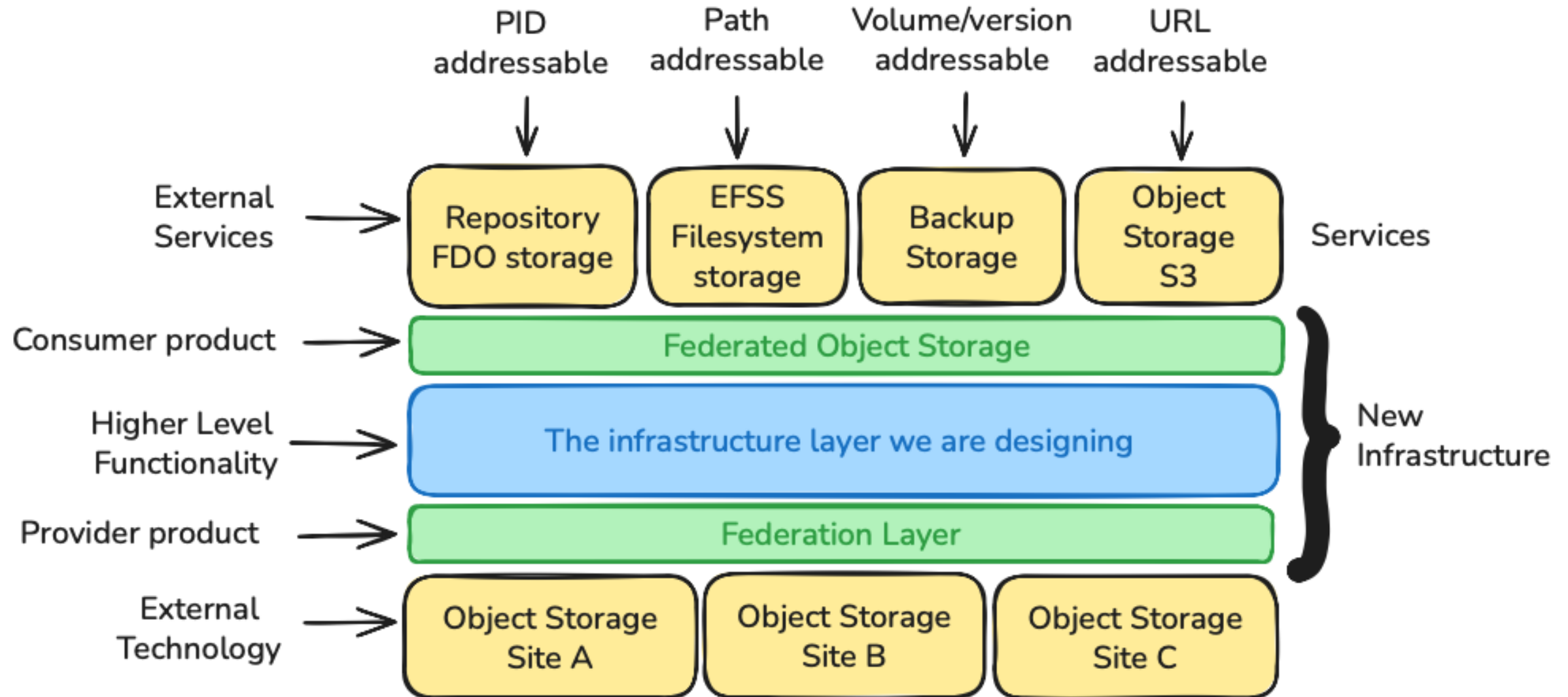
A sovereign, coherent infrastructure for European research

A credible alternative to or way of building your own

Easy to use as a foundation for other services

Long-term sustainability

A feasible roadmap



Infrastructure Design Principles

1

Build Infrastructure

Infrastructure is a building block for someone else's service

2

Use Existing Components

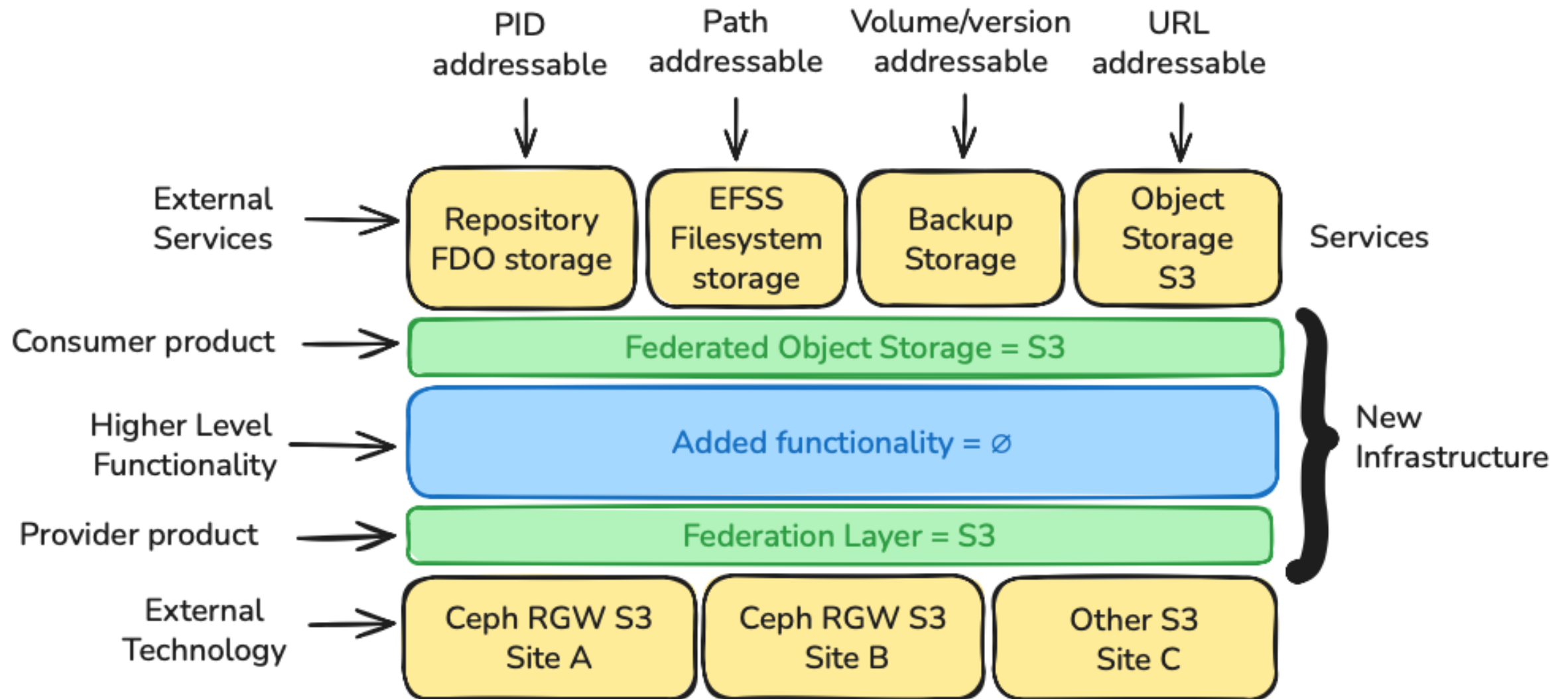
Utilise existing standards and components when possible

3

Cut to the Core

Do not internalise what can be externalised

"Solutions should come from the communities that own the problems"



Roadmap overview



Foster convergence among providers



Establish cross-site functionality



Introduce higher level functions



Establish a federated service

Year 1-3 Blueprints for Object Storage

Blueprints for a Ceph
Storage Cluster

Emphasis on:

Convergence

Compatibility

Sustainability

Compliance

Year 3-5 Higher Level Functions

Higher level functionality

Emphasis on:

Storage Abstractions

Data Management

Data Release

Cross-site Functionality

Years 5+ Federation

Federated service

Emphasis on:

Governance

Certification

Sovereignty

Productisation

Year 1-3: Blueprints for Object Storage

- Version 1 of a Blueprint document for building an object storage service for research
- A core set of independent but compatible provider sites
- A defined interoperability specification
- A test set against the interoperability specification

Year 3-5: Higher Level Functions

- A service ecosystem relying on the object storage services
- A growing set of independent compatible sites
- Cross-site functionality between sites
- Data management and data locality awareness
- First cross-border data-replication use cases
- (Version 2 of the blueprint document)

Year 5+: Federation

- Packaging the ecosystem to a federated service
- Contractual ties between sites
 - Also addressing GDPR, NIS2, and related requirements
- Heavy reliance on other European projects for federating the service
- (Version N of the blueprint document)

In short

- Aggregate of many suppliers
- Highly available
- Highly resilient
- Lukewarm storage
- High bandwidth
- High latency
- Highly trusted
 - Strong data protection where needed
- Long-term preservation
- Basis for the distribution of data



Towards a Pan-European Research Storage Infrastructure

Please join the discussion on Friday:

Side Meeting: 09:00–12:30 Friday 12 June,

Finlandia Hall: Room 25

GÉANT Clouds: **Sovereign Object Storage**, OCRE 2024 Framework, and Cloud Sovereignty

www.geant.org



Co-funded by
the European Union