

# Petabyte scale, OAIS/ISO 16363 conformant archive

David Giaretta <sup>1</sup>, Antonio Guillermo Martínez Largo <sup>2</sup>, María Fuertes <sup>3</sup>, Teo Redondo <sup>4</sup>

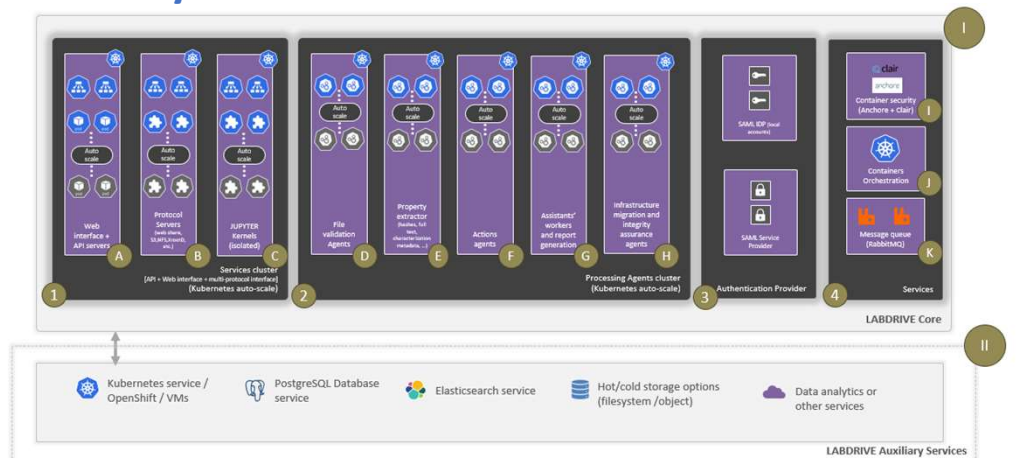
## PetaByte Scale

LABDRIVE has been tested to receive +610 million files and 15PB in a month (500TB/day data rate), scaling itself to more than 6500 Kubernetes pods to process the workload.

### Core Capabilities:



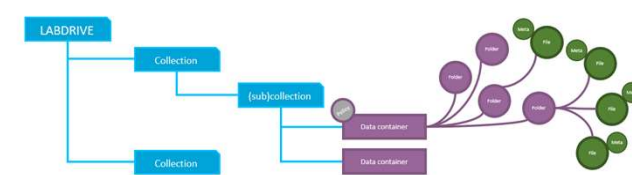
### Scalability:



### Supports the whole data lifecycle:



### Data Containers:



### Types of storage:



## Configuration to support OAIS / ISO 16363 Conformance

### OAIS Preservation Operations - supported by LABDRIVE

#### Data Object may be:

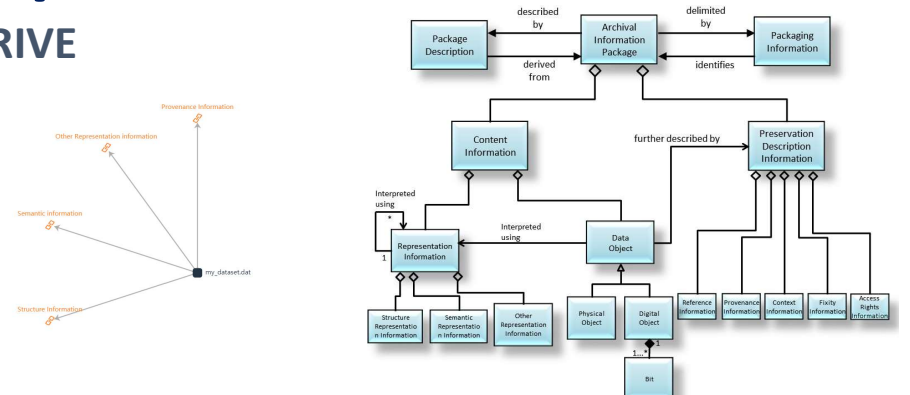
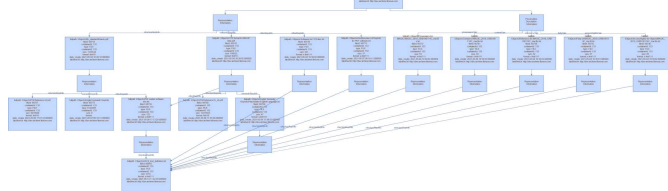
- Kept by archive, unchanged
- Kept by archive but may be changed
- Not kept by archive – handed over

Case 1 – Archive adds Representation Information

Case 2 – Transform the Data Object

Case 3 – Create and hand over complete AIPs

#### Representation Information Network:



OAIS Information Model:

OAIS Archival Information Package

This schema collects the mandatory pieces of information required for a complete OAIS AIP.

Search:

#	Name	Description	Data Type	IECode
1	Structure Representation Information	The Representation information that imparts information about the arrangement of and the organization of the parts or elements of the Data Object.	XML Data	StructureRep
2	Semantic Representation Information	The Representation information that further describes the meaning of the Data Object, and its parts or elements, beyond that provided by the Structure Representation Information.	XML Data	SemanticRep
3	Other Representation Information	A type of Representation Information which cannot easily be classified as Structure Representation Information or Semantic Representation Information, for example software. It is a type of information Object.	XML Data	OtherRep
4	Provenance	The Data Object for the information that documents the history of the Data Object. This information tells the origin or source of the Data Object, any changes that may have taken place since it was originated, and who has had custody of it since it was originated. The Archive is responsible for creating and preserving Provenance Information from the point of ingest; however, earlier Provenance Information should be provided by the Producer. Provenance information adds to the evidence to support Authenticity. This Data Object and its own Representation Information form a Provenance Information Object.	XML Data	ProvenanceObj

## ARCHIVER Project Framework

LIBNOVA CONSORTIUM (ARCHIVER Project' winner):

LABDRIVE PLATFORM is a Research Data Management and Preservation platform resulting of a joint effort and intense R&D. With it, Researchers can **do more** while Organizations **reduce risks and costs**.

ARCHIVER PROJECT:

Archiving and Data Preservation services for Research Environments for PB-scale datasets using commercial cloud services via the EOSC.

<https://archiver-project.eu/>



<sup>1</sup> PTAB Ltd, Dorset, UK, [david@giaretta.org](mailto:david@giaretta.org) | <sup>2</sup> LIBNOVA, Spain, [a.guillermo@libnova.com](mailto:a.guillermo@libnova.com) |

<sup>3</sup> LIBNOVA, Spain, [mfuertes@libnova.com](mailto:mfuertes@libnova.com) | <sup>4</sup> Formerly at LIBNOVA, now at Bidaidea, Spain, [teofilo.redondo@bidaidea.com](mailto:teofilo.redondo@bidaidea.com)