Re-introducing the REST VOL

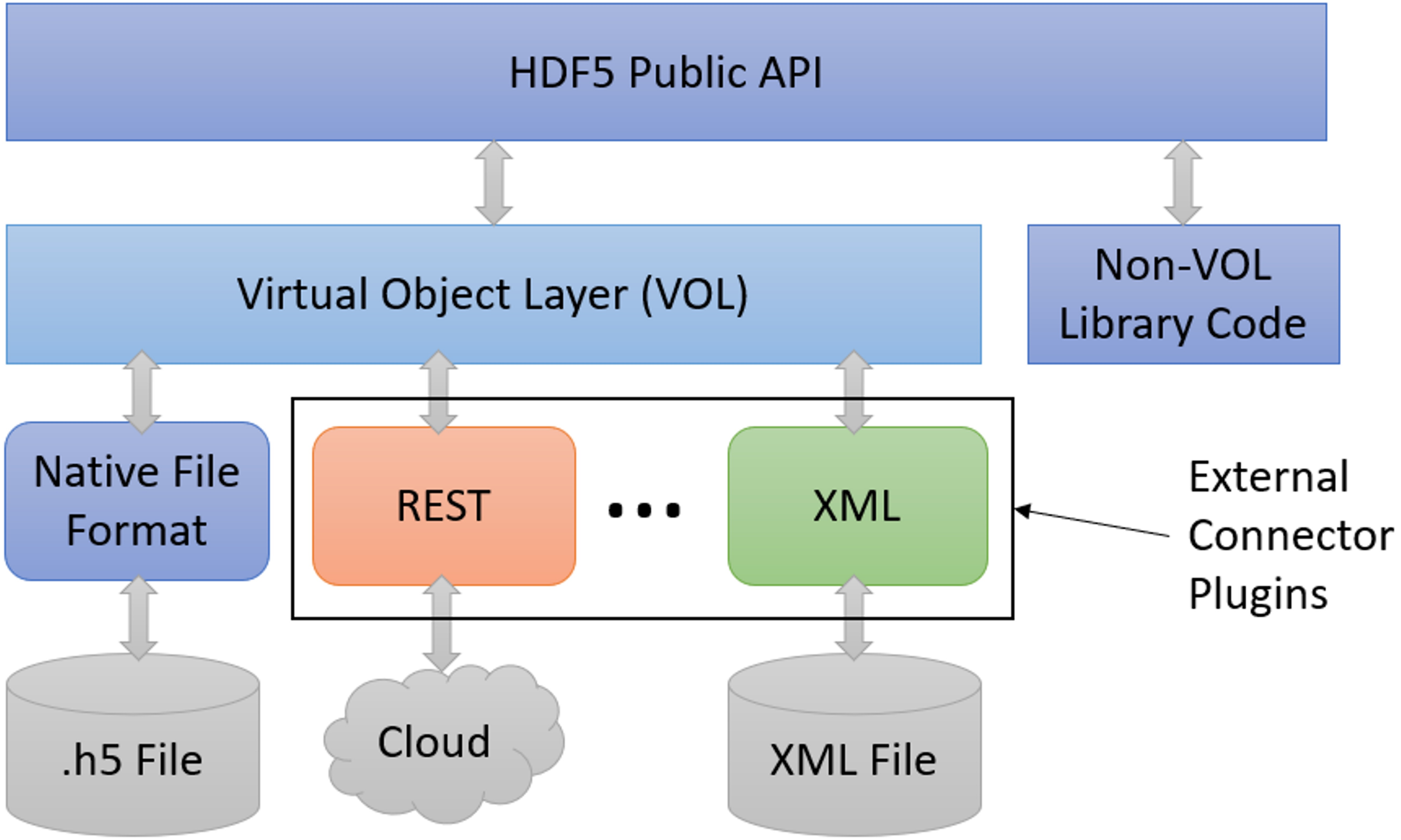
Matt Larson

The HDF Group

HUG 2023

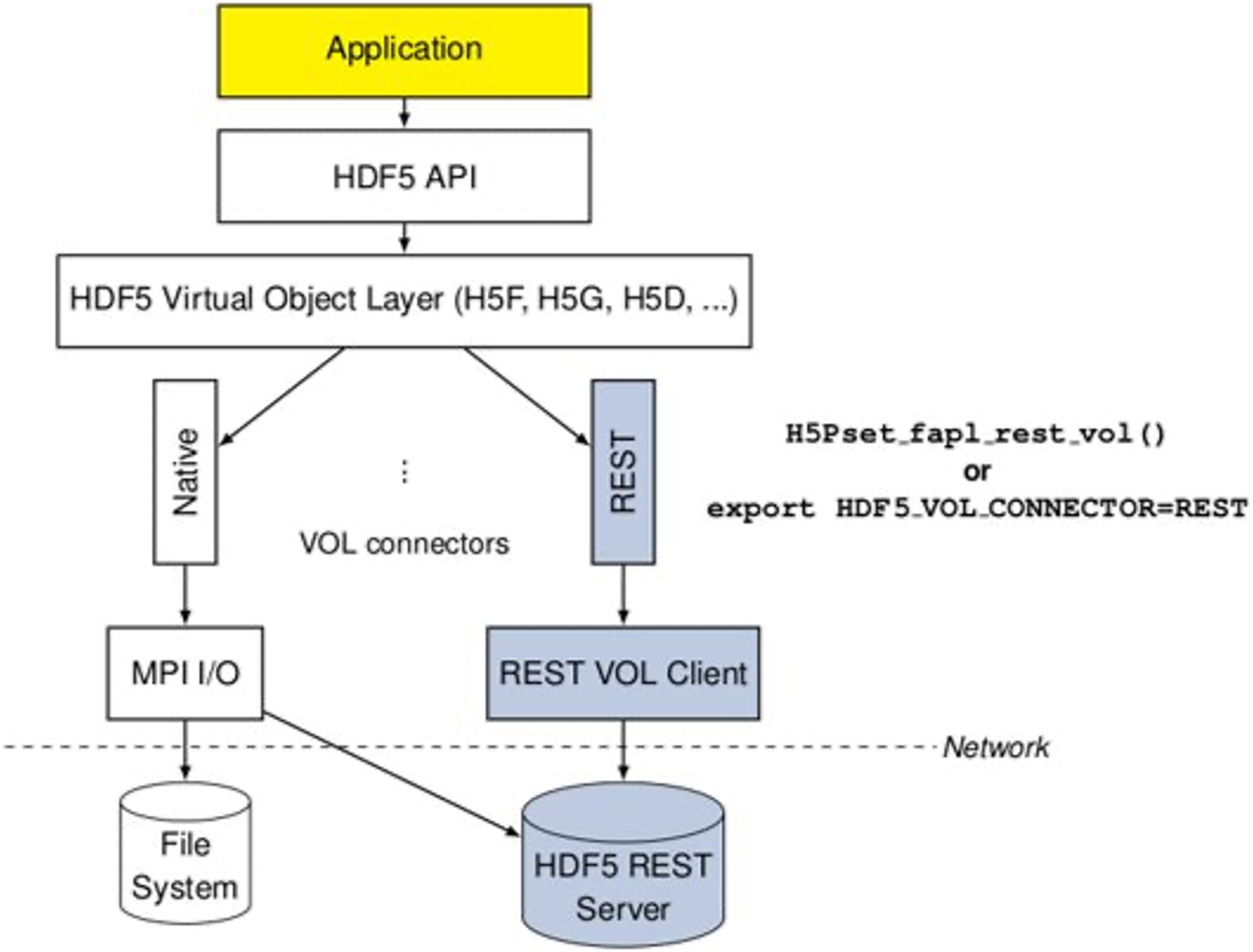
The Virtual Object Layer (VOL)

- Sits between library API and storage
- Changes storage operations
- Different connectors exist



The REST VOL

- Maps API calls to REST requests to external server
- Implements (most) core library functionality
- Originally developed 2017-2020 for HDF5 1.12
- Back in active development

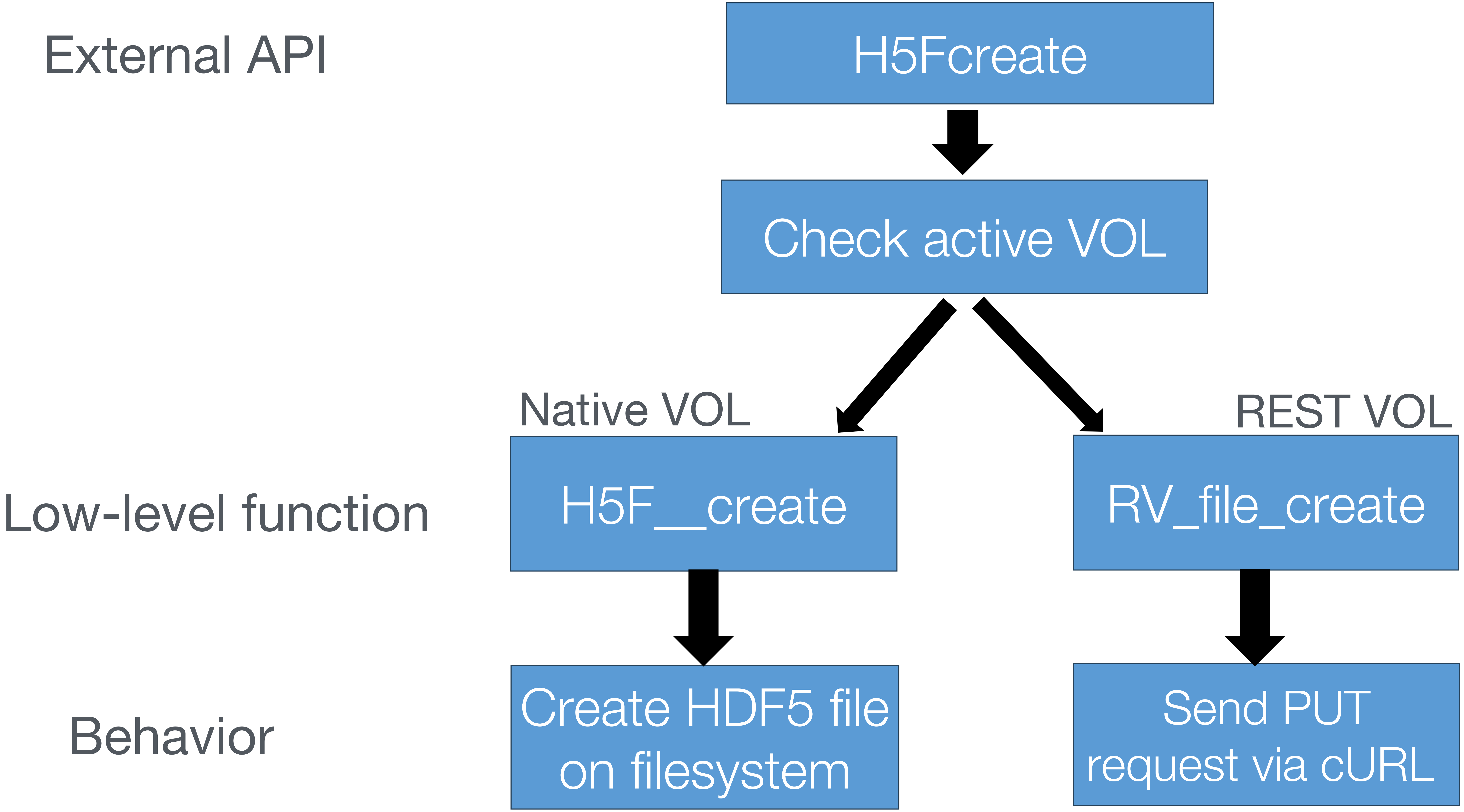


(e.g. HSDS)

github.com/HDFGroup/vol-rest

Example – H5Fcreate

The HDF Group



External API

H5Fcreate

Check active VOL

Native VOL

REST VOL

Low-level function

H5F__create

RV_file_create

Behavior

Create HDF5 file on filesystem

Send PUT request via cURL

HSDS (Highly Scalable Data Service)

The HDF Group

- HDF5 on the cloud via REST API

Storage Options

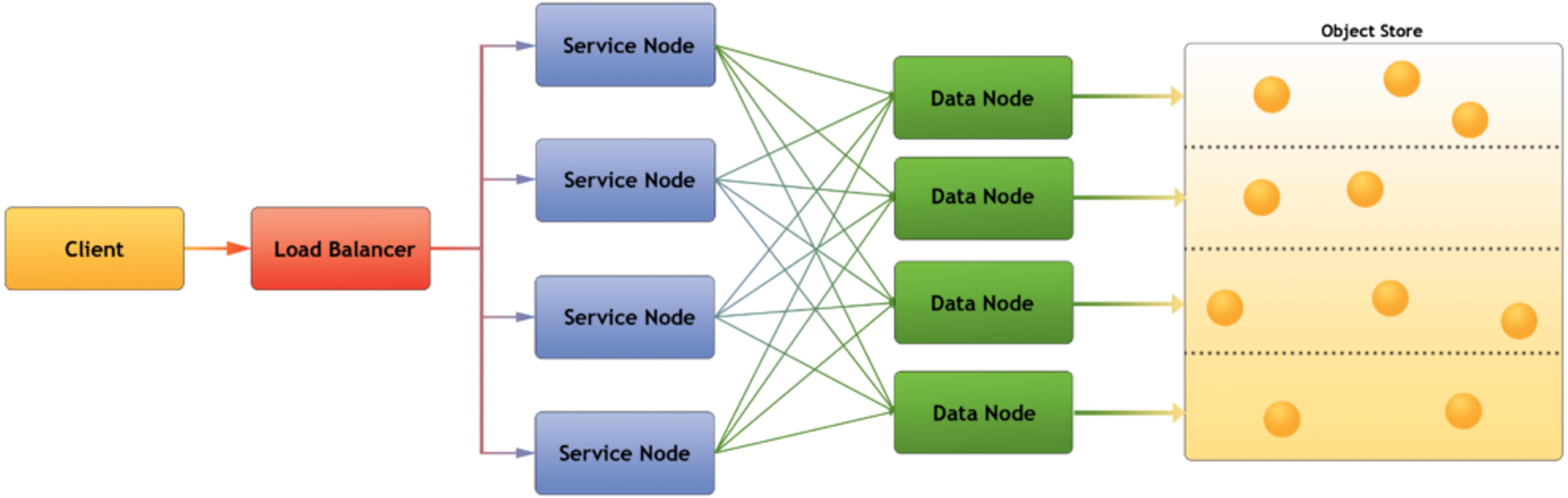
- AWS S3
- Azure Blob
- OpenIO
- …

Deployment Options

- Docker
- Kubernetes
- AWS Lambda
- …

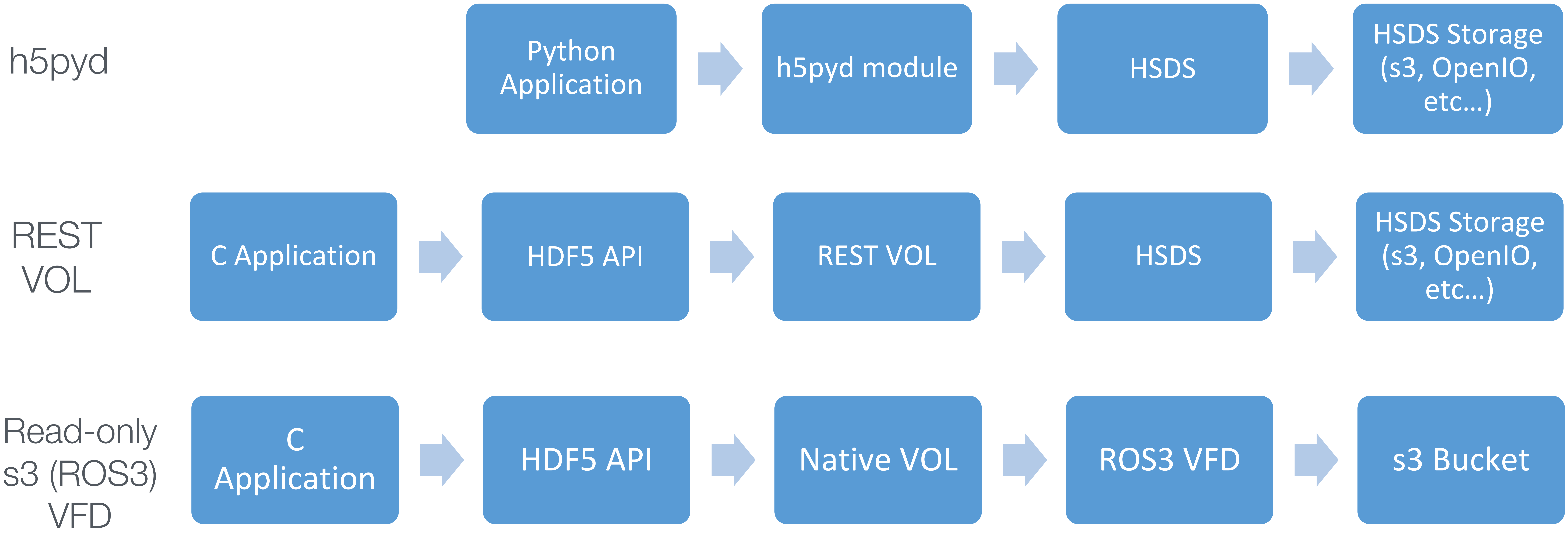
Access Options

- h5pyd
- HS CL tools
- **REST VOL**
- …



github.com/HDFGroup/hsds

HDF in the Cloud Options



h5pyd

Python Application → h5pyd module → HSDS → HSDS Storage (s3, OpenIO, etc…)

REST VOL

C Application → HDF5 API → REST VOL → HSDS → HSDS Storage (s3, OpenIO, etc…)

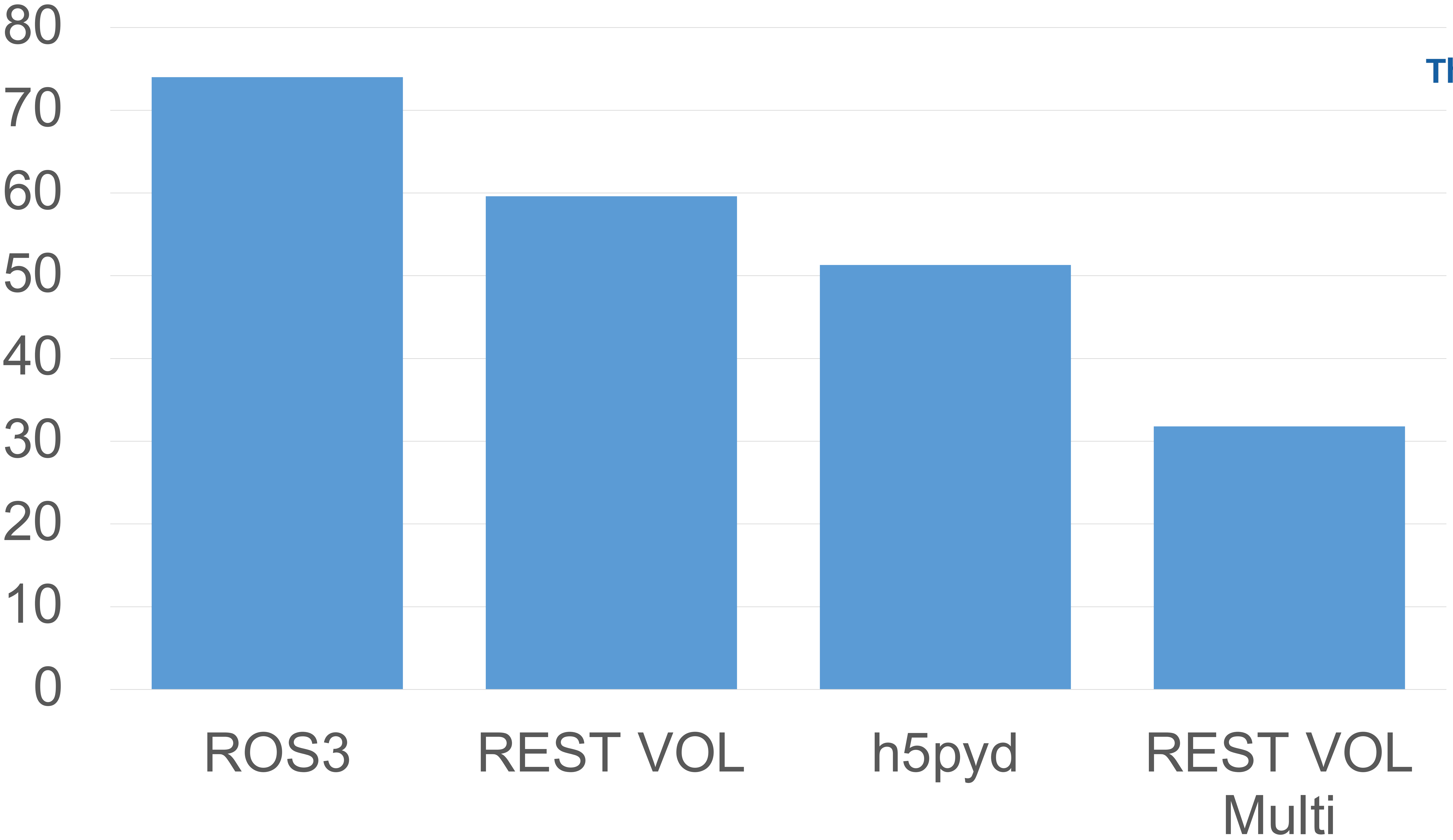
Read-only s3 (ROS3) VFD

C Application → HDF5 API → Native VOL → ROS3 VFD → s3 Bucket

Why to use the REST VOL

	h5pyd	ROS3 VFD	REST VOL
Language	Python	C	C
Read?	✓	✓	✓
Write?	✓		✓
Multi-Read/Write?	TBD		✓

Benchmark Time (seconds)



The HDF Group

github.com/mattjala/nasa_cloud/

How to use the REST VOL

- See also: User's Guide section 2.2.1 + 2.2.2
- First: Set up HSDS server
- VOLs sit under the API → Most code of your code is unchanged
- 2 ways to start using the REST VOL

Dynamically Loaded

- Default: All file access routed to HSDS
- Change FAPL (File Access Property List) to access filesystem

HOWTO:

- `HDF5_PLUGIN_PATH`
- `HDF5_VOL_CONNECTOR`

Manually Linked

- Default: File accesses work like normal
- `H5Pset_fapl_rest_vol()` to access HSDS

HOWTO:

- Include `rest_vol_public.h`
- Link against REST VOL library
- `H5rest_init(), H5rest_term()`

THANK YOU!

Questions & Comments?