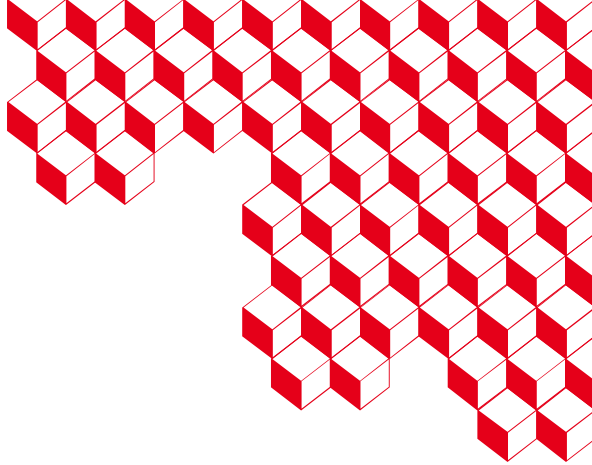


Phobos: an open-source scalable object store optimized for tapes

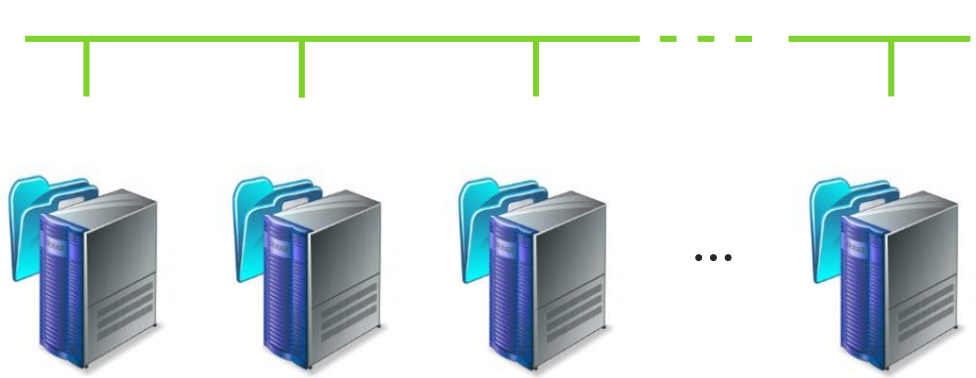
Features overview and Production case at CEA

Gauthier EVRAERD, gauthier.evraerd@cea.fr

Context

Needs for extremely scalable storage systems at a reasonable price

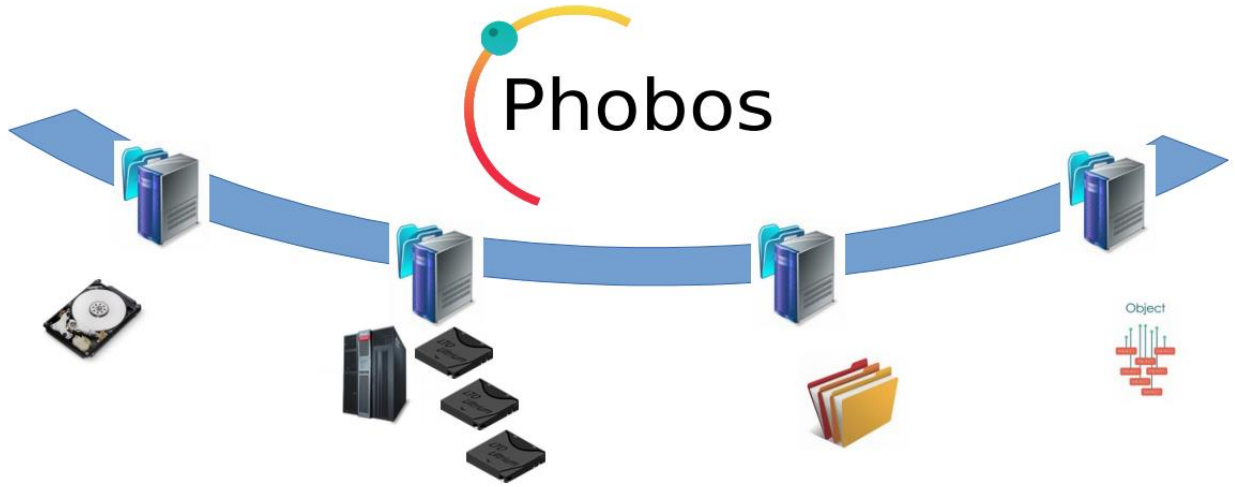
- Object store: horizontal scalability
- Tape library: safe long-term storage at low cost



Phobos: Parallel Heterogeneous Object Store

Goals :

- Manage a distributed set of storage resources on various storage technologies (HDD, tapes, object stores...)
- Implement the best I/O optimizations for each technology without compromise
  - E.g. for tapes: minimize mounts and data sync



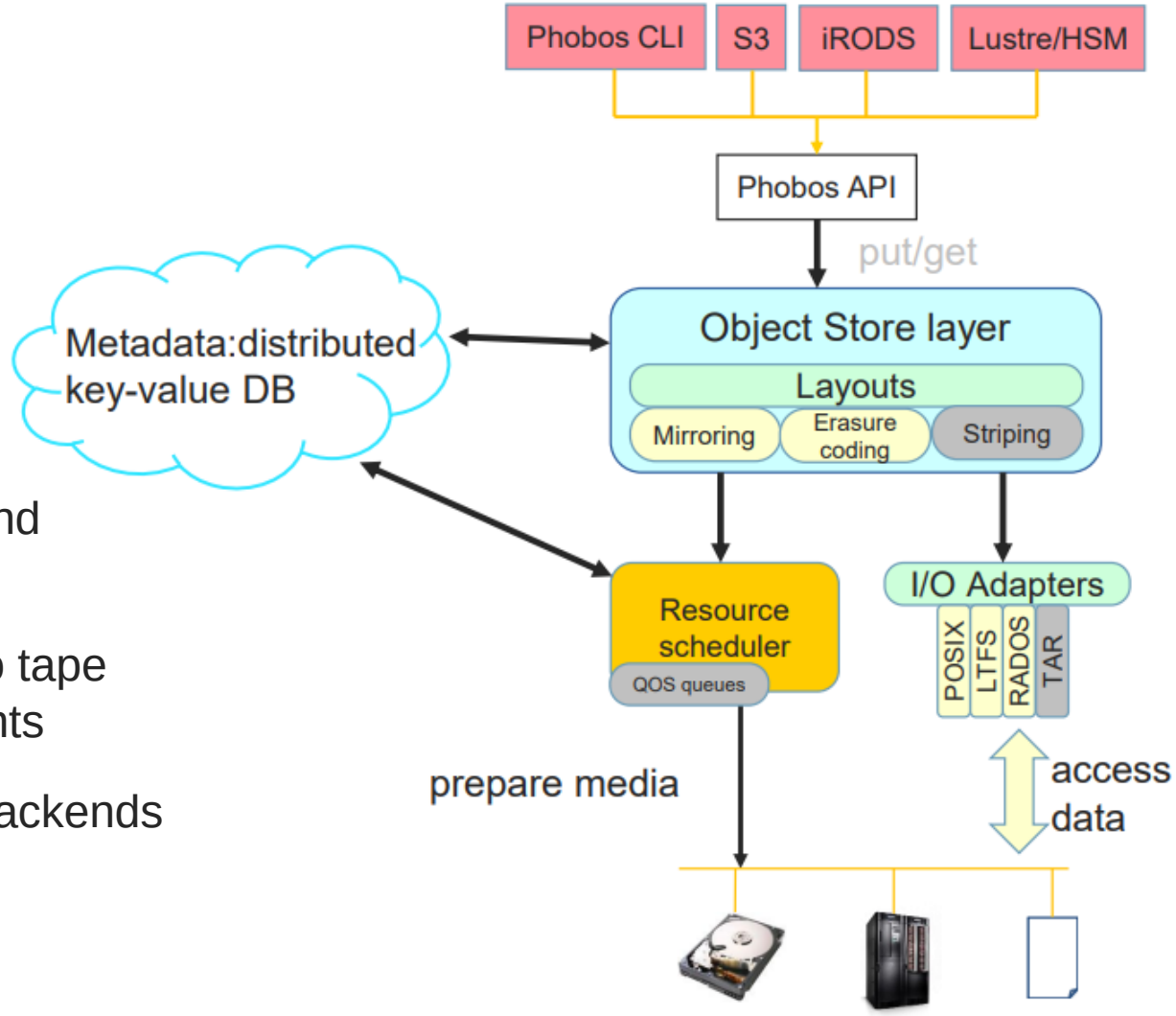
History of the project

- **2014-2015:** development of the initial version
- **2016:** multi-Petabyte of genomics data at TGCC
- **2019:** Phobos made open-source (LGPL v2.1), available on github
- **2022:** Phobos as Lustre/HSM backend at CCMD
- **2023:** First version of Phobos in production outside CEA, Stanford Research Computing, University IT, USA
- **2024-2025:** development of features to deploy Phobos at EXA and TGCC

Phobos: components overview

- **Front-ends:** API
  - Connectors: CLI, S3, iRODS, Lustre/HSM
- **Layout plugins:**
  - RAID1 (mirroring with n copies)
  - RAID4 (2+1 with XOR)

  The layout goal is to increase performance and provide fault tolerance
- **Resource scheduler:** optimizes stream to tape drives (grouped read), minimizes tapes mounts
- **I/O adapters:** support of multiple storage backends (POSIX, LTFS, RADOS)



1. Phobos features overview

Phobos features illustrated

Easy setup

- Drive setup

```
phobos drive add --unlock /dev/st1
```

- Tape addition & formatting

```
phobos tape add --type lto9 [073200-073222]L9

phobos tape format --unlock [073200-073222]L9
```

All done! Phobos is ready for I/Os!

Phobos features illustrated

Adding objects

```
phobos put file/to/put objid
```

Options available

- **Family**
  - Directory, tape, …

```
phobos put --family tape file/to/put objid
```

- **Layouts**
  - Raid 1

```
phobos put --layout raid1 --lyt-params repl_count=2 file/to/put objid
```

  - Raid 4 (2 + 1)

```
phobos put --layout raid4 file/to/put objid
```

Phobos features illustrated

Object versioning:

- Object uniqueness

```
phobos put /path/to/file objid1
```

→ fails if objid1 exists

- Creating new object version

```
phobos put --overwrite /path/to/file objid1
```

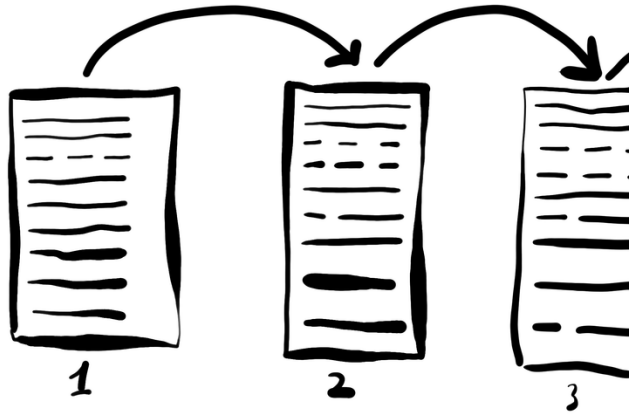
→ creates a new version of objid1

- Listing object versions

```
phobos object list --deprecated objid1
```

- Retrieving an old version

```
phobos get --version 1 objid1 file.out
```

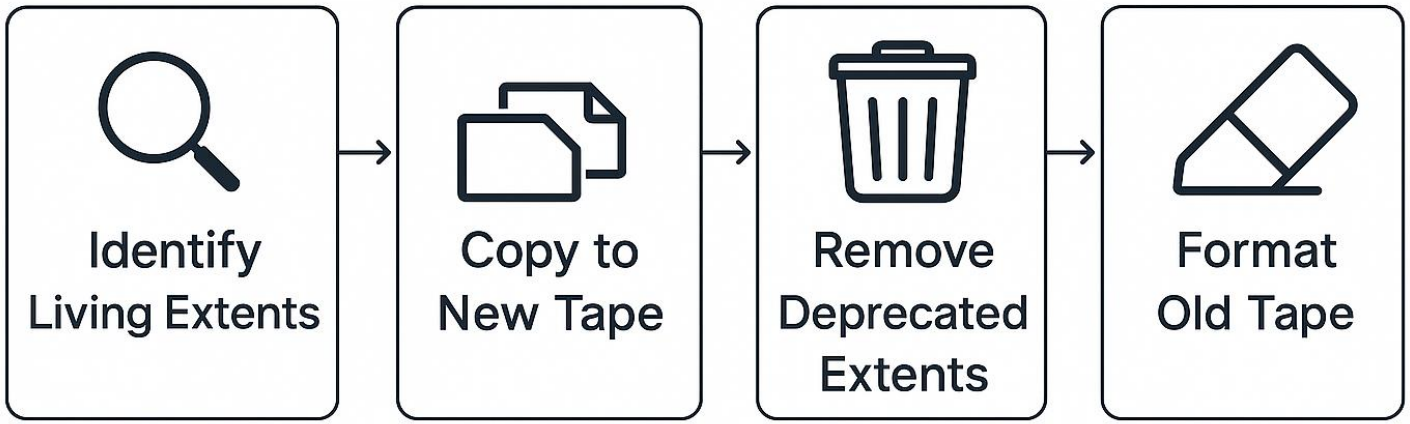


Phobos features illustrated

Repack: tape life cycle

- Optimize tape utilization
- Migrate data between tapes

```
phobos tape repack 073222L9
```



Phobos features illustrated

Resource partitioning with tags:

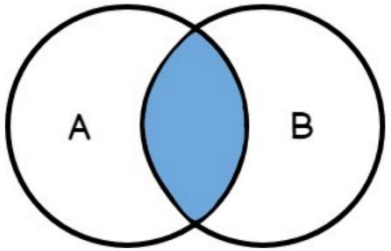
- Tagging resources

```
phobos tape update --tags classA,classB [073000-073099]L9
```

- Pushing data to specific resources

```
# push data to any media with tag “classB”
phobos put --tags classB /path/to/file objid1

# push data to a media with both tags “classA” and “classB”
phobos put --tags classA,classB /path/to/file objid2
```



2. Phobos 3.0: The object copy feature

Phobos objects copies

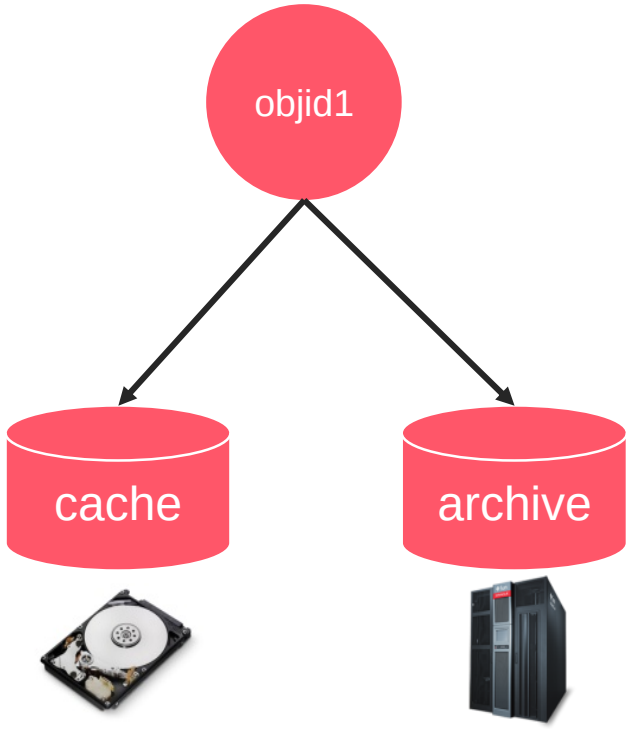
An object can now have multiple copies

This generic feature can be used to

- Migrate data between different storage support (HPSS → Phobos)
- Do internal HSM (Hierarchical Storage Management) policy (dir ↔ tape)
- Do security mirroring between multiple sites

Phobos objects copies

- Copy objects on different medias (POSIX, tape,…)
- A copy of an object represents a copy of its data on a medium (dir, tape, …).
- Creating a copy doesn’t create a new version of an object



Phobos objects copies

- **Create a copy:**

```
phobos copy create objid1 cache
```

- **Delete a copy:**

```
phobos copy delete objid1 cache
```

- **Listing copies:**

```
phobos copy list objid1
```

→ can take same options as the phobos put (layout, family, …)

Phobos objects copies

A new option --copy-name has been added to choose on which copy to work

- Put:

```
phobos put –copy-name cache /path/to/file objid1
```

- Get:

```
phobos get –copy-name cache objid1 file.out
```

The copy name can be linked with put parameters (layouts, family, …)

3. Production case at CEA

Lustre/HSM – Phobos backend

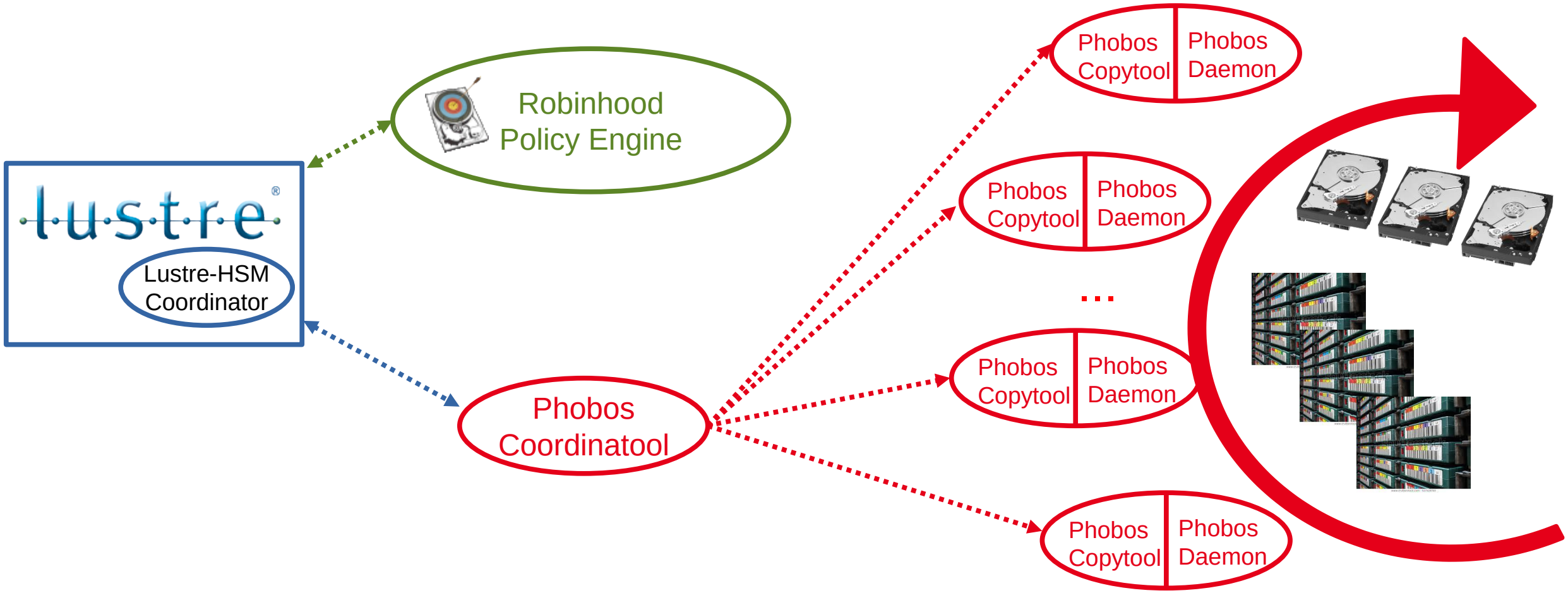
Phobos is deployed as the Lustre/HSM tape storage backend at CEA.

It is integrated into a complete toolchain, including

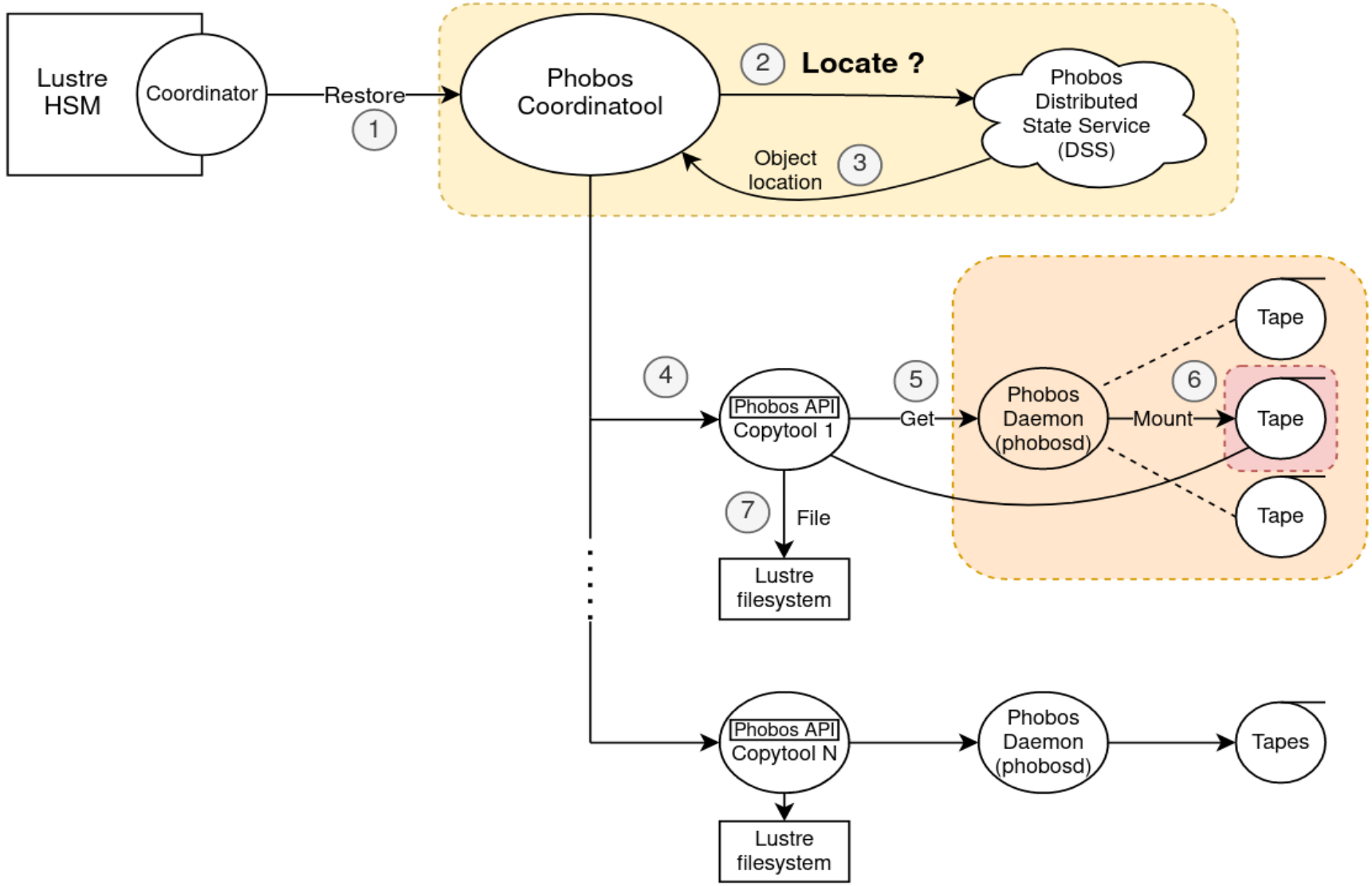
- Phobos Coordinatool: https://github.com/phobos-storage/coordinatool
- Phobos Copytool: https://github.com/phobos-storage/lustre-hsm-phobos
- Robinhood Policy Engine: https://github.com/cea-hpc/robinhood

This toolchains will also be used for the future Lustre/HSM tape backend of new EXA and Alice Recoque supercomputers

Lustre/HSM – Phobos tools chain



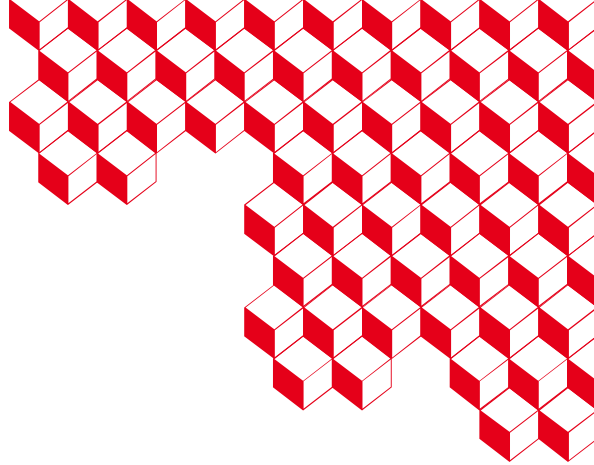
Lustre/HSM – Phobos tools chain



Contribution

Interested?

- Start here: https://github.com/phobos-storage
- Documentation and tutorials are available in Phobos's wiki
- Contributions are welcome, as well as feedback!



Thank you !

Gauthier EVRAERD, gauthier.evraerd@cea.fr

Sébastien GOUGEAUD, sebastien.gougeaud@cea.fr

Patrice LUCAS, patrice.lucas@cea.fr

Yoann VALERI, yoann.valeri@cea.fr