

The logo for CEA (Commissariat à l'énergie atomique et aux énergies alternatives) consists of the lowercase letters 'cea' in a white, rounded, sans-serif font. A horizontal green line is positioned below the letters. The logo is set against a white background within a dark red square frame.The Phobos logo features a stylized, multi-colored arc (yellow, orange, red) with a teal circle at its top end. The word 'Phobos' is written in a large, black, sans-serif font to the right of the arc. The entire logo is contained within a light red square frame.

# Phobos: an open-source scalable object store with tape libraries support

Patrice LUCAS, [patrice.lucas@cea.fr](mailto:patrice.lucas@cea.fr)

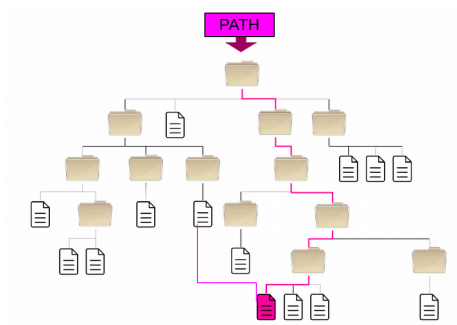
## Next scale of mass storage

- Exaflop supercomputers in the 2020's
- Huge amounts of data to ingest: petabytes per day
- Huge amounts of data to store: exabytes

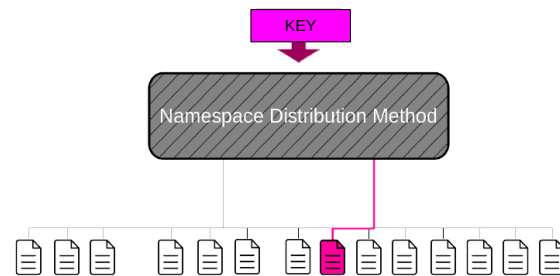


	Today	Tomorrow
Daily production	Hundreds of TB	Petabytes
Storage system capacity	Hundreds of PB	Exabytes

# Suppressing POSIX filesystem's bottlenecks



*Addressing entries  
in a POSIX file system*



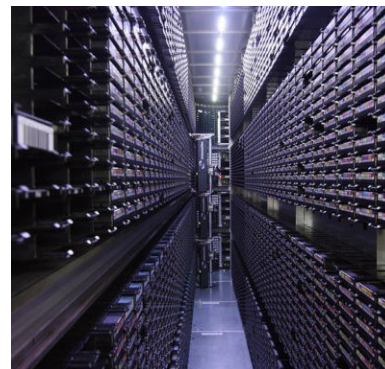
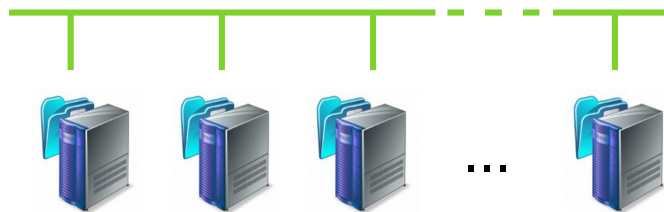
*Addressing entries  
in an object-store*

- Object stores have proved their scalability
- Widely adopted for Internet services, Cloud computing, social networks...



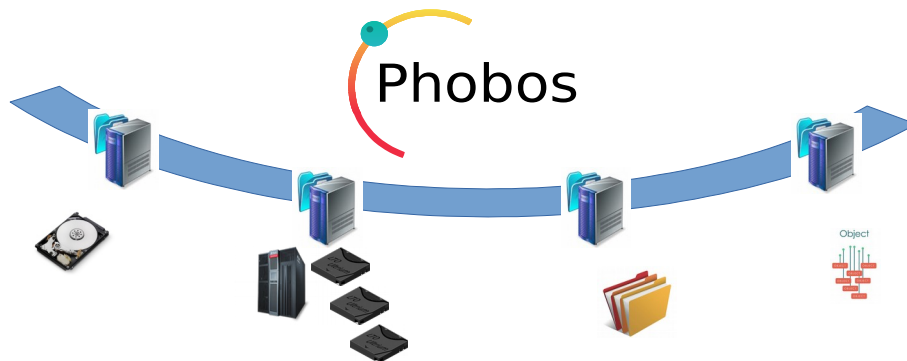
# Needs for extremely scalable storage systems at a reasonable price

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



# Phobos: Parallel Heterogeneous Object Store

- Goal : manage a distributed set of storage resources on different storage media technologies
  - Hard drives, magnetic tapes, file servers, object stores...
  - Optimized I/O for each storage technology
    - E.g. minimize mounts and data sync for tapes



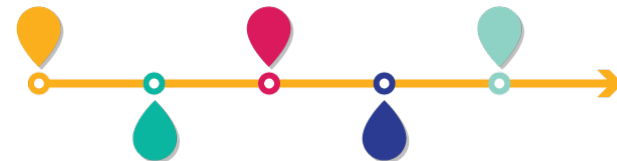
## Design guidelines

- Scalability and fault-tolerance
- Based on open formats, open protocols, interoperable
  - E.g. LTFS as tape filesystem (ISO/IEC 20919:2016)
- Simple and common interfaces (REST, object stores API)
- Simple administration (intuitive, admin-friendly CLI)
- Light, easy to deploy, easy to maintain
  - As of today: 14k lines of C, 2.5k lines of Python

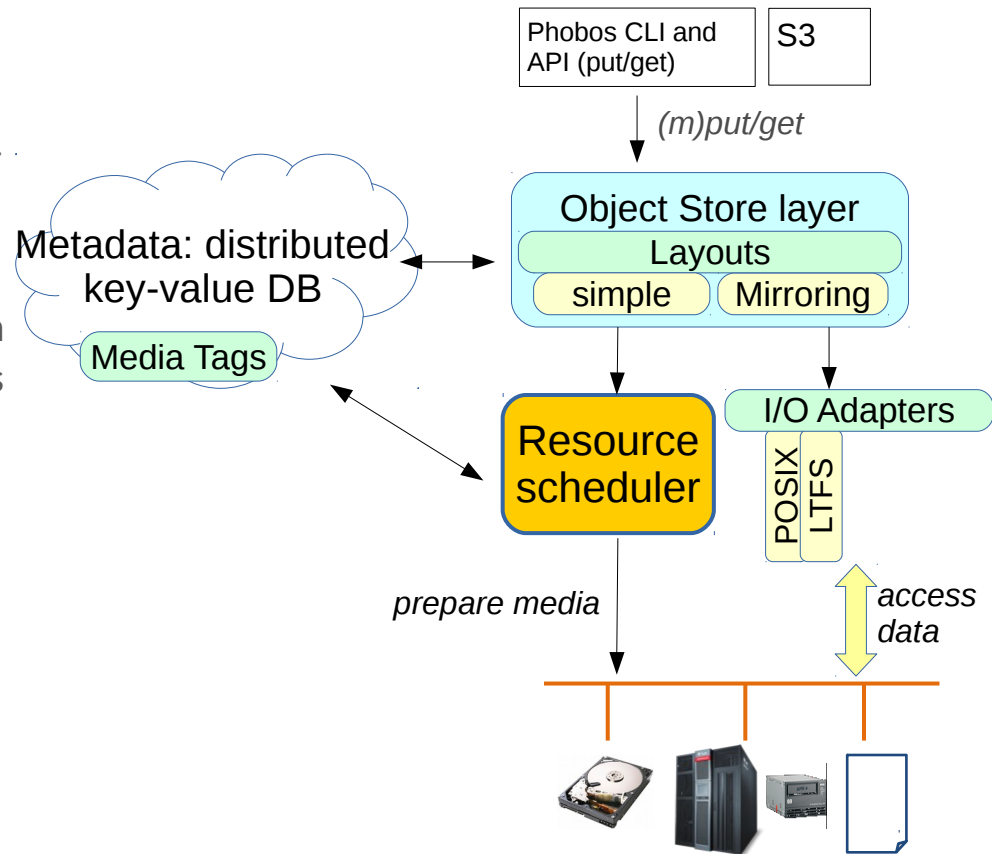


# History of the project

- 2013: first ideas
- 2014-2015: development of the initial version  
Scope:
  - Storage on tape, or in a filesystem
  - SCSI-controlled tape library and LTO drives
  - Single server
- 2016: Phobos v1.0 in production
  - Multi-Petabyte storage of genomics data
  - IBM TS3500 library, LTO5/6 drives
- 2019: Phobos made available on github as open-source (LGPL v2.1)
- 2020-2021: Working on the distributed version (v2.0)

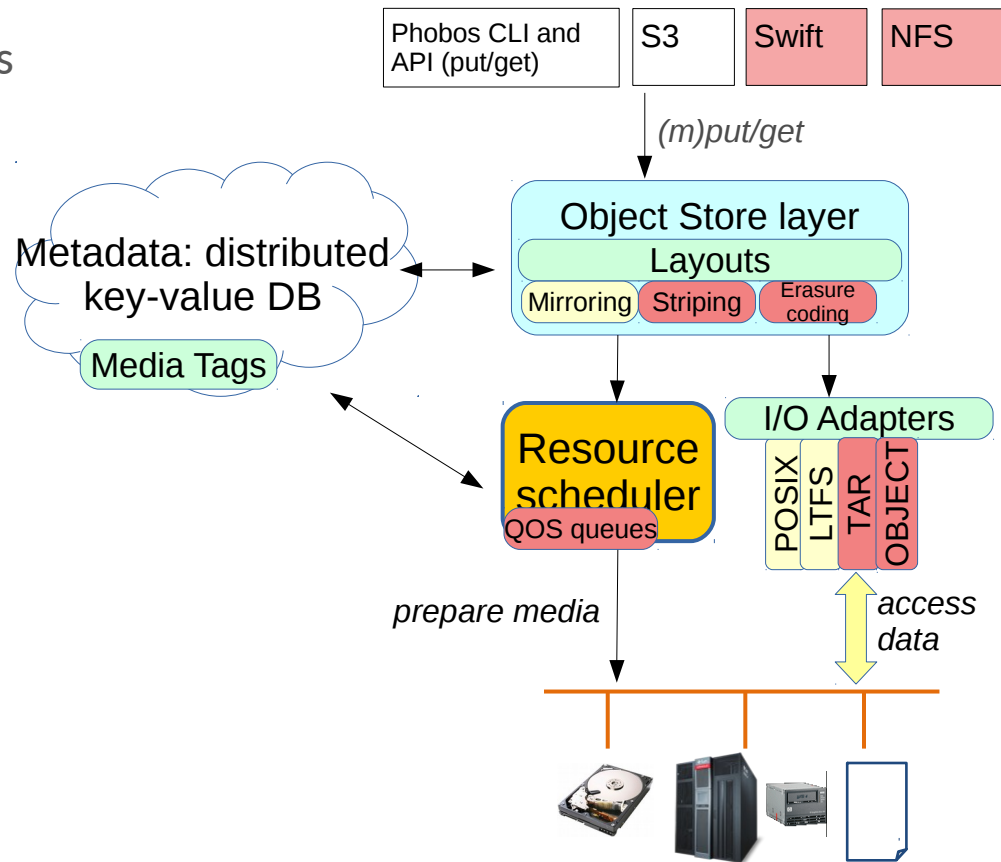


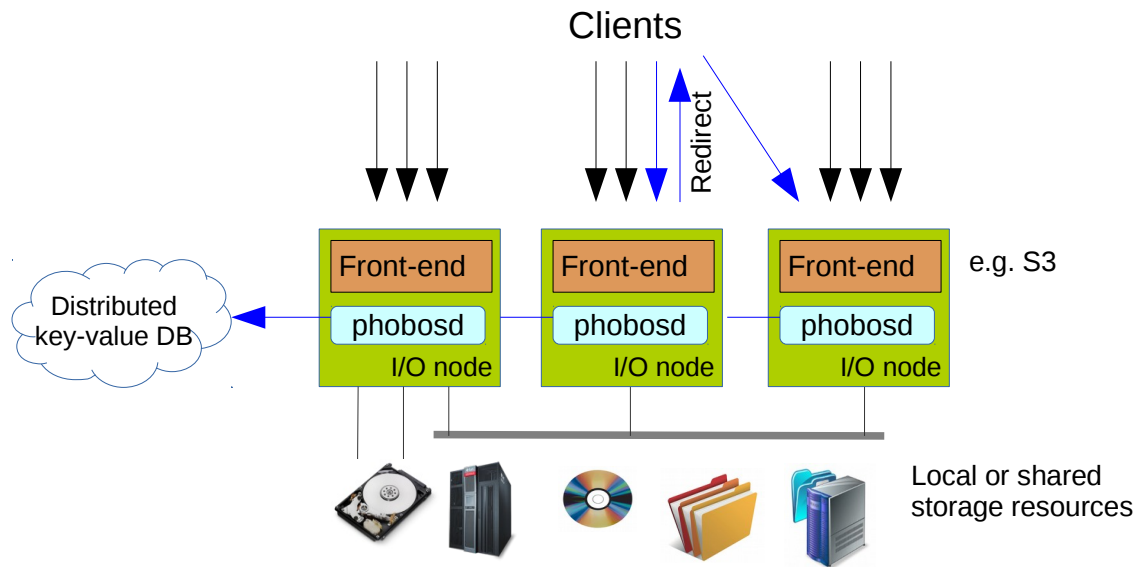
- **IO adapters:** multiple storage technologies (Posix, LTFS)
- **Layout plugins:** performance and fault-tolerance
- **Tags:** storage partitioning
- **Resource scheduling:** optimizes stream to tape drives, minimizes tapes mounts
- **Front-ends:** native API and S3
- Object versioning, undelete feature...
- **Key-value metadata schema:**
  - DB schema is NoSQL-ready
  - Currently uses PostgreSQL: can be parallelized thanks to sharding features
  - Saved within objects on media (recovery, tape import)





- Providing Swift and NFS connectors
- New layouts: striping, erasure-coding
- New IO adapters: NVMe, Object stores
- Media life cycle management: smart tape repacking policies, automatic migration between storage technologies
- Optimizing resource scheduler policies: prioritizing and grouping I/Os





- Synchronization in the distributed mode:
  - Through the distributed key-value DB (object location, resource reservation...)
  - Redirection of client requests to the preferred I/O node (max 1 hop)

- **Version number**
  - `phobos put -overwrite file existing_oid`
  - `Phobos object list --deprecated`
- **Generation number (UUID in addition of OID)**
  - `phobos delete existing_oid ; phobos put file existing_oid`
  - `phobos delete existing_oid ; phobos undelete deprecated_oid`
- **Getting deprecated object**
  - `phobos get --uuid aabbccdd --version 2 obj0123 /tmp/obj0123.back`

## Setting up a tape storage in a couple of commands

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

```
phobos tape add -t lto6 [073200-073222]L6
```

```
phobos tape format --unlock [073200-073222]L6
```

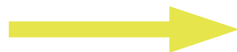
*That's done! Your system is ready for I/Os.*

# Example of deployments



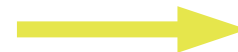
- Multi-petabyte genomics datasets
- In production since 2016

## DNA sequencers



## Phobos

- IBM TS3500 tape library (SCSI)
- LTO6 and LTO8 drives

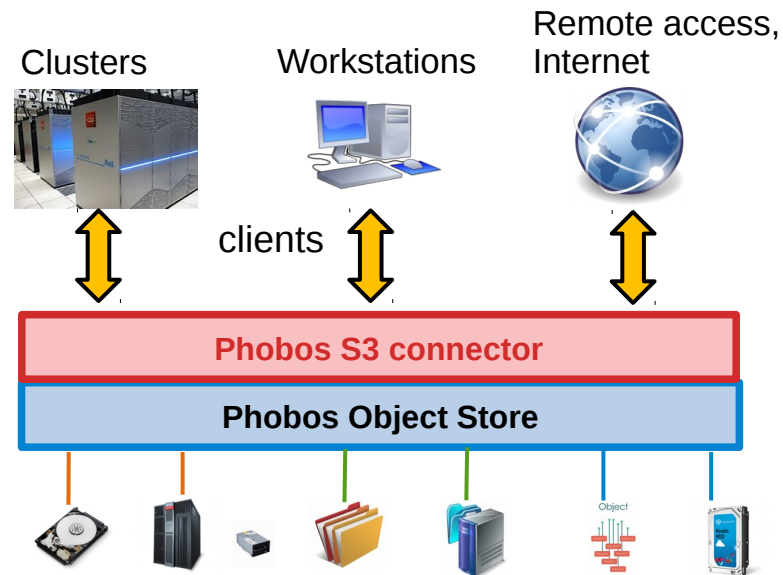


## HPC data clusters



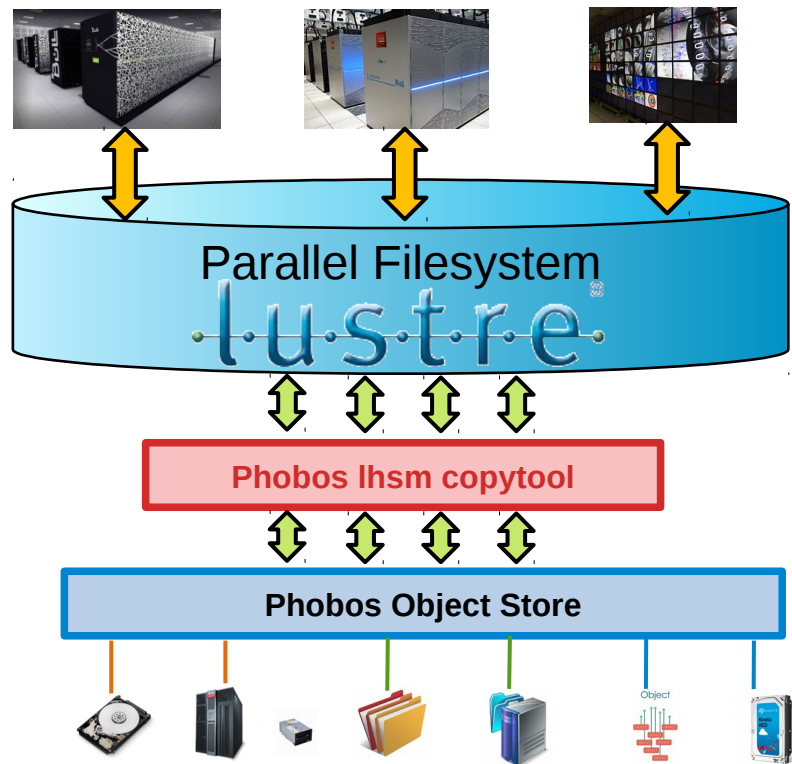
## Object store with an S3 interface

- S3 interface exposed to end-users
- Phobos: high-performance, scalable storage
  - Can manage a wide variety of capacitive storage, including tape libraries
  - Provides an easy/uniform management of these storages



## Lustre HSM backend

- Lustre: filesystem user front-end
- Phobos as capacitive backend (hierarchical storage)





- Collaboration with DDN and ICHEC:
  - Implementing a S3 server for Phobos
  - Integrating Phobos as Lustre-HSM backend
  - Contributions to the core of Phobos
  
- Collaboration with Atos, ECMWF, ICHEC, Seagate, Univ. of Mainz
  - In the framework of the EuroHPC project “IO-SEA”
  - Building a storage software stack for Exascale systems
  - Phobos used as the long-term storage component
  - New developments: scalability enhancements, erasure coding, media lifecycle management, administrative interface, LTFS tape import, smart tape request reordering, front-ends (Swift, POSIX)...



## Summary

- Tape object storage at scale (and more)
- Phobos is open-source, available on github:  
<https://github.com/cea-hpc/phobos>
- Contributions are welcome, as well as testers!



The logo for CEA (Commissariat à l'énergie atomique et aux énergies alternatives) features the lowercase letters 'cea' in a white, sans-serif font. A thin green horizontal line is positioned directly beneath the letters. The logo is set against a dark red background that has a faint, repeating pattern of small white circles.

DE LA RECHERCHE À L'INDUSTRIE

The Phobos logo is contained within a semi-transparent light pink square. It features a stylized orbital path represented by a curved line that transitions from red at the bottom to yellow at the top. A small cyan circle with a black dot in the center is positioned at the top of the curve, representing the planet Phobos. The word 'Phobos' is written in a large, black, sans-serif font to the right of the orbital path.

<https://github.com/cea-hpc/phobos>

**Thank you for your attention!**