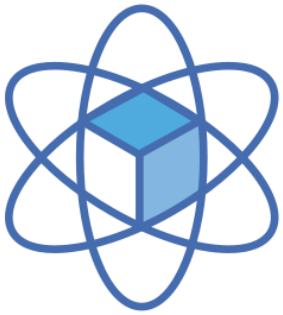


Testing CERNBox

Jakub T. Mościcki, IT/DSS

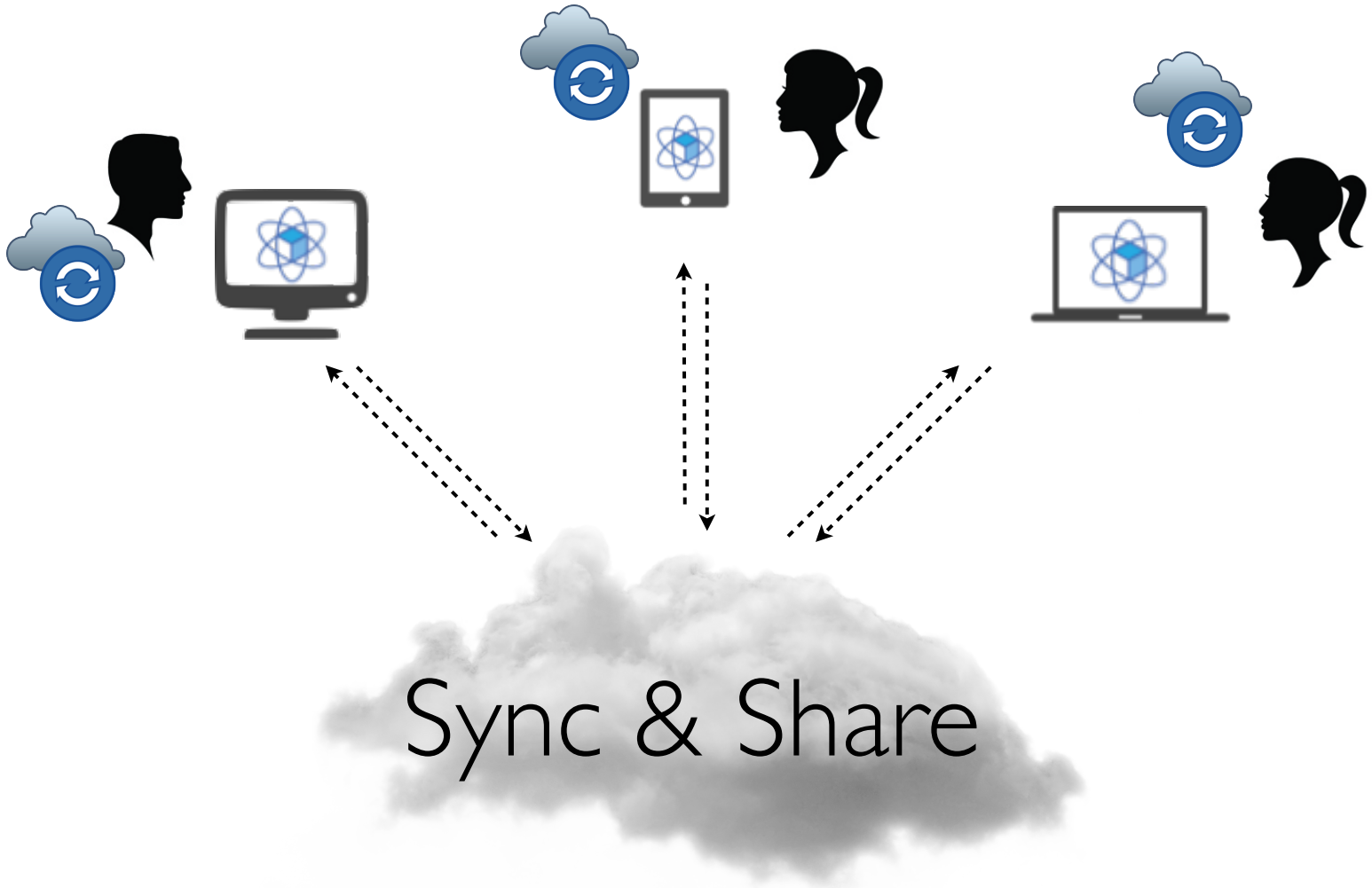
1st Developers at CERN Forum

28 September 2015

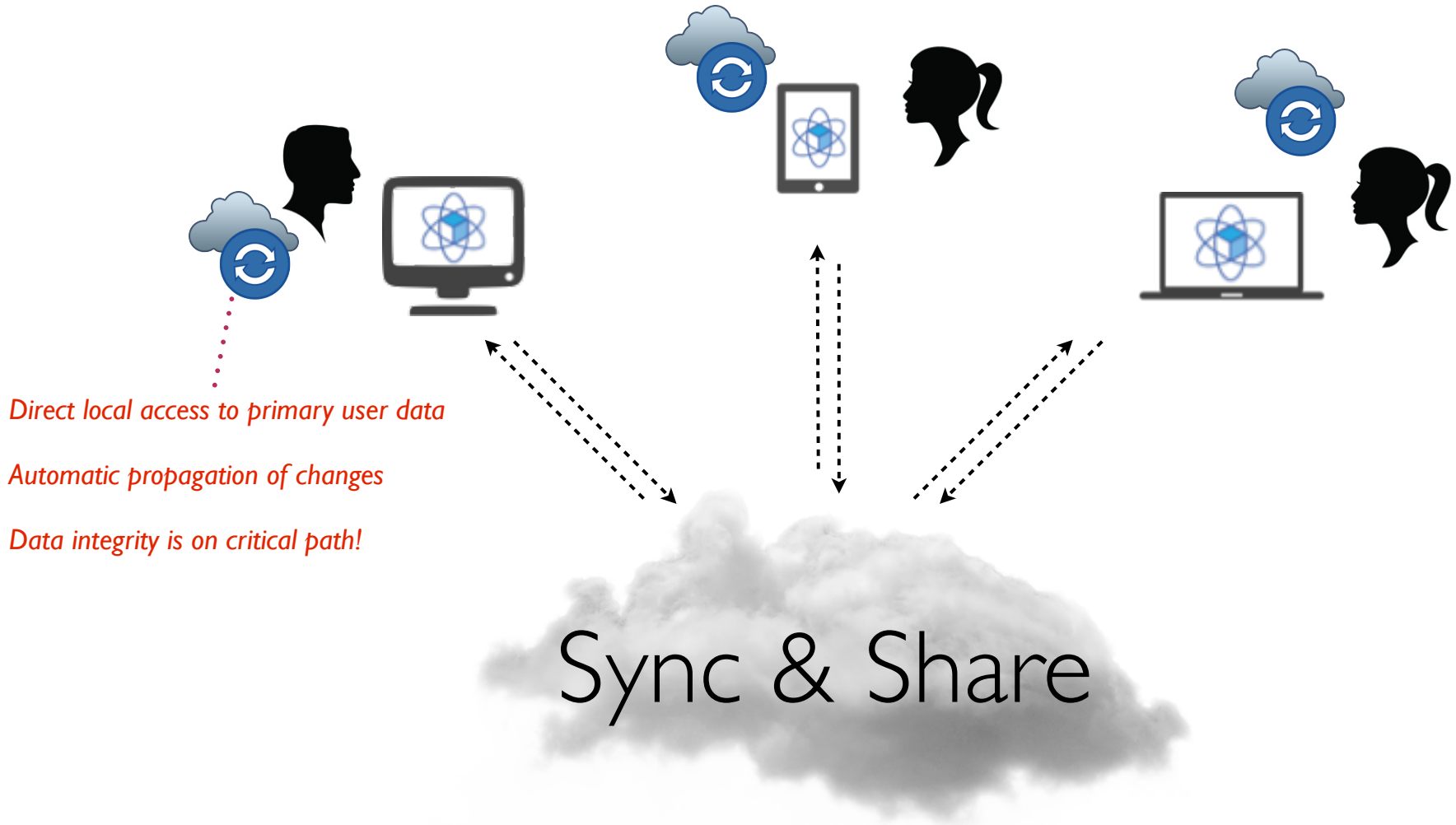
Setting the scene

- Why testing is critical for CERNBox...
- How we do it
- Experience & examples

Key Functionality

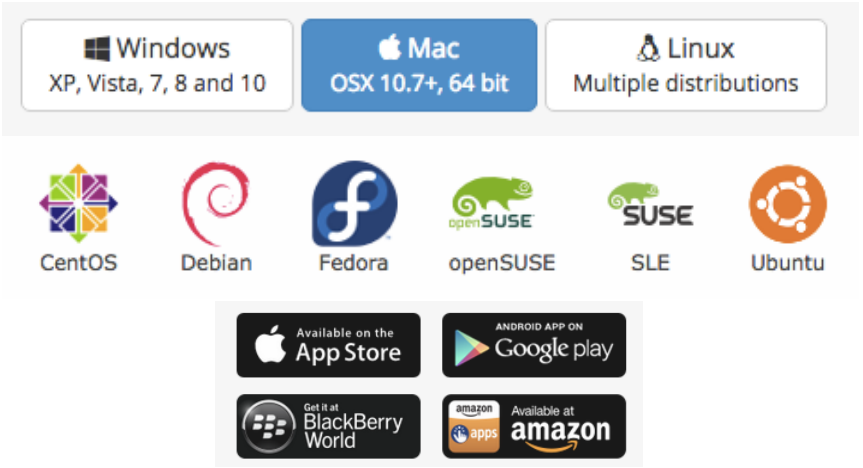


Key Functionality



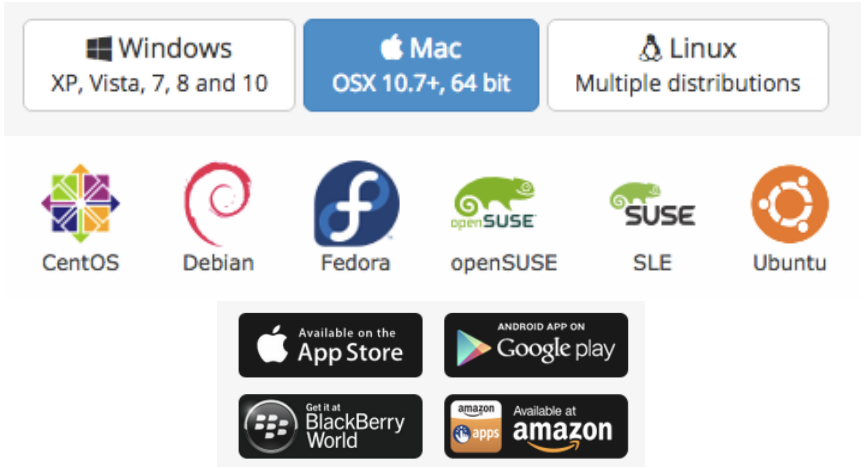
Multiplatform support

- Windows, MacOSX, N x Linux
  - case sensitive vs case-preserving
  - file locking semantics
  - filename limitations
  - bundles, junctures, device files
  - ...

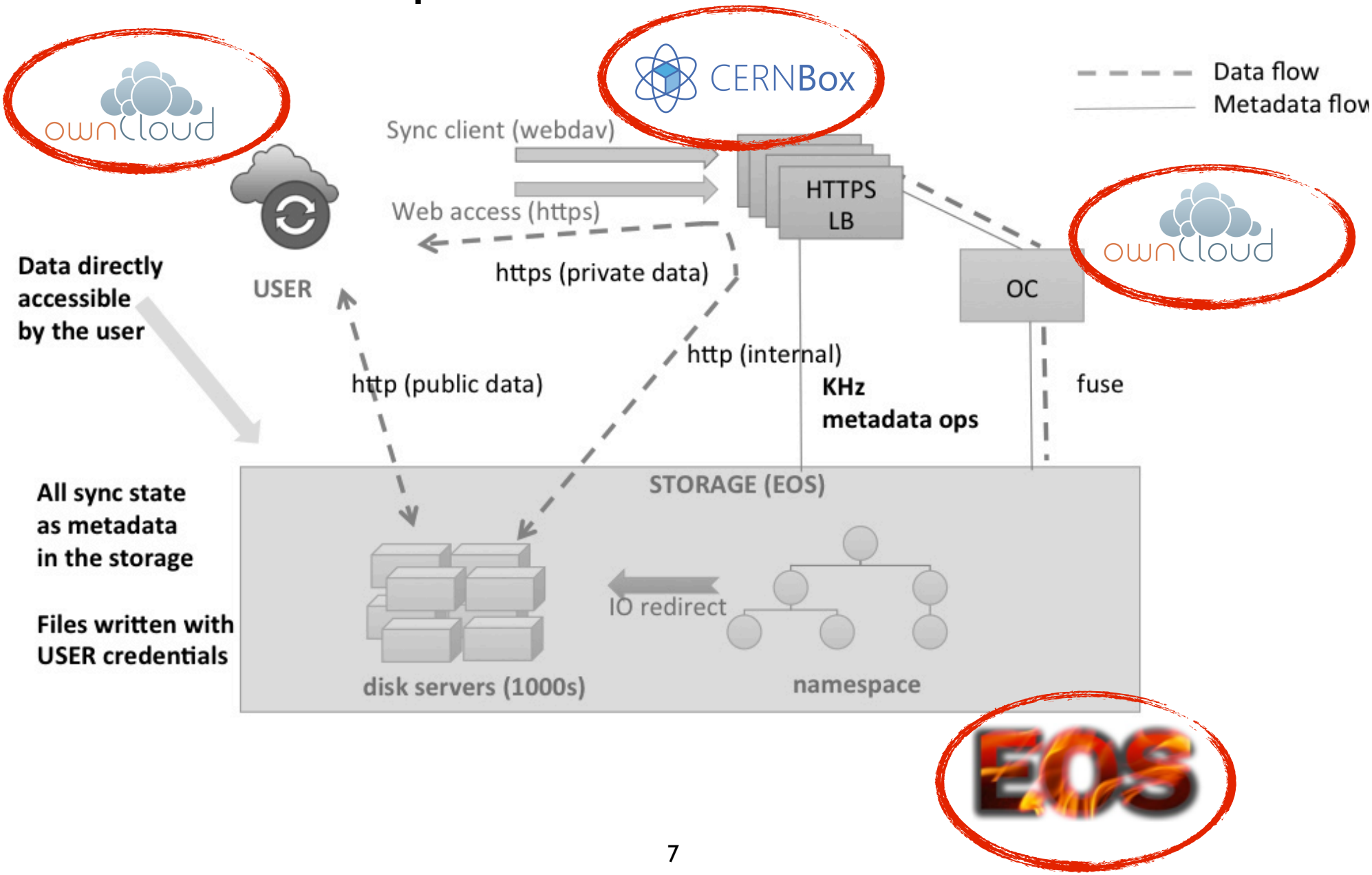


Multiplatform support

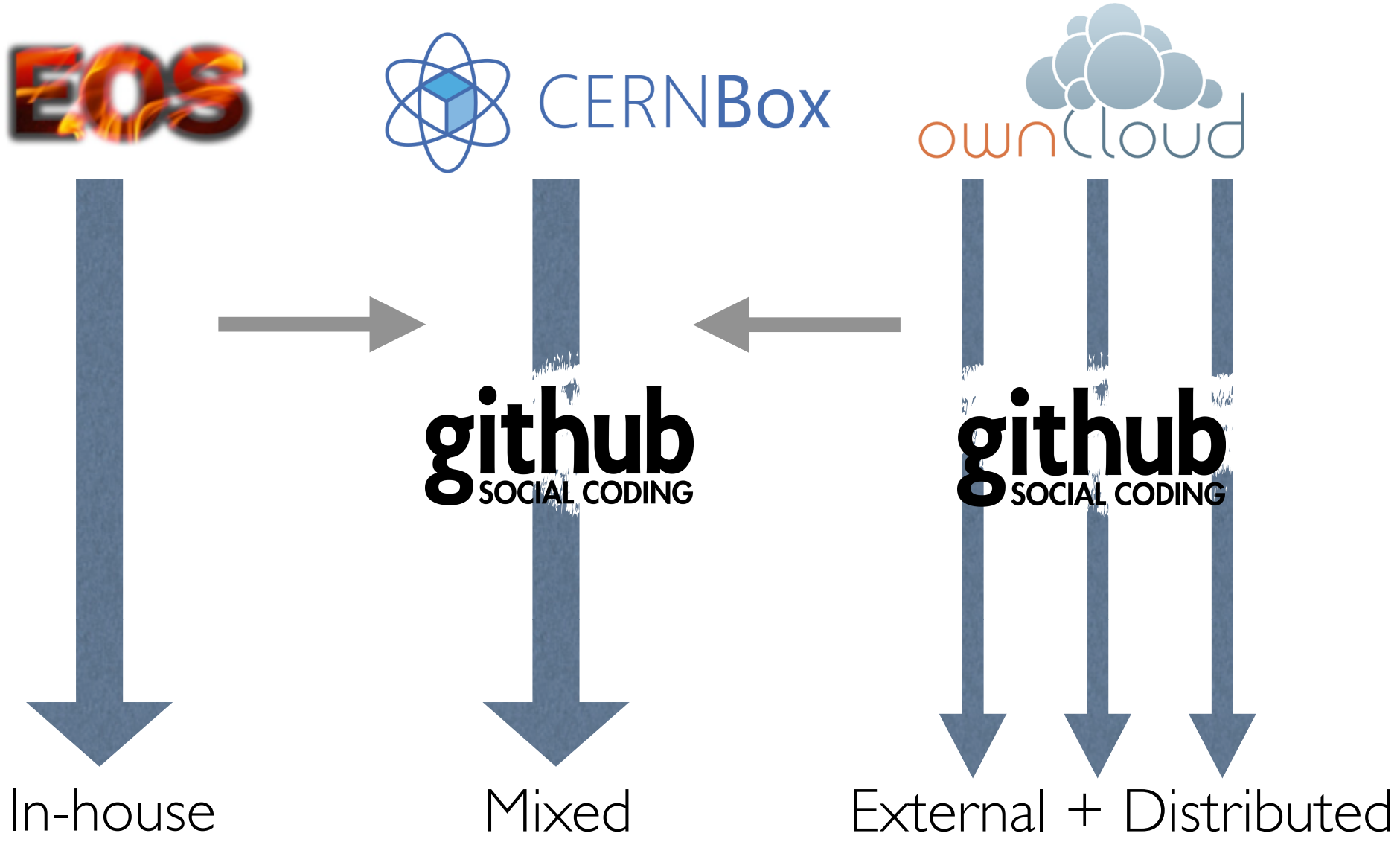
- Windows, MacOSX, N x Linux
  - case sensitive vs case-preserving
  - file locking semantics
  - filename limitations
  - bundles, junctures, device files
  - ...
- Shaky "Airport WiFi"
- Non-standard setups
  - Fuse mounted NTFS parition on Ubuntu/Windows dual boot system
  - NTFS hardlinks to represent 8.3 files
  - *.eml files constantly updated by Windows indexer behind the scenes
  - ...



Components involved



Processes involved



How do we test the synchronizer?

- Synchronizers are tricky
  - it's a background service: not driven by user commands (e.g. "git commit")
  - no clear transaction boundaries e.g.:
    - has user finished modifying a directory?
  - unclear semantics in general
    - unresolvable conflicting actions

Mysteries of Dropbox

- Benjamin Pierce's talk+video at CS3 Workshop 2014
  - https://indico.cern.ch/event/336753/session/1/contribution/28
- Erlang-based black-box testing to extreme
  - do not write tests: **generate them**
  - runtime check against formal specification
- Used in RT-systems testing (e.g. airplanes)
- Very clever and quite complex


smashbox

Smashbox

- Controlled test scenarios
  - explicit sync points (owncloudcmd)
  - easy to understand and write new tests
  - concurrency made easy
  - automation and parameter sweeps
    - e.g. many time sequences of the same scenario
- https://github.com/cernbox/smashbox

Example: simple consistency check

https://github.com/cernbox/smashbox/blob/master/lib/test_nplusone.py

```
@add_worker
def workerA(step):
    step(1,'Preparation')
    d = make_workdir()
    run_ocsync(d)

    step(2,'Add nfiles')
    for i in range(nfiles): create_hashfile(d,size=filesize)
    run_ocsync(d)

@add_worker
def workerB(step):

    step(1,'Preparation')
    d = make_workdir()
    run_ocsync(d)
    k0 = count_files(d)

    step(3,'Resync and check files added by workerA')
    run_ocsync(d)
    k1 = count_files(d)
    error_check(k1-k0==nfiles,'Wrong number of files')
    ncorrupt = analyse_hashfiles(d)[2]
    fatal_check(ncorrupt==0, 'Corrupted files (%s) found'%ncorrupt)
```

Example: simple consistency check

https://github.com/cernbox/smashbox/blob/master/lib/test_nplusone.py

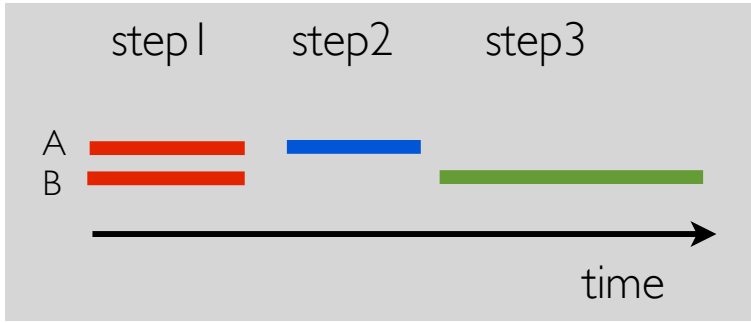
```
    @add_worker
def workerA(step):
    step(1,'Preparation')
    d = make_workdir()
    run_ocsync(d)

    step(2,'Add nfiles')
    for i in range(nfiles): create_hashfile(d,size=filesize)
    run_ocsync(d)

@add_worker
def workerB(step):

    step(1,'Preparation')
    d = make_workdir()
    run_ocsync(d)
    k0 = count_files(d)

    step(3,'Resync and check files added by workerA')
    run_ocsync(d)
    k1 = count_files(d)
    error_check(k1-k0==nfiles,'Wrong number of files')
    ncorrupt = analyse_hashfiles(d)[2]
    fatal_check(ncorrupt==0, 'Corrupted files (%s) found'%ncorrupt)
```



Status & Contributions

- Around 25 test cases currently implemented
- Tests contributed by 3rd parties
  - e.g. SWITCH, ownCloud
- Pull requests with core framework improvements
  - e.g. DelC, ownCloud

How we use it

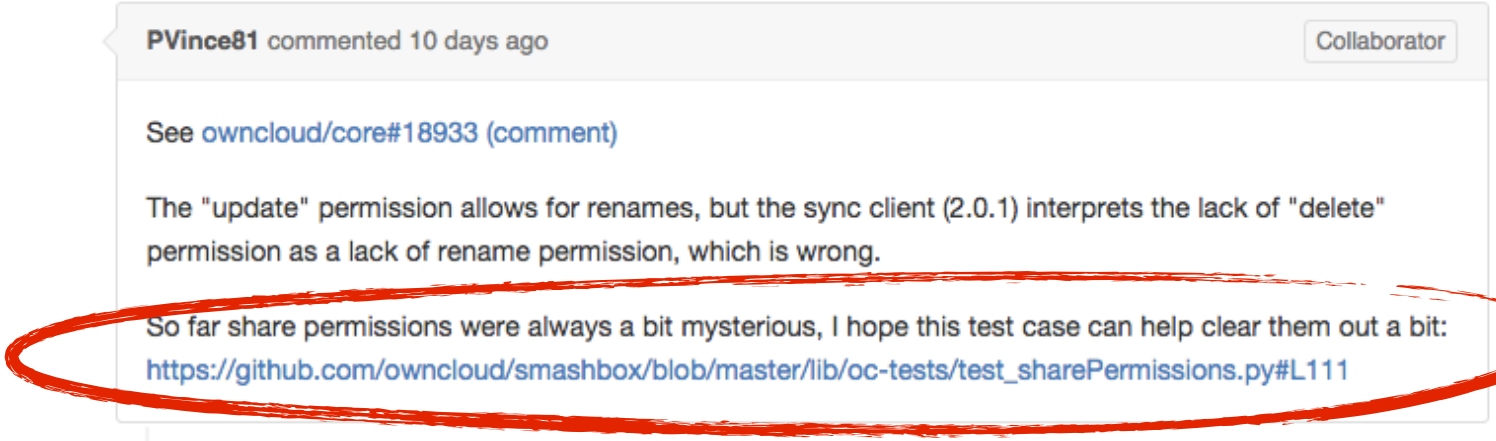
- Continous SLS-style test of cernbox.cern.ch
  - `nplusone`
- Smoke testing
  - software upgrades (client/server)
  - new CERNBox deployments
    - e.g. Taipei, Melbourne
- Performance/stress testing
  - Denmark (DeIC)

`nplusone` in practice

- **`nplusone`** detected first 16K block of a file copied twice into random offset with 1e-5 probability
- **Qt** bug in HTTP/SSL layer revealed and fixed
  - https://codereview.qt-project.org/#/c/110150/7
- Conclusion: do not take for granted even well established and popular libraries!

Lingua franca on github?

- Make it easy to communicate bugs
  - more efficiently: just refer to the test case
  - removes ambiguity how to reproduce it
  - automatically verify in the future
- Example:



Where to improve

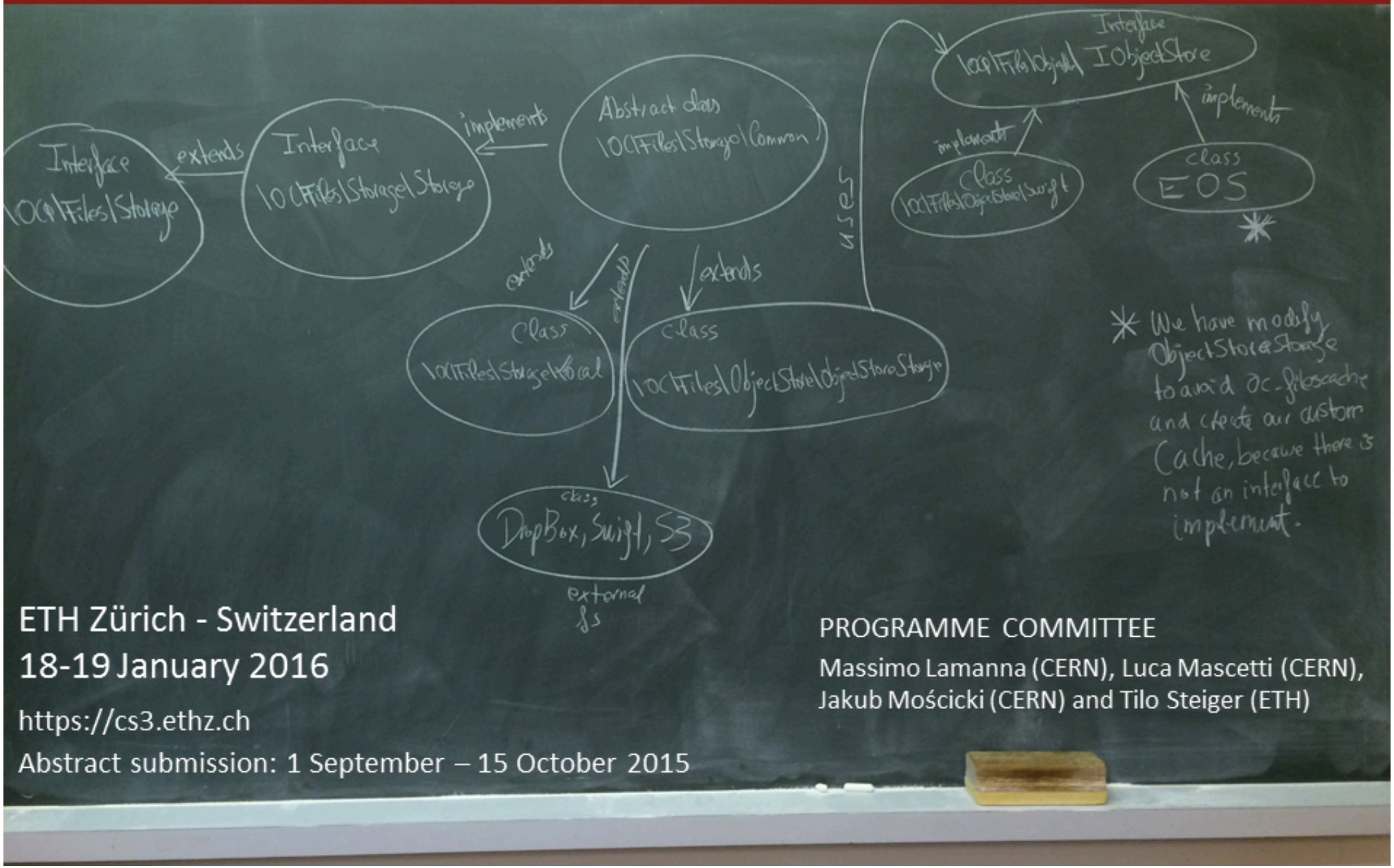
- We don't have expertise on Windows
  - IO patterns by MS Office applications!
- Smashbox on mobile platforms?
- Good ideas appreciated
- Volunteers appreciated :-)

More pointers

- **https://github.com/cernbox/smashbox**
- https://owncloud.org/blog/smashbox-in-action
- https://indico.cern.ch/event/336753
- http://cs3.ethz.ch

Cloud Services for Synchronisation and Sharing (CS3)

Cloud Storage Services for Novel Applications and Workflows



ETH Zürich - Switzerland

18-19 January 2016

https://cs3.ethz.ch

Abstract submission: 1 September – 15 October 2015

PROGRAMME COMMITTEE

Massimo Lamanna (CERN), Luca Mascetti (CERN), Jakub Mościcki (CERN) and Tilo Steiger (ETH)