FuS3FS
--A Cloud Filesystem

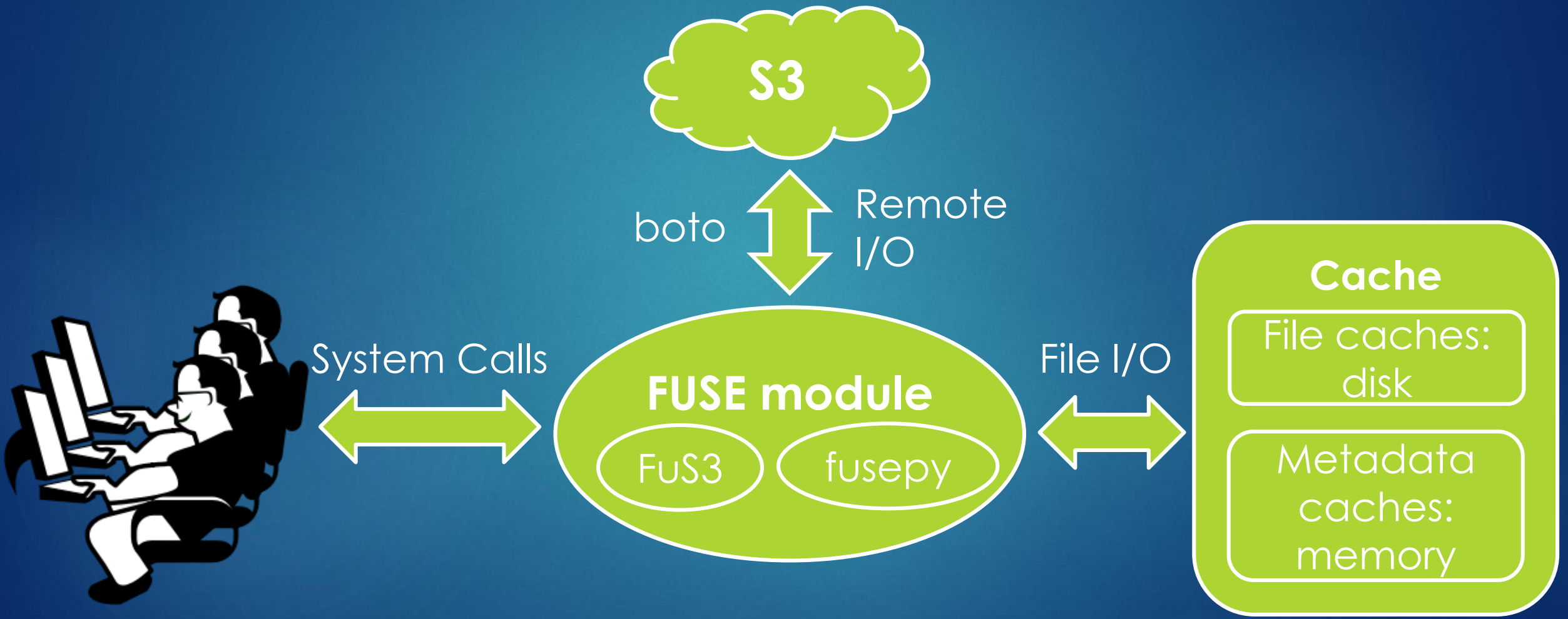
Kaijun Feng & Chao Luo

4/18/2016

Project goal

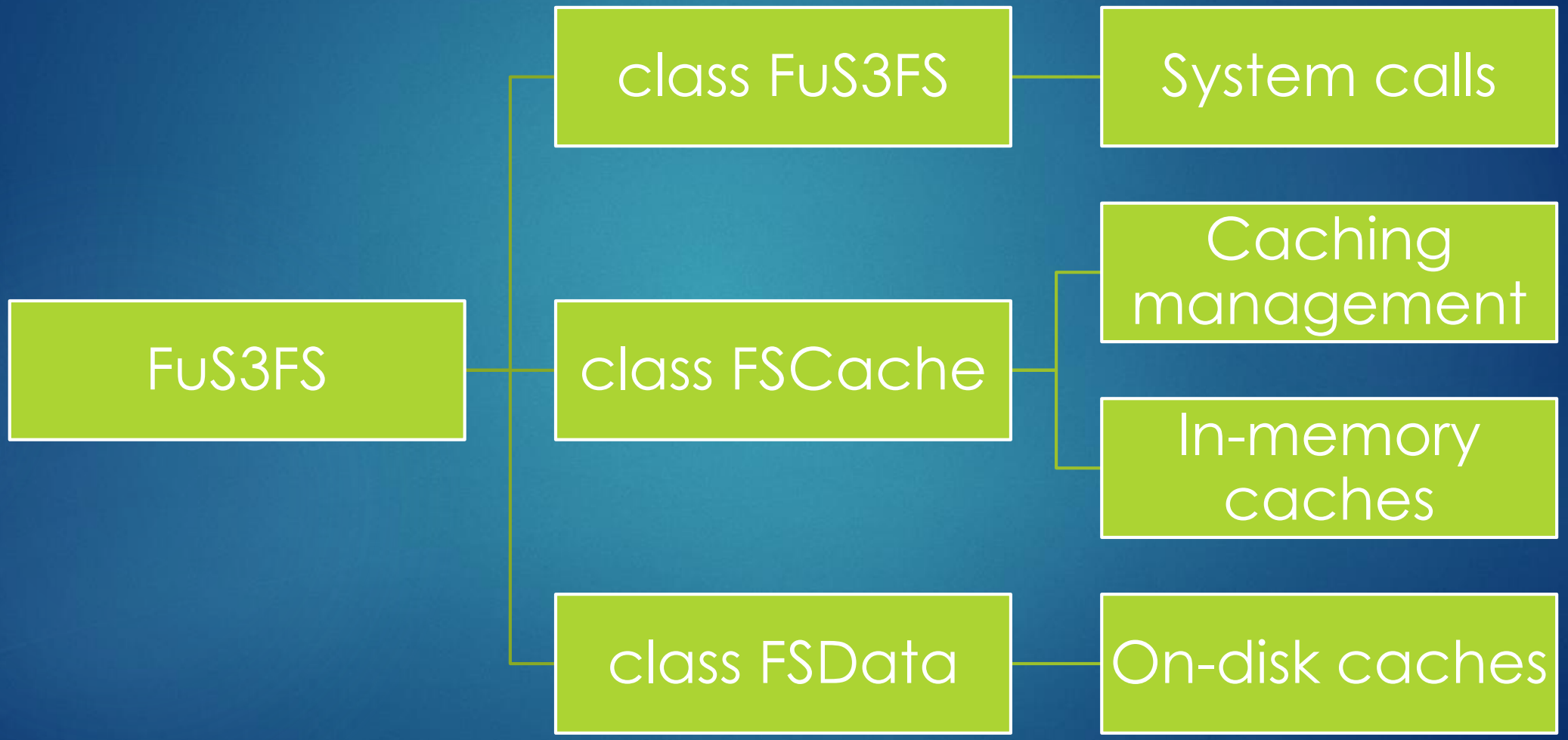
- Build a filesystem in the cloud that can be used by multiple users simultaneously.
- Works as a conventional filesystem for users, but files are stored on Amazon S3.

System diagram



- FUSE = Filesystem in Userspace

Code structure



Directory tree (readdir)

- With cache:

```
return dict[path]['dirs']
```

- Without cache:

```
self.s3 = boto.connect_s3()
self.s3_bucket = self.s3.get_bucket(self.s3_bucket_name…)
key_list = self.s3_bucket.list(full_path…)
//cache to dict[path]['dirs'] afterwards
```

Code snippet: open() in class FuS3FS

```
def open(self, path, flags):
        logging.debug("open '%s' '%i'" % (path,
flags))
        with self.cache.get_lock(path):
            self.cache.add(path)
            if not self.check_data(path):
                self.mknod(path, flags)
            self.cache.get(path, 'data').open()
        return 0
```

Example

- `./fus3fs.py s3://your-bucket /local/dir`

Directory operations:

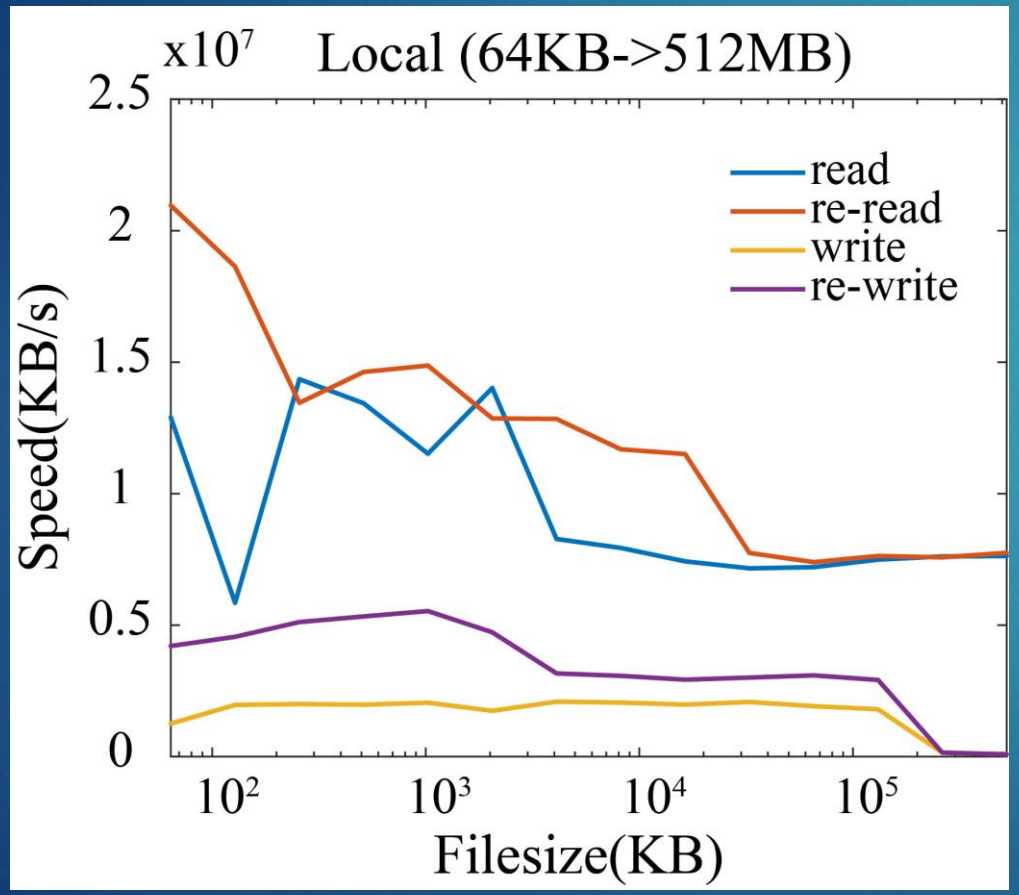
```
[ec2-user@ip-172-31-56-244 s3bucket]$ mkdir test_dir
[ec2-user@ip-172-31-56-244 s3bucket]$ ls
createDB.py      htmltowords.py  local_result.xls  read_test.py
fus3_result.xls  image           part-r-00000      test_dir
[ec2-user@ip-172-31-56-244 s3bucket]$ cd test_dir/
[ec2-user@ip-172-31-56-244 test_dir]$
```

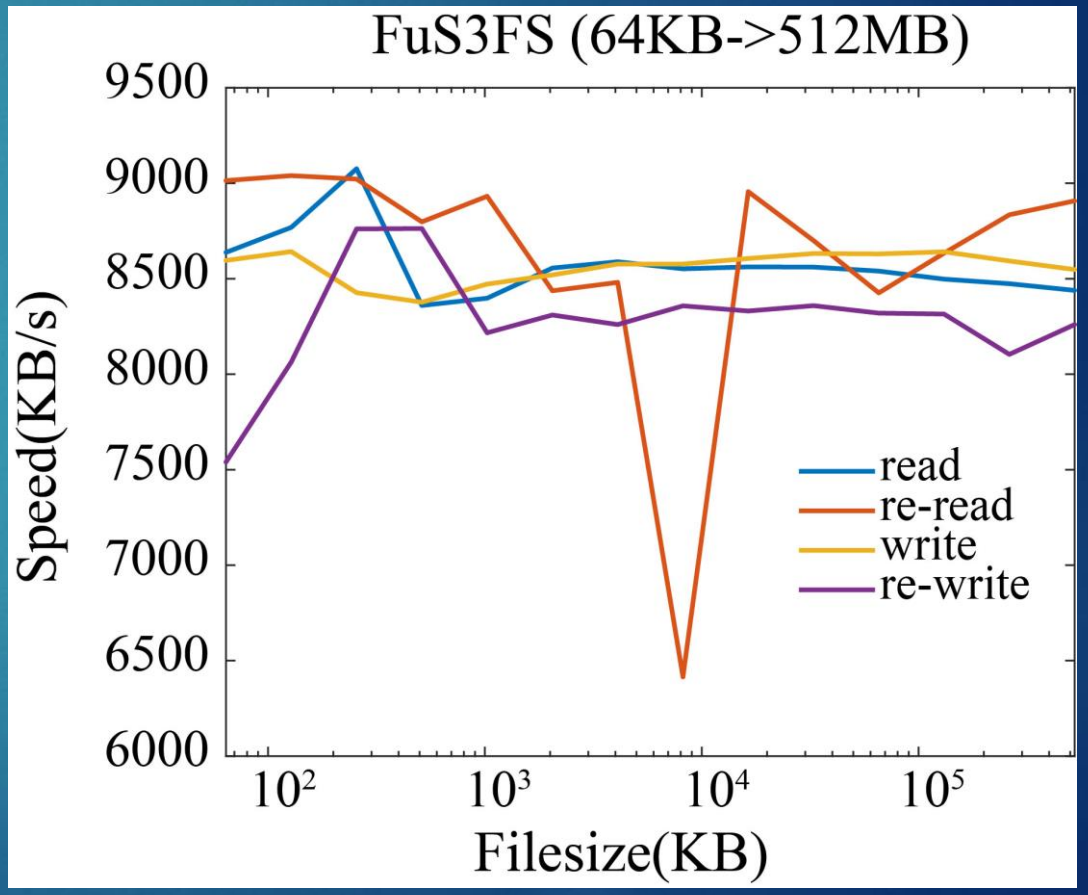
File operations:

```
[ec2-user@ip-172-31-56-244 s3bucket]$ cat read_test.py
#! /usr/bin/env python
with open('testfile','r+') as testfile:
    testfile.read()
testfile.close()

[ec2-user@ip-172-31-56-244 s3bucket]$ ls -l read_test.py
-rwxr-xr-x 1 ec2-user ec2-user 101 Apr 18 02:18 read_test.py
[ec2-user@ip-172-31-56-244 s3bucket]$
```

Performance evaluation: IOZONE





Performance evaluation: python script with 10MB file

		real (s)	user(s)	sys(s)
Write tests	Local	0.024	0.008	0.012
	FuS3 1st run	1.051	0.008	0.008
	FuS3 2nd run	0.071	0.008	0.008
Read tests	Local	0.015	0.008	0.004
	FuS3 1st run	1.624	0.008	0.004
	FuS3 2nd run	0.049	0.012	0.000

Problem & future work

- Consistency: need to invalidate cache when other users make a change to cached files.
- Proposed solution: Amazon SNS

Questions?