

EOS XDC Developments

Mihai Patrascoiu
CERN / IT- ST- AD

About XDC

- The eXtreme DataCloud – a 2 year, EU-funded software development and integration project started in February 2018
- Targets current and next generation e-Infrastructures
- Goal: improve existing Data Management Services by adding missing functionalities requested by research communities

XDC website: <http://www.extreme-datacloud.eu/>

XDC and EOS



Areas of development:

- Smart Caching
 - **XCache**
 - **Storage inclusion**
- QoS data management

(upcoming)

Involved:

- Oliver Keeble
- Andreas Peters
- Mihai Patrascoiu
- Fabrizio Furano

Smart Caching

Caching mechanism to allow transparent data access for the dynamic extension of data centers [..]

(XDC proposal)

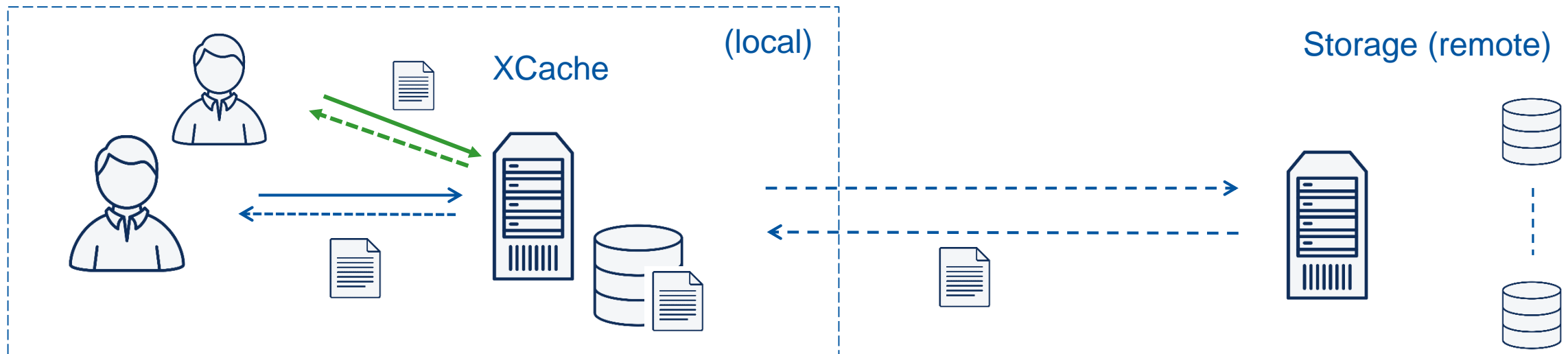
Chosen approach: XCache

- Integrate, Evaluate, Develop and Test

XCache: XRootD server plugin [PSS + PFC]

Envisioned scenario

- CPU acquisition (e.g.: from public cloud), data hosted at CERN
- Objective: maximize CPU performance
- Approach: deploy XCache server at the acquired site



Why XCache?

- Block level caching
- Easy to deploy with EOS
- Supports both XRoot and HTTP protocols
- Leverage CERN IT-ST expertise with XRootD framework
- WLCG DOMA Working Group also looking into caching [DOMA-Access]

XCache – work done

- Evaluate viability of XCache + EOS storage
- Integrate with HTTP access plugin (XrdHttp) [→ bug fixes]
- Identity forwarding plugin ([Github link](#)) [requires SSS protocol]
- Set-up authentication at XCache node:
 - Allows trusted x509 certificates + XDC VOMS group
(vo.indigo-datacloud.eu)

XCache – Identity forwarding plugin

```
190122 12:45:16 4925 mipatras.21017:26@xdc-eosdev ofs_open: 0-600 fn=/eos/xdc/testing/hello.txt
[AuthForwardClientCred] Registering sec entity: id=26 name=mipatras
190122 12:45:16 4925 XrdFileCache_Manager: info Cache::Attach() root://26@xdc-test-mgm.cern.ch:1094//eos/xdc/testing/hello.txt?oss.lcl=1
190122 12:45:16 4925 XrdFileCache_Manager: debug Cache::GetFile /eos/xdc/testing/hello.txt
190122 12:45:16 4925 root oss_Open_ufs: fd=2048 flags=0 mode=600 path=/cache/eos/xdc/testing/hello.txt.cinfo
190122 12:45:16 4925 XrdFileCache_IO: info IOEntireFile::initCachedStat successfully read size from info file = 222 root://26@xdc-test-mgm.cern.ch:1094//eos/xd
dc/testing/hello.txt?oss.lcl=1
190122 12:45:16 4925 root oss_Open_ufs: fd=2048 flags=2 mode=600 path=/cache/eos/xdc/testing/hello.txt
190122 12:45:16 4925 root oss_Open_ufs: fd=2049 flags=2 mode=600 path=/cache/eos/xdc/testing/hello.txt.cinfo
190122 12:45:16 4925 XrdFileCache_File: debug Read existing info file. /eos/xdc/testing/hello.txt
190122 12:45:16 4925 XrdFileCache_Manager: debug Cache::inc_ref_cnt /eos/xdc/testing/hello.txt
190122 12:45:16 4925 XrdFileCache_Manager: debug Cache::Attach() root://26@xdc-test-mgm.cern.ch:1094//eos/xdc/testing/hello.txt?oss.lcl=1 location: <deferred
open>
```


Drawbacks

- XCache doesn't allow writes
- Proxy server always uses his identity when fetching files from storage
- XCache doesn't evict stale files

Drawbacks – Upcoming work

- XCache doesn't allow writes
 - Work in progress
- Proxy server always uses his identity when fetching files from storage
 - Identity forwarding plugin – revamp with XRootD 4.9 changes
- XCache doesn't evict stale files
 - Incorporate file notifications (inotify over SSE [slide 25])

Storage inclusion

EOS will absorb external storage resources, of diverse types (including HTTP), and operate them as a cache

(XDC proposal)

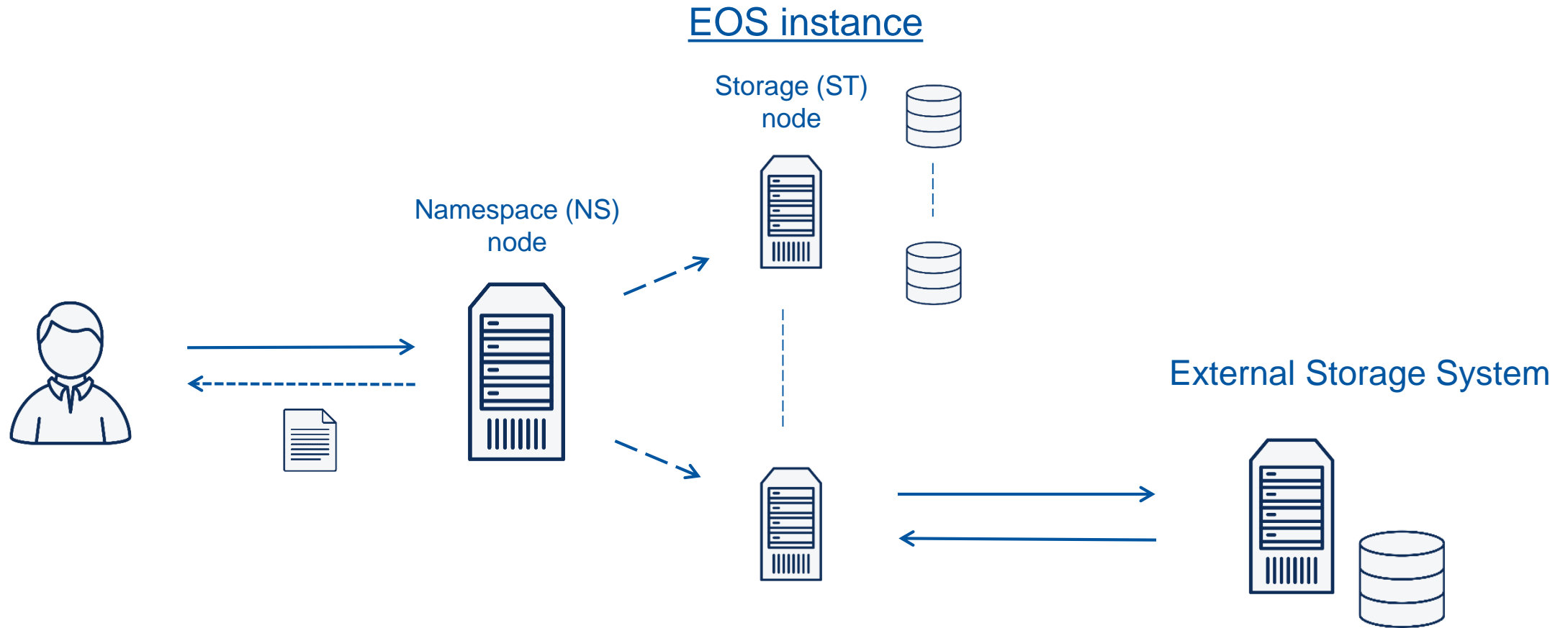
Chosen approach:

- Storage inclusion
- File adoption

Envisioned scenario

- Satellite sites pledge resources to a central system
- Management of storage around cloud procurements
 - The capability of the central system can be used, thus replica and placement policies respected uniformly
- Data ingestion from external systems

Storage inclusion - overview



Storage inclusion

- Use storage hosted externally through different protocols, such as S3, WebDav and XRoot
- Operations on the provided storage done exclusively through EOS
- Implemented by allowing filesystems to map to these protocols:

s3:// | http:// | https:// | root://

- File access mechanisms brought up to date
- Credential handling

Storage inclusion

```
fs add <uuid> <host:port> <s3://url | http://url | https://url | root://url >
[<schedgroup>] [<status>]
```

host	port	id	path	schedgroup	geotag	boot	configstatus	drainstatus	active
xdc-test-fst1.cern.ch	1095	8	/fst1_zerochecksum	checksum.0	xdc-test	booted	rw	nodrain	online
xdc-test-fst4.cern.ch	1095	5	https://dcache-xdc.desy.de/Users/mipatras/	dcache.0	xdc-test	booted	off	nodrain	online
xdc-test-fst4.cern.ch	1095	6	http://dcache-xdc.desy.de/Users/mipatras/	dcache.0	xdc-test	booted	rw	nodrain	online
xdc-test-fst1.cern.ch	1095	1	/fst1	default.0	xdc-test	booted	rw	nodrain	online
xdc-test-fst2.cern.ch	1095	2	/fst2	default.0	xdc-test	booted	rw	nodrain	online
xdc-test-fst1.cern.ch	1095	3	/fst1_lpath	lpath.0	xdc-test	booted	rw	nodrain	online
xdc-test-fst1.cern.ch	1095	17	/fst1_lpath/eos_lpath	lpath_test.0	xdc-test	booted	rw	nodrain	online
xdc-test-fst4.cern.ch	1095	4	s3://cs3.cern.ch/	s3.0	xdc-test	booted	rw	nodrain	online
xdc-test-fst2.cern.ch	1095	18	s3://s3.cern.ch/eostest/	s3_test.0	xdc-test	booted	rw	nodrain	online

File adoption

- Discover and import metadata information about files existing on external storage resources
- Imported files behave like regular EOS files
- From here on, imported files should be managed only from within EOS

Example: `eos fileinfo <path>`

no.	fs-id	host	path	physical location
0	18	xdc-test-fst2.cern.ch	s3://s3.cern.ch/eostest/	s3://s3.cern.ch/eostest/eos/s3_test/4kb.dat

Adopting files – Import functionality added to file system

```
fs import start <fsid> <src_path> <dst_path>
```

fsid = FS to perform the scan

src_path = location reachable by FS <fsid> to scan

dst_path = location within EOS namespace where to import files

Returns <import_uid>

```
fs import query <import_uid>
```

Started: [timestamp] Elapsed: [elapsed_time (h:m:s)]

Imported: [number] Failed: [number] [files/sec]

Import operation

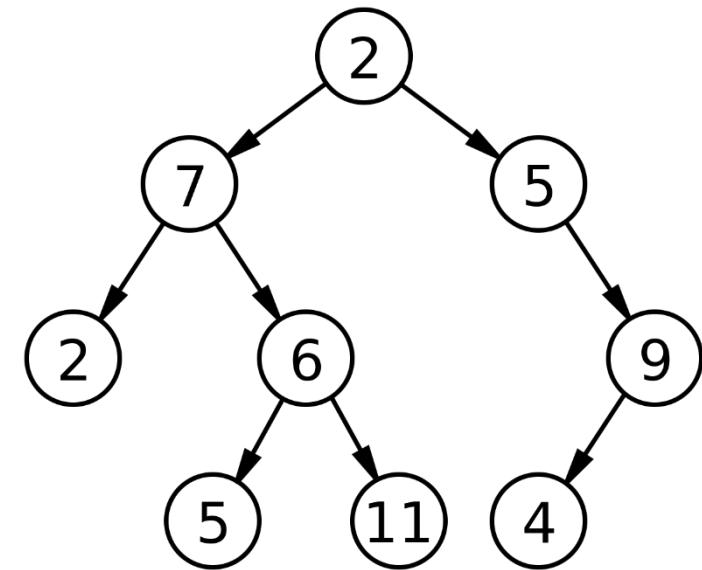
Dedicated Import thread on the FST waiting for start signal.

Once started:

- Traverses the root in a breadth-first manner
- In each directory, lists all the files
- For each file, a stat is performed
- Information is sent to the MGM

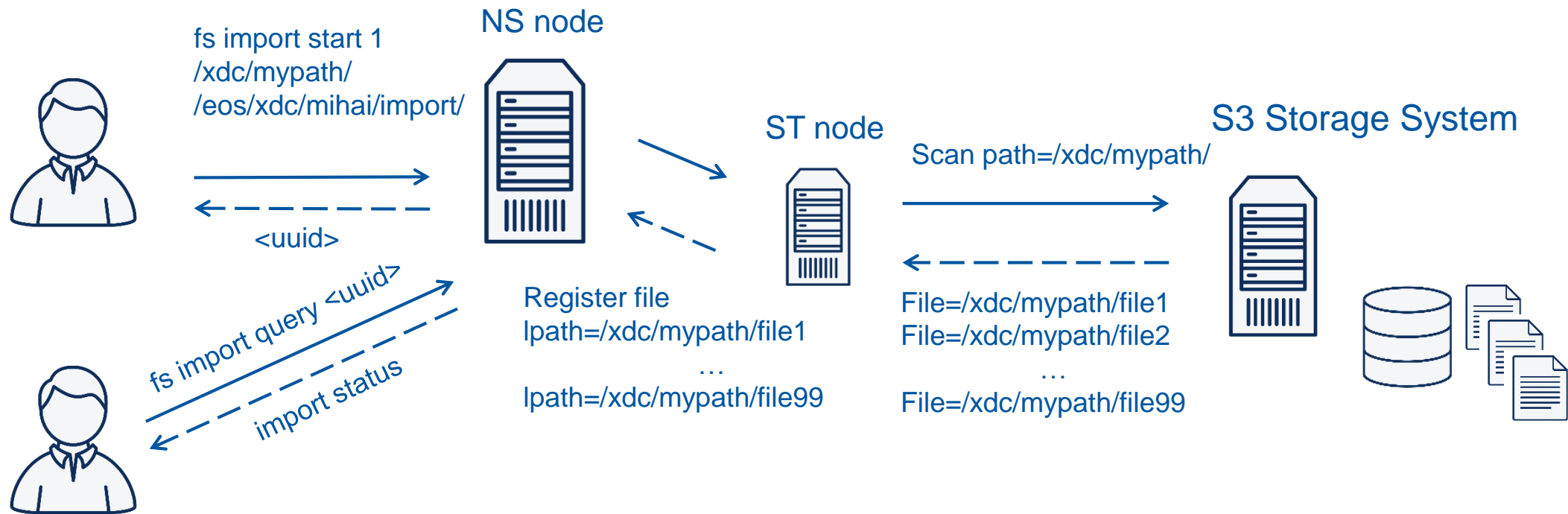
Once finished:

- Import end signal is sent to the MGM



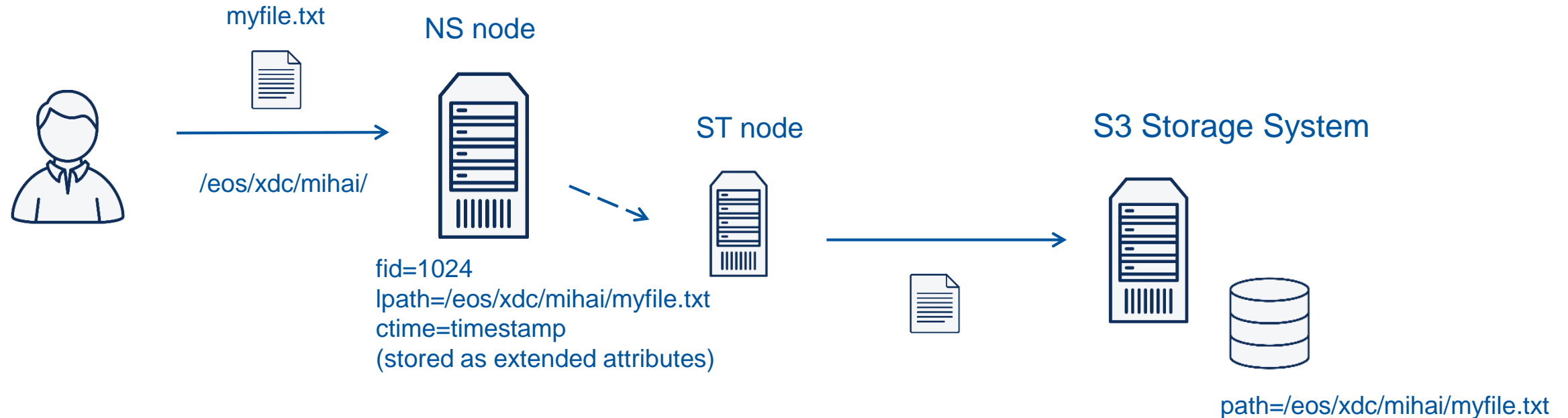
[2 7 5 2 6 9 5 11 4]

Importing files – Overview



Working with an imported file

MGM keeps track of additional information, such as *lpath* and *ctime*.



Imported files drawbacks

- Direct external storage access leads to inconsistent state
- Handling of external attributes is more difficult (shadow file `.filename.xattr`)
- Everything is tunneled through the FST

Imported files drawbacks – upcoming work

- Direct external storage access leads to inconsistent state
- Handling of external attributes is more difficult (shadow file .filename.xattr)
- Everything is tunneled through the FST
 - Redirect directly to external storage if the client matches the protocol
 - Simple redirect – implicit assumption that the client has access
 - Generate signed links – if authorized access is needed

Development Process

- Separate XDC branch (EOS gitlab)
- The EOS CI / CD process is leveraged
- New 'extstorage-test' CI job
- RPMs published to storage-ci/eos/citrine-xdc/commit/el-7/
- Documentation pushed to cern.ch/eos-docs website

<https://eos-docs.web.cern.ch/eos-docs/configuration/logicalpath.html>

<https://eos-docs.web.cern.ch/eos-docs/configuration/import.html>

Ext-storage test

1. Register s3://s3.cern.ch/eostest/ filesystem & local filesystem
2. Upload 50 files to EOS S3 filesystem
3. Check files arrived in S3
4. Import files from other S3 directory into EOS
5. Replicate imported files from S3 filesystem to local filesystem
6. Drop S3 replicas
7. Check availability in EOS
8. Check absence in S3

Planned developments (1-year timeline)

- XCache write functionality
- File notifications
 - Send/consume inotify events over SSE (Server Sent Events)
 - Integrate with EOS and XCache
 - Explore creating WebDav specific event-set
(joint work with DESY institute)
- QoS (unresolved at the moment)

Thank you for your time!

Icons: [Computer Network Icons collection - openclipart.org](https://openclipart.org)

