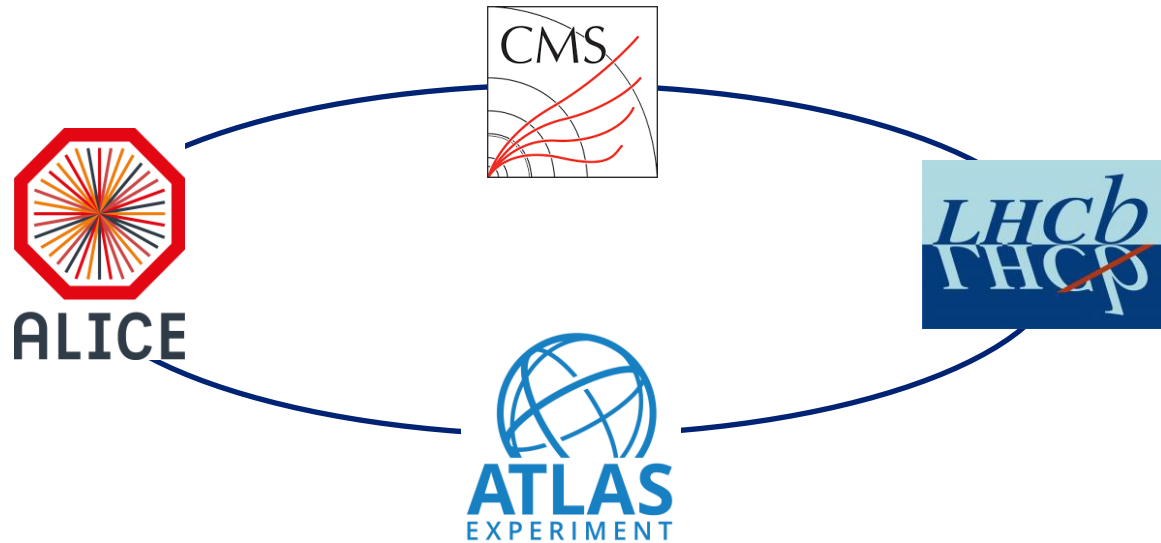


CTA best practices for data taking workflows

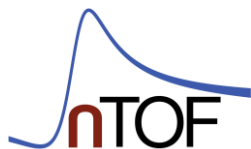
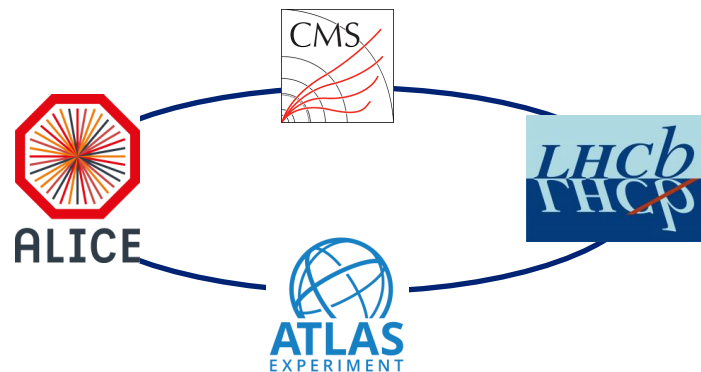
Author: Volodymyr Yurchenko

Co-authors: Vladimir Bahyl, Eric Cano, Michael Davis, David Fernandez Alvarez, Aurelien Gounon, Oliver Keeble, Julien Leduc, Steven Murray, Cedric Caffy

Large experiments at CERN



Large and smaller experiments at CERN



Small and medium experiments at CERN

- **CAST** - A Solar Axion Search Using a Decommissioned LHC Test Magnet
- **NA61/SHINE** - Study of Hadron Production in Hadron-Nucleus and Nucleus-Nucleus Collisions at the CERN SPS
- **NA62** - Proposal to Measure the Rare Decay $K^+ \rightarrow \pi^+ \nu \nu$ at the Cern SPS
- **DUNE** - fundamental research in the field of Neutrino Accelerator Physics
- **COMPASS** - COmmon Muon and Proton Apparatus for Structure and Spectroscopy experiment
- **AMS** - Alpha Magnetic Spectrometer in a module on the International Space Station
- **n_TOF** - Neutron Time-Of-Flight experiment studies neutron-nucleus interactions for neutron energies
- **TOTEM** - Total Cross Section, Elastic Scattering and Diffraction Dissociation at the LHC
- **ILC/CLIC** – the experiment at future particle collider called the International Linear Collider

- Visit <https://home.cern/science/experiments> for more info!

Non-LHC experiments

- Small team of one or two people
- Focus on a specific physics tasks
- Own implementation of online (**data acquisition – DAQ**) and offline (**reprocessing and analysis**) workflows
- Preferred set of commands, protocols
 - API or CLI

Two types of online workflows

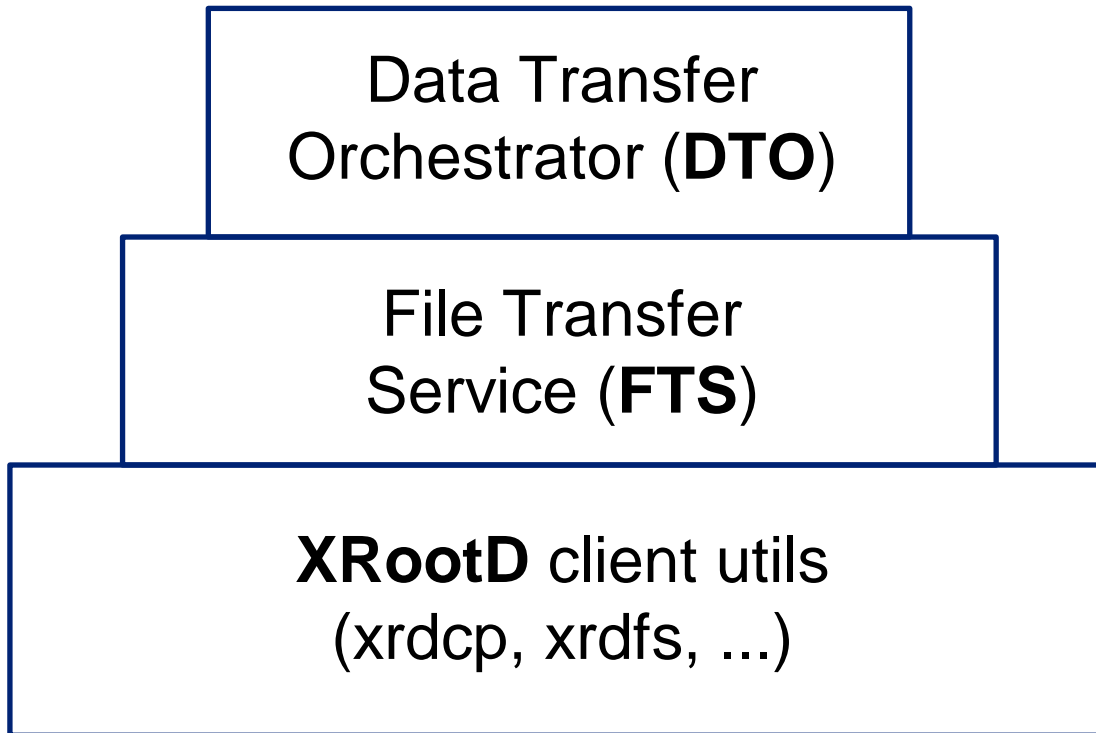
- **Archival:** the experiment writes raw data directly to CTA, files will be recalled sometime in the future
 - This is the simplest use case, as we only need to configure the direct connection from the experiment's machines to the CTA instance
- **Reprocessing:** raw data needs to be both archived and accessible for immediate reprocessing
 - Solution: write one file copy to the tape backend, write another replica to the disk backend
 - Typically, the internal disk buffer of the experiment is not very big, CERN offers a **big disk storage**, called **EOS**
- Workflows implementation in CTA – see [EOS+CTA WorkFlows: Tape Archival and Retrieval](#)

Constraints

CTA best practices are driven by the following constraints:

- Protocols
 - XRootD – currently in production
 - HTTP – deployment is being finalized, will become a recommended protocol in the future
- List of accounts that have access to the CTA instance must be restricted
- Small, fast buffer for writes and reads
 - For some recall use cases we need a disk cache with the garbage collector (see [ALICE and the CTA Garbage Collectors](#))

Tools for the file transfer



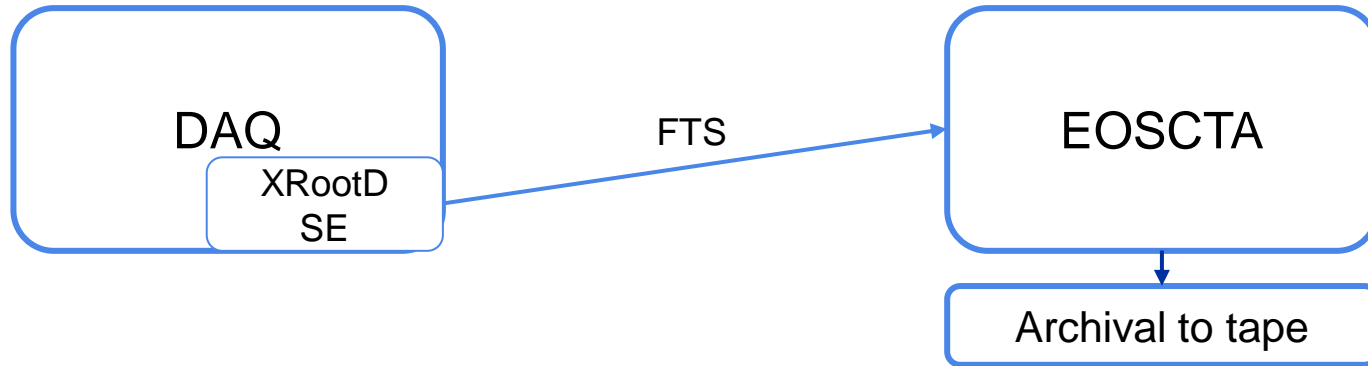
Archival with xrdcp

- xrdcp may be an option when the **simplest direct copy** is needed
 - no need for DAQ storage element
 - requires additional manual management in case of failures



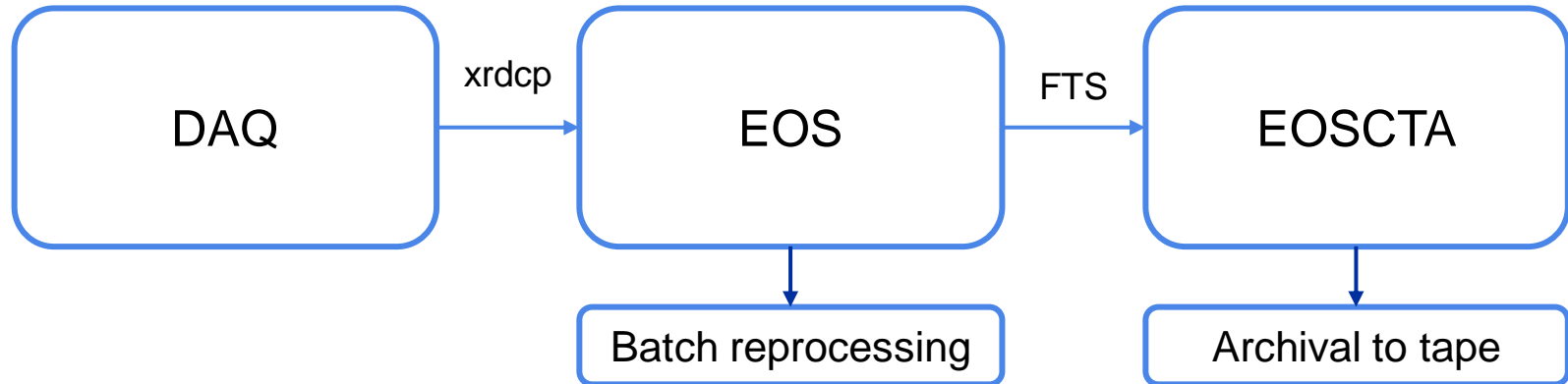
Archival with FTS – File Transfer Service

- FTS is the **recommended** way of file transfer because:
 - Tape transfers are entirely managed by FTS
 - automatic retry mechanism
 - resubmit only failed files
 - checksum validation
 - archive monitoring feature – check-on-tape in CTA
- FTS requires the **storage element (SE)** installed on DAQ side
 - Possible implementation: simple XRootD server in a Docker container



Archival + reprocessing with xrdcp

- First transfer is done with **xrdcp**
- Second transfer is managed by **FTS** for stability and better error handling

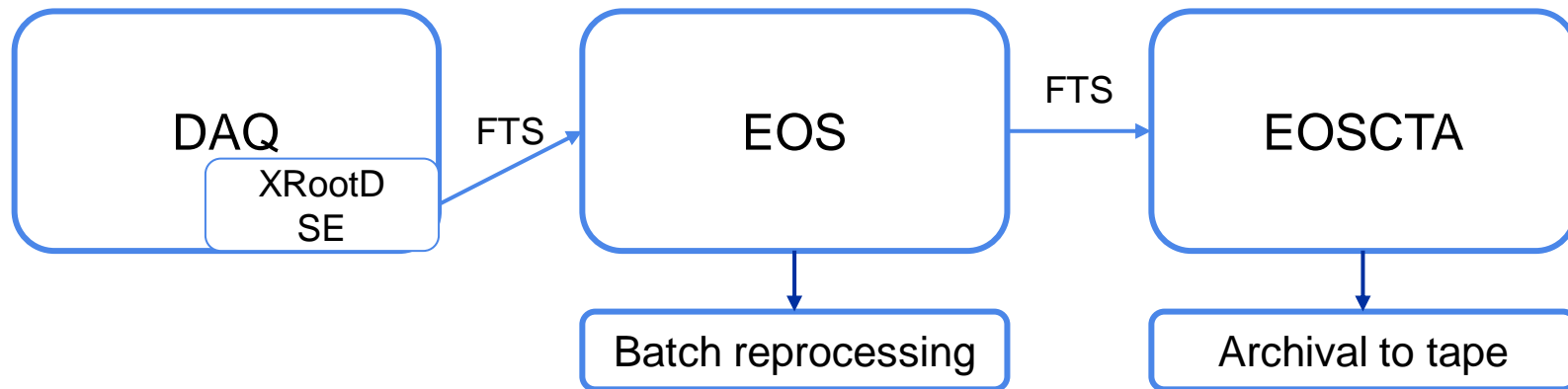


Archival + reprocessing with FTS

FTS “**multi-hop**” transfer feature allows client to define several targets – the file is copied to the big disk storage (EOS), then from EOS to the tape instance (CTA) in one shot.

This chain may have as many “hops”, as required – it can be used for file transfers through T1’s.

In case of copy errors FTS doesn’t resubmit the full chain – the last successful replica is taken for retry.



File safely on tape

File is safe when:

- its **checksum** is OK from DAQ to CTA tapes
 - some protocol combinations break this: for example ftp:// DAQ source
- **“on tape” status flag** is set
 - FTS check-on-tape feature soon in production
 - XRootD “BackupExists” / “on_tape” flag

Example of xrdcp transfer

Copying a local file to CTA directly:

```
$ export XrdSecPROTOCOL=gsi,unix
```

```
$ xrdcp test.ini
```

```
root://eosctapublicpps//eos/ctapublicpps/path/test.ini
```

```
210224 09:44:18 15781 cryptoss1_X509CreateProxy: Your identity:  
/DC=ch/DC=cern/OU=Organic Units/OU=Users/CN=vyurchen/CN=758460/  
CN=Volodymyr Yurchenko
```

```
[248B/248B] [100%] [=====]  
[248B/s]
```

Example of FTS transfer

```
$ fts-rest-transfer-submit -s https://fts3-pilot.cern.ch:8446 -f submit.json
```

where the content of submit.json is:

```
{
  "files": [
    {
      "sources": [
        "root://experiment-daq.cern.ch/path/to/file"
      ],
      "destinations": [
        "root://eospublic.cern.ch/eos/ctapublic/path/to/file"
      ]
    },
    {
      "sources": [
        "root://eospublic.cern.ch/eos/ctapublic/path/to/file"
      ],
      "destinations": [
        "root://eosctapublic.cern.ch/eos/ctapublic/path/to/file"
      ]
    }
  ],
  "params": {
    "multihop": True
    "archive_timeout": 604800
  }
}
```

multihop param instructs FTS to perform a sequential transfer from DAQ to EOS, from EOS to CTA

archive_timeout is the time in seconds to wait until CTA confirms the file is safely on tape
CTA recommendation is to set this timeout to 1 week

File status with XRootD

```
$ XrdSecPROTOCOL=gsi,unix xrdfs root://eosctapublicpps.cern.ch query prepare 0
/eos/ctapublicpps/archivetest/path/test.ini | jq
{
  "request_id": "0",
  "responses": [
    {
      "path": "/eos/ctapublicpps/archivetest/path/test.ini",
      "exists": true,
      "path_exists": true,
      "on_tape": true,
      "online": false,
      "requested": false,
      "has_reqid": false,
      "req_time": "",
      "error_text": ""
    }
  ]
}
```

"on_tape": true is specific to the tape backend. It tells if the file is safely on tape.

"online": false indicates that the file is not in the disk buffer. CTA automatically evict the file from disk as soon as it is archived

Monitoring with FTS

- command line tool

```
$ fts-transfer-status -s https://fts3-pilot.cern.ch:8449 <job_id>
```

```
Request ID: <job_id>
```

```
Status: FINISHED
```

```
Client DN: <DN>
```

```
Reason:
```

```
Submission time: 2021-02-22T17:41:47
```

```
Priority: 5
```

```
VO Name: <VO>
```


Monitoring with FTS

- Web interface: https://fts3-pilot.cern.ch:8449/fts3/ftsmon/#/job/<job_id>

Showing 1 to 1 out of 1

SUBMITTED DELETE READY STAGING ARCHIVING ACTIVE STARTED CANCELED FAILED 1 FINISHED NOT_USED

First Previous 1 Next Last

File ID	File State	File Size	Throughput	Remaining	Start Time	Finish Time	Staging Start	Staging End	Archiving Start	Archiving End	
+ 1447915675	FINISHED	2.99 GiB	160.95 MB/s	-	2021-02-23T19:07:45Z	2021-02-23T19:08:05Z	2021-02-22T17:43:28Z	2021-02-23T19:07:43Z			 Log

DTO – Data Transfer Orchestrator

- A generic tool that pushes data for DAQ for small experiments
- Acts as a sync daemon on a specific directory
- Provides the file catalogue on the users side
- Runs on the DAQ machine
- Requires gridFTP / XRootD server
- Uses FTS
- Developed by @Marco Boretto, @Cristina Voineag (NA62)
- <https://gitlab.cern.ch/ep-dt-di/daq/dto>



CTA authentication policy

- The limitations do not come from the technical, but from operational side. Our goal is to **restrict write access** to the smallest set of users / accounts.
- Preferred authentication method is **grid certificate** which maps to a CTA privileged user.
- Ideally, the experiment should only write files with its DAQ service account.

Conclusion

CTA best practices for DAQ workflows are based on the experience of integrating small and medium CERN experiments.

One of the experiments is already in production with CTA, the rest will be migrated during 2021.

Recommendations for the preferred way of interaction with the tape backend include the preferred file transfer tools, protocols, authentication policy and checks to ensure files are safely on tape.