

DUNE Rucio Deployment and Plans for the Token Era

Steven Timm (Fermilab) / Doug Benjamin (BNL)

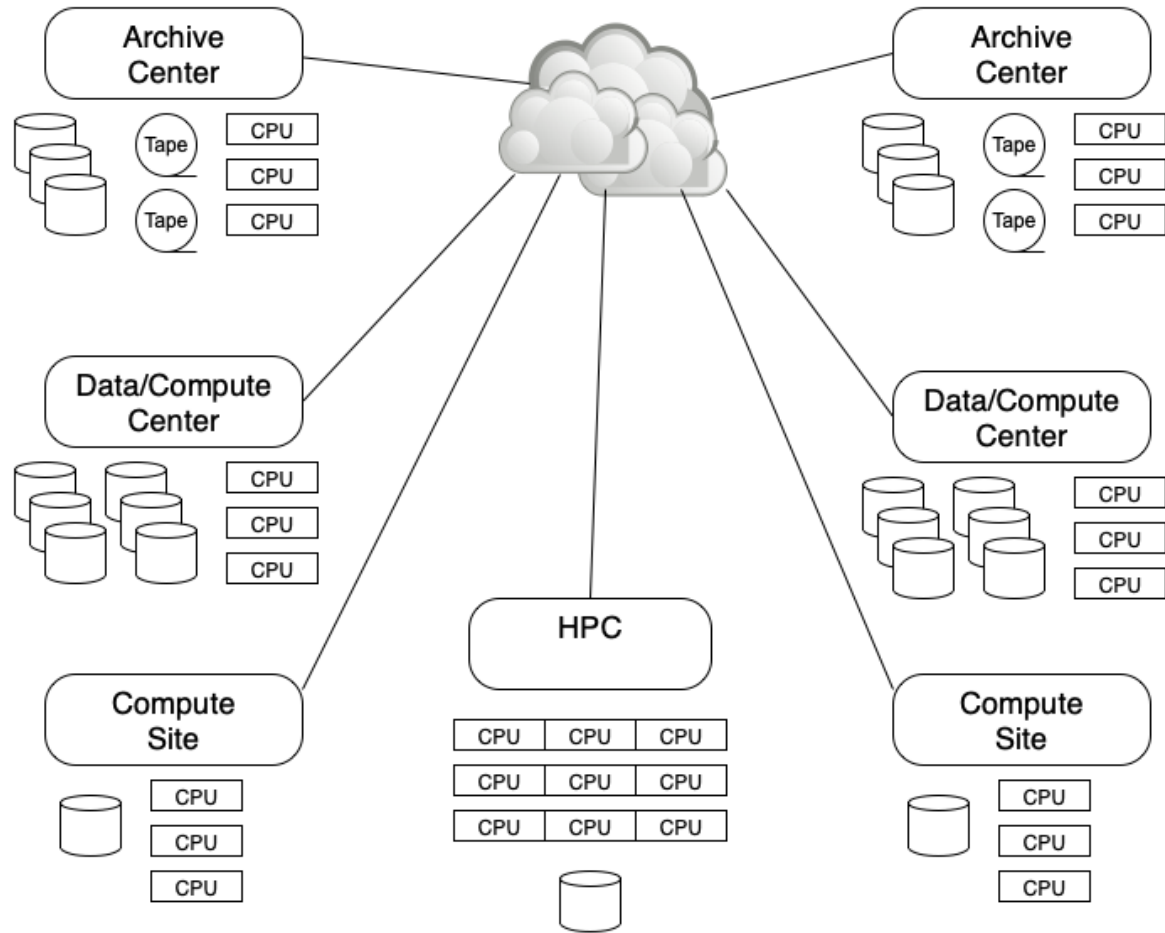
DUNE Data Management Group

For the DUNE collaboration

Rucio Workshop 10 Nov 2022

DUNE Compute and Storage Facilities

- 17 active RSEs around the world. 21PB data
- Primary tape sites at FNAL and CERN, (raw data)
- Secondary tape sites at RAL and IN2P3 (reconstructed data)
- Topology – data at “local” “nearby” “stream”
- Data model: Copies of current MC and reco on disk in US and Europe.



DUNE Software Terminology

Name	Purpose	SUP	Token-ready
Rucio	Replica manager	ATLAS	
FTS3	File transport	CERN	Testing
Declaration/Ingest Daemon	Declare files to Rucio	FNAL	* (Planned)
MetaCat	Metadata catalog	FNAL	* (Planned)
Data Dispatcher	File URL delivery to jobs	FNAL	* (Planned)
dCache (Enstore)	Storage systems	FNAL	Yes
EOS (CTA)	Storage systems	CERN	Yes
Other storage providers			?
HTCondor	Batch system	WISC	Yes
GlideinWMS	Workload system	FNAL	Yes
Workflow manager	Workflow system	GRIDPP	No
Jobsub_lite	Batch submission	FNAL	Yes

DUNE Data Challenge and Processing Workflows

- New Workflow System being scale tested in Data Challenge 4.
 - 5 copies * 500 TB of test data spread out around the world.
- Ultra-late-binding: Workflow task and file are chosen at run time
- Generic jobs are submitted to sites
 - Handle the tasks of requesting a workflow and file,
 - **Use production credentials to access storage on behalf of the user.**
 - Data is downloaded or streamed from nearest Rucio RSE
 - Use Rucio upload to send results back to Rucio RSE
 - This generates a lot of single file rules, want to get around this eventually.
- Currently all in-job storage access done by X.509 proxy
- Expect switching to tokens will be straightforward
- All inputs and outputs under production ownership now,
 - Want to keep it that way in token era.



CILogon Token Issuer

- Operated by CILogon on behalf of Fermilab-based VOs
- Driven by LDAP server which is filled from Fermilab's FERRY DB
- Authentication for tokens can be via Fermi Kerberos or the Fermi IdP
- Issues tokens completely compatible with WLCG schema (see next slide)
- Fermi users can use “htgettoken” utility
 - (don't have to make an oidc client, etc.)
 - Common commands such as job submission, file transfer, get tokens if necessary, transparent to the user
 - Long lived token stored in the Vault server
 - Short-lived bearer tokens stored on volatile disk area.
- Unlike IAM, CILogon is not an all-in-one server--does not support legacy voms-proxies (will keep legacy VOMS server up)
- Any oidc-client workflows or SCIM scripts will need to be tested against this server too.

Physics Groups dCache

- First exposure of DUNE users to token-based storage
- 11 extra token roles created, each with extra scope to create files in one group directory—quotas based on GID

```
{  
  "wlcg.ver": "1.0",  
  "aud": "https://wlcg.cern.ch/jwt/v1/any",  
  "sub": "dunendsim@fnal.gov",  
  "nbf": 1666573161,  
  "scope": "storage.create:/dune/scratch/users/timm  
storage.create:/dune/persistent/physicsgroups/dunendsim compute.create  
compute.read compute.cancel compute.modify storage.read:/dune",  
  "iss": "https://cilogon.org/dune",  
  "exp": 1666583966,  
  "iat": 1666573166,  
  "wlcg.groups": [  
    "/dune",  
    "/dune/neardetsim"  
  ],  
  "jti":  
    "https://cilogon.org/oauth2/453f0e3107421319a1b837bf2cd9ff13?type=access  
Token&ts=1666573165975&version=v2.0&lifetime=10800000"  
}
```

Permission Layout: Legacy vs. Tokens

- Legacy situation (X.509 proxy / NFS v4.1):
 - Fermilab dCache has each individual user mapped to individual user id.
 - Unix permissions govern which users can write in which directories.
 - “dunepro” production user owns all the raw data and production outputs, (managed by Rucio).
 - Other sites—DUNE VO maps to one or two users. “dune” and “dunepro”
 - All data at external sites managed by Rucio (except small interactive space @ CERN).
- Tokens on dCache:
 - By default all tokens issued by the dune issuer map to dune:dune
 - Can override this with dCache inheritance—then files will inherit user:group of the directory
 - This allows us to have group quotas
- Corner cases exist:
 - Files can and do wind up with different owners
 - X.509 proxies can write directories tokens can’t, and vice versa.
- GOAL: want to keep everything manageable by the production user in the era of tokens.

Constraints From External Software

- All DUNE pilot jobs already being submitted with tokens
- Grid Community Toolkit already end-of-life (as of May 2022)
 - Some key software still available in EPEL but not assured.
- User job submission with tokens is in beta “Jobsub-lite”
- HTCondor 9.0 supporting GSI goes away March 2023
- There are rough plans to carry X.509 proxy in jobs for writing storage
 - Will have to maintain this hybrid mode at least until Rucio token support is ready (2024)?
 - May have to maintain this hybrid mode until our last DUNE SE supports tokens (end of LHC Run 3).

Questions on future token-exchange workflows

- DUNE has variety of storage providers
 - EOS, XRootD, dCache, Enstore, CTA, DPM, ECHO, STORM, Ceph
 - Details of who owns the files are crucial
 - dCache @ Fermilab may be different than dCache @ BNL
 - DUNE's goal to try to make all files movable/readable/deleteable by the rucio daemon
- Expiration times of tokens—how long
- Do they have refresh tokens on them
- Scenarios to avoid:
 - User uploads file to Rucio which the rucio daemon then can't move or control.
 - Multiple users owning various branches of a hashed rucio directory tree
- Load on token issuer?
- Synchronization processes—will they work with CILogon?
- What about Rucio and FTS3 integration—2 levels of token exchange?

Future Plans

- All our European storage elements shared with some other WLCG VO's
- Would be nice to have documented a hybrid configuration so we can test tokens in 2023-2024 time frame.
- Goal: Be ready to use tokens for all storage when Rucio token support is complete
- Will participate in WLCG data challenge 2024 using tokens.
- Thanks to team: Wenlong Yuan, James Perry, Brandon White, Dennis Lee, Fermilab Federated ID project team.

BACKUP SLIDES

CILogon Web Auth Screen Shot



```
[tim@snowball ~]$ htgettoken -r fardetsim -i dune --vaultserver htvaultprod.fnal.gov
Attempting OIDC authentication with https://htvaultprod.fnal.gov:8200

Complete the authentication via web browser at:
https://cilogon.org/device/?user_code=TXL-F6T-3ZM
No web open command defined, please open URL manually
Waiting for response in web browser
```

Consent to Attribute Release

htvaultst-dune-vault requests access to the following information. If you do not approve this request, do not proceed.

- User Code: TXL-F6T-3ZM
- Your CILogon user identifier
- Your name
- Your email address
- Your username and affiliation from your identity provider

Select an Identity Provider

Fermi National Accelerator Laboratory

☐ Remember this selection

Log On

Cancel

By selecting "Log On", you agree to the [privacy policy](#).



CILogon User Code Verification



You have successfully approved the user code. Please return to your device for further instructions.