

rucio-analysis - added functionality

- Wrappers for:
 - getting/setting metadata
 - adding updating subscriptions
- SKAO tests of metadata based subscriptions
 - Simple metadata-based replication test
 - Upload data, set metadata
 - Create subscription to create replication rule
 - 'SDP-SRC interface' test
- <https://github.com/ESCAPE-WP2/rucio-analysis>



SDP-SRC Interface test

- Create empty dataset at ND RSE1
- Set metadata on dataset
- Create metadata-based subscription for dataset to replicate to RSE2
- Upload data to ND-RSE1 site via GFAL
- Register data to dataset
- (Manual step) Verify replication of data1 to RSE2 (may take a few minutes)
- <https://github.com/ESCAPE-WP2/rucio-analysis/blob/master/etc/tasks/skao/tests/sdp-src-simulation.yml>



SDP-SRC test run

```
[root@730afb0338e3 rucio-analysis]# python3 src/run.py -v -t etc/tasks/skao/tests/sdp-src-simulation.yml
2021-07-22 15:22:10,908 [CreateSubscription] task DEBUG 658 Constructing instance of CreateSubscription()
2021-07-22 15:22:10,947 [CreateSubscription] task INFO 658 Executing CreateSubscription.run()
2021-07-22 15:22:10,948 [CreateSubscription] sdp_src INFO 658 Creating a new dataset
2021-07-22 15:22:10,948 [CreateSubscription] helpers INFO 658 Checking to see if DID (rjoshi_test:test_dataset_2) already exists...
2021-07-22 15:22:11,343 [CreateSubscription] helpers DEBUG 658 Adding DID rjoshi_test:test_dataset_2 of type DATASET
2021-07-22 15:22:11,508 [CreateSubscription] sdp_src INFO 658 Setting metadata on dataset rjoshi_test:test_dataset_2
2021-07-22 15:22:11,510 [CreateSubscription] sdp_src WARNING 658 AttributeError("'Client' object has no attribute 'set_metadata_bulk',)
2021-07-22 15:22:11,510 [CreateSubscription] sdp_src INFO 658 Bulk metadata call may have failed, trying to set metadata fields individually.
2021-07-22 15:22:11,657 [CreateSubscription] sdp_src INFO 658 Creating/updating subscription test_subscription_rj_2
2021-07-22 15:22:11,825 [CreateSubscription] sdp_src INFO 658 Creating subscription test_subscription_rj_2
2021-07-22 15:22:11,990 [TestUploadNondeterministic] task DEBUG 658 Constructing instance of TestUploadNondeterministic()
2021-07-22 15:22:11,994 [TestUploadNondeterministic] task INFO 658 Executing TestUploadNondeterministic.run()
2021-07-22 15:22:12,209 [TestUploadNondeterministic] upload_nondeterministic INFO 658 Uploading file with path https://fal-pygrid-30.lancs.ac.uk/dpm/lancs.ac.uk/home/skatelescope.eu/rucio-ska/dev/nondeterministic/nonde
terministic_test_220721T15.22.12/project0/dataset0/0_7dc44879-1164-4a68-980f-eeb0a50ad1a0
2021-07-22 15:22:13,171 [TestUploadNondeterministic] upload_nondeterministic INFO 658 Uploading file with path https://fal-pygrid-30.lancs.ac.uk/dpm/lancs.ac.uk/home/skatelescope.eu/rucio-ska/dev/nondeterministic/nonde
terministic_test_220721T15.22.12/project0/dataset0/1_cb424e4b-b3b8-42d5-9b22-037b7deee507
2021-07-22 15:22:13,860 [TestUploadNondeterministic] upload_nondeterministic INFO 658 Writing file list to https://fal-pygrid-30.lancs.ac.uk/dpm/lancs.ac.uk/home/skatelescope.eu/rucio-ska/dev/nondeterministic/ingest/ru
n_220721T15.22.12
2021-07-22 15:22:14,426 [TestUploadNondeterministic] upload_nondeterministic INFO 658 Finished in 2s
2021-07-22 15:22:14,449 [AuxRegisterNondeterministic] task DEBUG 658 Constructing instance of AuxRegisterNondeterministic()
2021-07-22 15:22:14,452 [AuxRegisterNondeterministic] task INFO 658 Executing AuxRegisterNondeterministic.run()
2021-07-22 15:22:15,137 [AuxRegisterNondeterministic] register_nondeterministic INFO 658 Adding replica for 0_7dc44879-1164-4a68-980f-eeb0a50ad1a0 at LANCASTER_ND at with did rjoshi_test:0_7dc44879-1164-4a68-980f-
eeb0a50ad1a0
2021-07-22 15:22:16,574 [AuxRegisterNondeterministic] helpers INFO 658 Checking to see if DID (rjoshi_test:test_dataset_2) already exists...
2021-07-22 15:22:16,708 [AuxRegisterNondeterministic] helpers DEBUG 658 DID already exists. Skipping.
2021-07-22 15:22:16,709 [AuxRegisterNondeterministic] register_nondeterministic INFO 658 Attaching did to dataset rjoshi_test:test_dataset_2
2021-07-22 15:22:16,869 [AuxRegisterNondeterministic] register_nondeterministic INFO 658 Adding rule to keep it there for 3600s
2021-07-22 15:22:17,071 [AuxRegisterNondeterministic] register_nondeterministic INFO 658 Adding replica for 1_cb424e4b-b3b8-42d5-9b22-037b7deee507 at LANCASTER_ND at with did rjoshi_test:1_cb424e4b-b3b8-42d5-9b22-
037b7deee507
2021-07-22 15:22:18,482 [AuxRegisterNondeterministic] helpers INFO 658 Checking to see if DID (rjoshi_test:test_dataset_2) already exists...
2021-07-22 15:22:18,622 [AuxRegisterNondeterministic] helpers DEBUG 658 DID already exists. Skipping.
2021-07-22 15:22:18,622 [AuxRegisterNondeterministic] register_nondeterministic INFO 658 Attaching did to dataset rjoshi_test:test_dataset_2
2021-07-22 15:22:18,782 [AuxRegisterNondeterministic] register_nondeterministic INFO 658 Adding rule to keep it there for 3600s
2021-07-22 15:22:19,033 [AuxRegisterNondeterministic] register_nondeterministic INFO 658 Finished in 5s
[root@730afb0338e3 rucio-analysis]# rucio list-rules rjoshi_test:test_dataset_2
ID ACCOUNT SCOPE:NAME STATE[OK/REPL/STUCK] RSE_EXPRESSION COPIES EXPIRES (UTC) CREATED (UTC)
-----
5017a94c0bd646679dd7ec17106618c4 rjoshi rjoshi_test:test_dataset_2 REPLICATING[0/1/0] MANCHESTER 1 2021-07-22 16:22:16 2021-07-22 15:22:16
[root@730afb0338e3 rucio-analysis]# rucio list-rules rjoshi_test:test_dataset_2
ID ACCOUNT SCOPE:NAME STATE[OK/REPL/STUCK] RSE_EXPRESSION COPIES EXPIRES (UTC) CREATED (UTC)
-----
5017a94c0bd646679dd7ec17106618c4 rjoshi rjoshi_test:test_dataset_2 REPLICATING[0/2/0] MANCHESTER 1 2021-07-22 16:22:16 2021-07-22 15:22:16
[root@730afb0338e3 rucio-analysis]# rucio list-rules rjoshi_test:test_dataset_2
ID ACCOUNT SCOPE:NAME STATE[OK/REPL/STUCK] RSE_EXPRESSION COPIES EXPIRES (UTC) CREATED (UTC)
-----
5017a94c0bd646679dd7ec17106618c4 rjoshi rjoshi_test:test_dataset_2 REPLICATING[0/2/0] MANCHESTER 1 2021-07-22 16:22:16 2021-07-22 15:22:16
[root@730afb0338e3 rucio-analysis]# rucio list-rules rjoshi_test:test_dataset_2
ID ACCOUNT SCOPE:NAME STATE[OK/REPL/STUCK] RSE_EXPRESSION COPIES EXPIRES (UTC) CREATED (UTC)
-----
5017a94c0bd646679dd7ec17106618c4 rjoshi rjoshi_test:test_dataset_2 OK[2/0/0] MANCHESTER 1 2021-07-22 16:22:16 2021-07-22 15:22:16
```

Create dataset, set metadata

Create subscription

Upload data to non-deterministic RSE

Register data

Attach to dataset

Data replicated to MANCHESTER



Outcomes

- Non-deterministic issue encountered with transmogrifier daemon
- Unable to delete subscriptions via Rucio API/CLI
- Transmogrifier only considers DIDColumn-based metadata

Next steps

- Upload to 'faraway' non-deterministic RSEs (AARNET and IDIA)
- QoS based subscription replication

