

Minutes for 2020-09-16 T2.2 meeting

Present

Aleem Sarwar
Frederic Gillardo
Gareth Hughes

Marcelo Soares
Marek Szuba
Nadine Neyroud

Paul Millar
Rohini Joshi

Apologies

Martin Barisits

Agenda

News

News from WP2 meeting last week: WP2 is moving from Vidyo to (the CERN instance of) Zoom because CERN has decided not to extend the contract with Vidyo. From QoS side, we can probably take the same approach as with Vidyo *i.e.* use the same room (linked to from CERN Indico) as main WP2 meetings. Aleem will contact Simone and Xavi to ask about switching from Vidyo to Zoom.

Round table

Streams info can be found at:

<https://indico.in2p3.fr/event/19937/contributions/75972/attachments/56927/75787/ESCAPE-T2.2-update.pdf>

Stream A: prototyping and demonstration

Plans for the demo-day

https://indico.in2p3.fr/event/22196/contributions/86029/attachments/59734/80871/ESCAPE_datalake_performance_dress_rehearsal.pdf

Simulating analysis vs actual analysis jobs

CTA

https://docs.google.com/document/d/1rGThyJwCsT_Clyb6dkioRRFQU-cvCKNXdudb9Waa37w/edit?usp=sharing

LAPP using HT-Condor,
For CTA we need to generate fake data, then simulate ingest.

For reconstruction (demonstrate CTA specific workflow)

For CTA, HDF5 format is used.
Analysis read the header of the HDF5 file.

Fake data should be a valid HDF5 file.

Currently, the analysis involves downloading the complete file to the worker node, and access from there.

Datasets are well-defined. Access all the files in the dataset. Files are 2 GiB in size.

A job would download all files in a dataset, to process them.

There would be several jobs running concurrently.

For data injection

Dataset of 20 GiB, expect once per 15 minutes.

Inject the data into one of the storage elements at LAPP.

The distributed work-flow on data-ingest, where data is hopped from initial site to a back-up/secondary site would be “nice to have”.

Simulating tape would be “nice to have”.

SKA

https://docs.google.com/document/d/1uludSciEt5uR7NkSt9jL_-hOEgzFqUUotLAdzwwHP0M/edit#heading=h.398og476hihs

For analysis

There are datasets from SKA science data challenge 1.

Continuum emission “images” 4 GiB files, each. (Some of) the files are different frequency bands.

There is a work-flow that goes with these files that 'solves' the data challenge; identifies sources within the image and records what sources are present in a catalogue.

It takes 3--4 hours to complete, on a VM with sufficient CPU and memory.

This could be simulated by simply downloading, running the workflow and uploading the results back up again.

Also looking into a pulsar search based use case

These are closer to HEP / "embarrassingly parallelisable".

Looking at lots of small files (~ a few MiB), but large number $\sim 10^3$, which would be simulated pulsar observations.

Unclear how these would be grouped together into datasets (and containers) based on day of observation/observation ID/etc.

For data injection

The pulsar use-case would be a good example of data injection.

Unclear what at frequency we would need to inject these files.

Multi-hop transfers are "would be nice" -- would need to be two SEs already in the data lake.

ATLAS

<https://docs.google.com/document/d/1Ur2kLndoEbccP8wQWg3dKFo5uQ0sWhd0l701ZUxxMUs/edit#heading=h.398og476hihs>

Skipped, as nobody from ATLAS was here.

FAIR

- CBM experiment only:
https://indico.in2p3.fr/event/21945/contributions/84643/attachments/59520/80433/20200715-QoS_CBM_v1.pdf
- PANDA contribution expected soon

For analysis

The two CBM guys who have been working on integrating data-lake access into their workflows are currently away for extended periods of time. Having received a status update from the one who has only just left, we will almost certainly *not* have live use-cases ready for early November.

Use of GSI-ROOT for data-movement exercises in non-FAIR use cases (including setting custom QoS tags) is very much welcome. Both SciToken and VOMS authn/authz are enabled, the former only for WebDAV for now though (must upgrade xrootd to 5.0 for native-xrootd token support).

For data injection

Mostly using ROOT-formatted data. Relatively large file size: current tests use files of around 1 GiB. Data rates currently supported by the RSE are “hundreds of megabytes” per second.

Marek will come up with scripts (etc.) to simulate data injection if no real-world input is available in time.

For monitoring

Adding some additional monitoring for the Rucio server (e.g. REST API response time, as previously discussed) will not be ready any time soon --- currently no way to feed external metrics into our monitoring infrastructure. WP2 plans to have an InfluxDB instance outside of CERN which will serve as a data source for our Grafana instance, it has only just recently been decided though. More discussion offline, particularly involving Rohini --- they will most likely host that instance.

Stream B: engagement with experiments

NTR

Stream C: software developments

Rucio REST API (Aleem)

Aleem will update Rucio REST API access on ESCAPE wiki page/google doc already available there.

Marek: make sure to include instructions about installing the CERN CA bundle. In the long run, it would make sure to switch the Rucio server to a certificate issued by a publicly recognised CA --- CERN issues such certificates upon request, or we could just deploy Let's Encrypt certificates on that server.

We can upload data, specify QoS and check credentials. We need to have a Rucio authentication token, which expires relatively quickly.

Had to by-pass the client checking the certificate authority when verifying the Rucio server host certificate --- the current host certificate has been issued by the CERN Grid CA, which is not publicly trusted. Trusting that CA is possible but depending on the client and the system, can be tricky.

Request for the CA change on the Rucio server: Riccardo and Rizart are informed, Marek has created a JIRA ticket (see <https://jira.skatelescope.org/browse/EDLK-59>). Alba (or whoever else has got admin access to the escape-rucio K8s pod) will likely need to fix this.

AOB

The next meeting will be on 2020-09-30, the usual time.