

<https://indico.in2p3.fr/event/19943/>

News

Xavi

ESCAPE EB 3rd of April:

- 1) Based on article 51 (Cause of Force Majeure) in the grant agreement, we prepared a statement on Covid-19 impact in the deliverables and milestones of the ESCAPE project. To be sent to the EU project office. Some of the WP are more affected than others but certainly will have a general impact. (Question: Pandey: What's the timescale of the impact?
Xavi: Unclear, mainly has to do with other work packages. We can not quantify it yet as the duration of this situation is still unknown.)
- 2) We are trying to understand the participation on the EU call INFRAEOSC03. A position statement letter being prepared together with the other science clusters (ESCAPE, PANOSC, EOSC-life, SSHOC and ENVRI-FAIR)
- 3) The ESCAPE External Advisory Board (EEAB) report prepared after the Brussels General Assembly was positive.
- 4) Understanding possible collaboration with the ARCHIVER project (EU funded). Archiver project is trying to understand the use case of long term storage in the cloud by sciences. I was invited to the follow-up meeting to understand whether we can collaborate, this will help address one of the ESCAPE goals for trustworthy digital repositories (ISP16363 self-assessment) since we target the same goal. Also two of the partners are common: PIC and DESY. Currently this is under discussion.

Reports

Report by Aris on Datalake Status

Metadata: Rucio VO specific metadata storage plugins are almost there, need to start developing our own solution. Will need to gather parties interested.

SSR: Right now we have a naive view of the amount of storage use per site. WLCG is using SSR to get a more real view. Need to implement also on the ESCAPE Datalake. Every site publishes a json file, a service collects them and publishes them centrally.

Xavi: You should start an email thread about those activities.

SKA interested in metadata

Riccardo on Data Access

From the Brussels progress meeting, the CERN caching infrastructure has been updated and is stable. The CERN based prototype is generic, works, we can work towards a common infrastructure. It's time to converge.

Regarding XRootD+VOMS, the prototype is ready to be tried out by other partners within ESCAPE. Please follow up with me in case you want to replicate and join the effort wrt XCache.

HTTP+tokens is somewhat tricky. We saw a presentation from Diego, but his test wrt HTTP is completely decoupled from XRootD. Working towards integrating both XRootD+VOMS and HTTP+tokens. In contact with XRootd Developer, not clear if we can have both protocols enabled on the same port/path.

Paul will probably start looking into this soon.

Xavi:

Will be interested to deploy this somewhere with non-lhc/events based data: LSST, CTA and MAGIC.

Ghita:

Have a task force with LAPP-ANNECY, idea being to build a cache, configure authentication with x509 and token from user to endpoint behind the cache. Next is to connect with the endpoints at CC and LAPP, inject data from CTA, LSST, and ATLAS and make tests of processing this data from close and remote sites through cache. For now, work is on understanding the authentication.

Xavi:

Action: Can you do a small update/presentation about it (2-4 weeks)?

Ghita:

YES! That's why we have Paul understanding the authentication and stuff. Trying to containerize, not yet kubernitisation.

FG:

For CTA data, waiting for approval by managers to use data

Riccardo: Is that your top priority atm?

Ghita:

Yes, we want to understand tokens and new storages (Xcache). Only involvement for in2p3-cc, can not talk about others in IN2P3.

Riccardo: so we can push it a little bit

Killian (GSI): We have developed xrootd plugin, we can contribute to the xcache infrastructure.

Riccardo: do you plan to have ESCAPE GSI xcache instance

Killian: we are doing something similar but we can contribute

Xavi: Maybe there is enough interest and quorum to organise a dedicated meeting on content delivery and caching activities among all interested people and sites. **Action on Riccardo to evaluate and eventually propose this.**

Report by Richard on Monitoring

Presenting monitoring
Document linked to get you started
Grouped with folders
Would like to have transfer matrix like the madash
When people reach out, we can document and explain available data, then decide which grafana plugins to use and so on.

Xavi:

Action: All task leaders provide input to Richard (please Richard send an email to start the thread)

Andrea: This could be integrated on IAM.

We have done it for other instances, should be something that works out of the box for flat authorization, we have not exposed more fine grained auth.

Would be convenient for people.

Xavi: lets take this offline

Workflows: Xavi mentioned a quick update on Yan's behalf, due to connectivity issues was not noted correctly.

Pandey & Yan Grange (absent temporarily): We are trying a couple of things.

1) We have a container based LOFAR software (singularity based). It can also include a small amount of test data on which a LOFAR pipeline can be run. The pipeline includes steps all the way from flagging (removing bad data) to calibration and final imaging steps. The amount of test data is quite small but good enough to do basic tests.

2) Preparing a table which includes the range of resource requirements (computing/storage) for various LOFAR data types. This will help in estimating what kind of resources do we need to do the basic tasks.

3) Upload the software on Zenodo as per the instructions by Thomas. This will enable the entire ESCAPE community to have access and help in its improvement.

Rosie (brief task 2.4 update): We have reached out to SA partners for their involvement in the ESCAPE data lake, waiting to hear back. Very much interested in the dashboard development.

Andrea: Last week we had the webinar, well attended, no recording possible, will re record part of the webinar by my self to have reference of the main parts of it that we can use as documentation

Need to have remote hackathon with WP5 Zheng,

Xavi: Can we try to zoom?

Moved to Zoom

<https://cern.zoom.us/j/359358038?pwd=>

Meeting ID: 359 358 038

Password:

Agustin PIC presentation:

Slides:

<https://indico.in2p3.fr/event/19943/contributions/81571/attachments/58196/78026/20200408-rucio-wp2-v2.pptx>

- Introduction to pic
- Injection scheme
- Generating data, Rucio has 2 ways to register the data
- Will register data on PIC-INJECT, then replicate to INFN-NA, CNAF-STORM, DESY-DCACHE
- For the next steps, we would like to implement '/' in the naming of files, in order to archive more informative naming files and support pre-existing naming conventions. I.e.
RAW/M1/CrabNebula/2020_03_13/20200313_M1_05089889.008_D_CrabNebula-W 0.40+215.raw.gz . Equally important, we will try to develop an algorithm that builds the folder path from the name-space of the file at the destination RSE.
- Finally, Configure we will need help in order to configure a RSE INJECT for gridftp endpoint in La Palma Observatory.

Questions to Agustin:

Aris: Can you share the scripts and jupyter notebook you used?

Rohini same for us

Agus: Rohini, could you send me your email (or a suited platform for sharing)? So I can share with you the code?

Rohini: Yes, I can find you via rocketchat!

Xavi: Would be nice to have some actual data, (MAGIC?)

Gonzalo: Right now we are learning the technologies.

Delfino: Once we have some auth and understanding we can push the data. We need to have ACLs because its production Data. Other issues is with using an older version of root on our side, we need to check compatibility. We need to investigate also remote access. INFN might be first to try it out since they work with xcache and magic data.

Riccardo: Need to make clear that the lake is not production/reliable for production use.

Delfino: Ofc this is for copy of the data and does not put at risk any of the files. Also it's scaled down.

Riccardo: If we get tokens out of the discussion, ACLs should be there.

Delfino: We promised that we will get rid of x509

Andrea: Tokens should also work

Riccardo: Actually its not straightforward, we face issues with xrootd, we are getting the dev version from directly contacting main dev

Andrea: ofc, for demonstration it should be OK, for production we need more time, we mostly need better documentation, will invest some effort. Then I need users to try it and make comments.

Xavi: **If possible we would like to have an update on this around mid May?**

That was agreed

Xavi: Despite the current situation things are progressing quite well in our community. This is remarkable.

Riccardo: Were workflows mentioned?

Xavi: This will be on the minutes (hopefully)

Action list:

- **Aris: Metadata and Storage Reporting script for the participating**
- **Ghita and Paul Musset:** 2/ 4 weeks time: Presentation on XCache project@IN2P3
- **Riccardo:** Start a thread for interest on the XCache initiative.
- **Riccardo:** evaluate the idea of having a dedicated meeting on content delivery and caching activities among all interested people and sites. This will be summarised afterwards in the WP2 fortnightly meeting.
- **Rizart** to start a thread, tasks leaders to provide input to Rizart for the monitoring.
- **Rizart and Andrea Ceccanti:** Understand IAM as authenticating mechanism for the monitoring
- **Agustin/Gonzalo:** Agreed to have an update on the “gamma-ray data lake injector” from PIC in May.