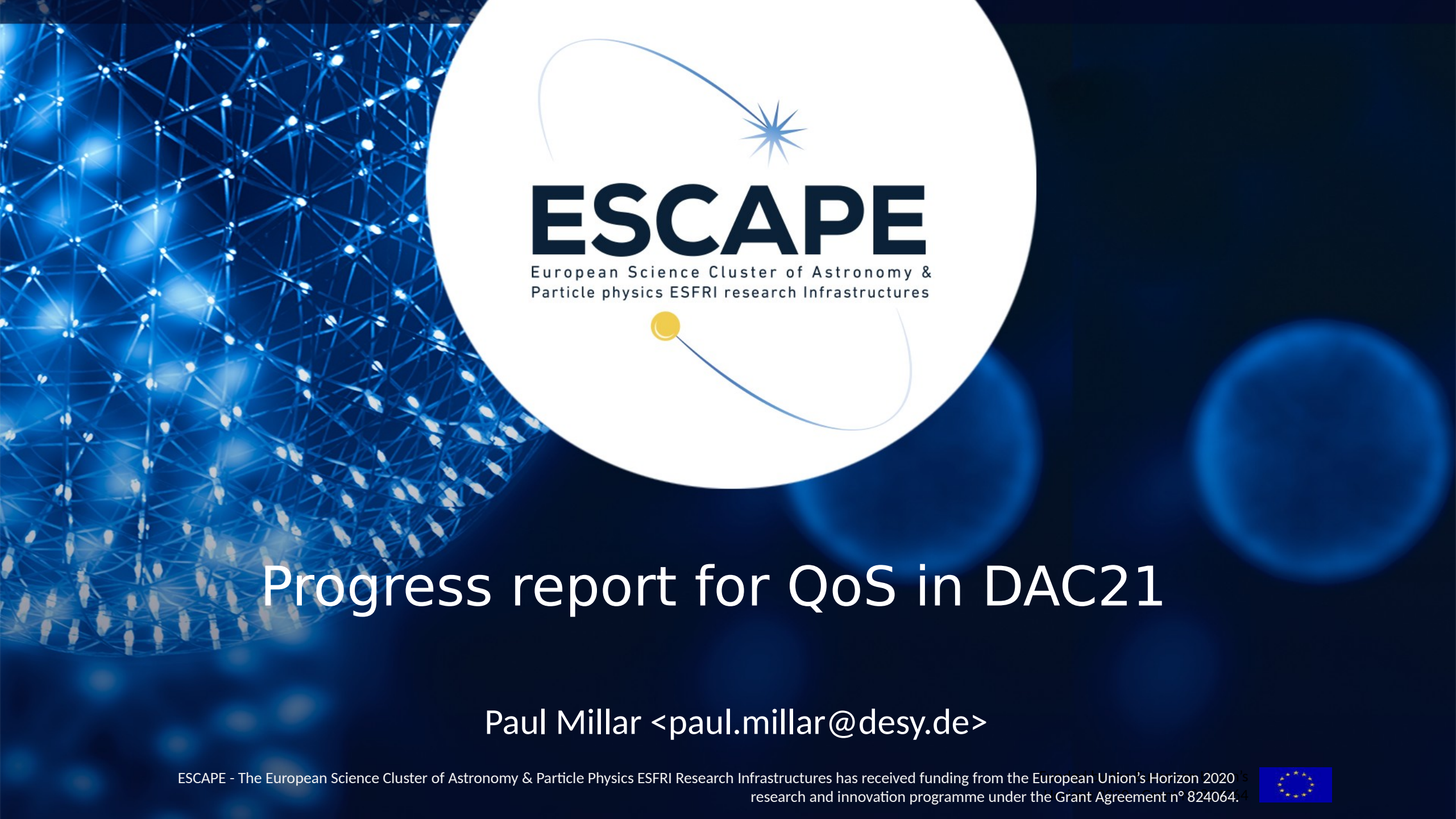

Progress report for QoS in DAC21

Paul Millar <paul.millar@desy.de>

ESCAPE - The European Science Cluster of Astronomy & Particle Physics ESFRI Research Infrastructures has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement n° 824064.

Current status

- Tape storage:
  - SurfSARA: RSE exists (w/ caveats),
  - PIC: RSE exists (w/ caveats),
  - DESY: storage configured and being tested (ETA: 1–2 days),
  - CNAF: under investigation.
- EC storage:
  - CERN: RSE exists for EC(4+2), …
  - SurfSARA: underlying storage has been offered.
- Other possibilities:
  - AWS (RSE exists), CEPH (no clear candidates)

Funded by the European Union's Horizon 2020 - Grant N° 824064

Common ~~problems~~ themes

Funded by the European Union's
Horizon 2020 - Grant N° 824064

Flush triggers

鹿威し / 添水

- Once data is written, it doesn't immediately get written to tape.
  - It spends "some time" sitting on a disk. This is to optimise tape usage.
- Data is written to tape when one of **three** things happens:
  - Spent too long on disk; too much unflushed capacity; too many unflushed files.
- Different policies in WLCG:
  - ATLAS → storing on disk with a promise it'll end up on tape is good enough.
  - CMS → the transfer is successful only once the file ends up on tape.
- With CMS adopting Rucio, this is now an option in Rucio.

Funded by the European Union's Horizon 2020 - Grant N° 824064

Automated testing

- What active testing do we want to do with tape RSEs?
- Tape drives are a precious resource:
  - Tapes are cheap, tape ***drives*** are expensive, tape drives installed in a silo even more so.
  - “Blindly” copying files could create unacceptable load on a site’s tape infrastructure.
- Options include:
  - All tape RSEs have a similar disk RSE: so test the disk RSE only.
  - Transfer files to a tape RSE and delete them quickly, before they are flushed to tape.
  - Transfer files to a tape RSE and verify that they are flushed to tape.
  - Transfer files to a tape RSE, verify that they are flushed, verify that they can be staged back.

Funded by the European Union’s
Horizon 2020 - Grant N° 824064

rse_type ~~bug~~

- During setting up tape endpoints a bug was discovered
  - rse_type is a Rucio setting that enables additional handling, because it can take longer to copy data off of tape.
  - Tape-backed RSEs should have rse_type=TAPE.
  - The correct value was used in CRIC, but not propagated to Rucio configuration.
- The problem was investigated and is now fixed.
  - My thanks go to Rizart and Aleem for sorting this out.

Funded by the European Union's Horizon 2020 - Grant N° 824064

Endpoint-specific notes

Funded by the European Union's
Horizon 2020 - Grant N° 824064

SurfSARA tape

- Status: available as RSE **SARA-DCACHE-TAPE**
- Available capacity: **100 TiB**.
- Available for: **all ESFRI communities**.
- Flush policies: currently not known.
- Data will be wiped after DAC21.

Funded by the European Union's
Horizon 2020 - Grant N° 824064

PIC tape

- Status: available as RSE **PIC-DCACHE-TAPE**.
- Available capacity: currently unclear.
- Available for: **CTA**.
- Flush policy: 24 hours, 1,000 pending files, 1 TiB outstanding.
- Question: how to handle authorization?

Funded by the European Union's
Horizon 2020 - Grant N° 824064

DESY tape

- Status: storage is configured and is being tested, RSE to be added.
- Available capacity: **100 TiB**.
- Available for: **all ESFRI communities**.
- Flush policy: 2 hours, 60 pending files, 20 GB
- Data will be wiped after DAC21.

Funded by the European Union's
Horizon 2020 - Grant N° 824064

CERN EC

- Status: available as RSE **EULAKE-EC**.
- Available capacity: currently unknown.
- Available for: **all ESFRI communities**.
- Coding: EC(4+2)

Funded by the European Union's
Horizon 2020 - Grant N° 824064

SurfSARA EC

- Status: S3/Swift storage has been offered.
- Available capacity: **50 TiB**.
- Available for: **all ESFRI communities**.
- Coding: EC(8+4) (TBC)

Storage is an S3 endpoint. My suggestion: we put the S3 secret into Rucio, as part of the RSE configuration.

Funded by the European Union's Horizon 2020 - Grant N° 824064

Other storage

- AWS storage exists as RSE **AWS_WEBDAV** – However, it's currently unclear whether this will be part of DAC21.
- CEPH – SKA are interested in general, but no plans for DAC21 (unless a site steps up and offers it!)

Funded by the European Union's Horizon 2020 - Grant N° 824064

Thanks for listening

Funded by the European Union's
Horizon 2020 - Grant N° 824064

- Milestones (Further details in this Google docs):
  - M1. **The use-cases have been identified** (due **end of May**)

    All ESFRI communities have a detailed plan on what operations they would like to demonstrate. Any missing features or resources are identified.
  - M2. **The resources needed to support QoS use-cases have been identified and deployed** (due: **end of June**)

    Sites have been identified that will provide access to this hardware, and has been deployed.
  - M3. **The code needed to drive the use-cases is feature-complete** (due: **end of July**)

    This milestone is about developing the code to drive DAC21 QoS and caching use-cases and work-flows. The code is not expected to be feature-complete but not fully tested.
  - M4. **All the QoS use-cases have been exercised at small-scale** (due: **end of August**)

    This milestone is about catching the more obvious problems.
  - M5. **The QoS use-cases have been demonstrated at the desired scale** (due: **end of September**)

    Verify the work-flows work at scale; may require coordinated access.
  - M6. **DAC21 has completed successfully**. (2021-11-xx)

Funded by the European Union's Horizon 2020 - Grant N° 824064