



ESCAPE

European Science Cluster of Astronomy &
Particle physics ESFRI research Infrastructures

QoS experiments input and storage endpoints deployment

Paul Millar (on behalf of WP2.2)



High-level overview

- Storage QoS is a promise (between scientist and resource-provider) on the characteristics of stored data.
 - How fast, what durability, what cost, what parallel-access profile, what deletion rate, ...
- Storage QoS is used for two (related) purposes:
 - Cut down cost of storage
 - Getting the most from computational resources.
- In ESCAPE we want to:
 - Encourage a **discussion** within ESFRI communities on how data is stored
 - Allow people to gain **operational experience** with handling differing QoS
 - Identify **missing functionality**, based on actual or planned use-cases.
 - Provide **development feedback**, to resolve problem areas.

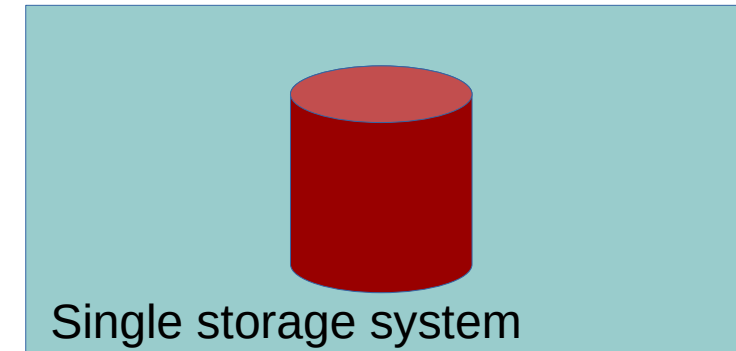
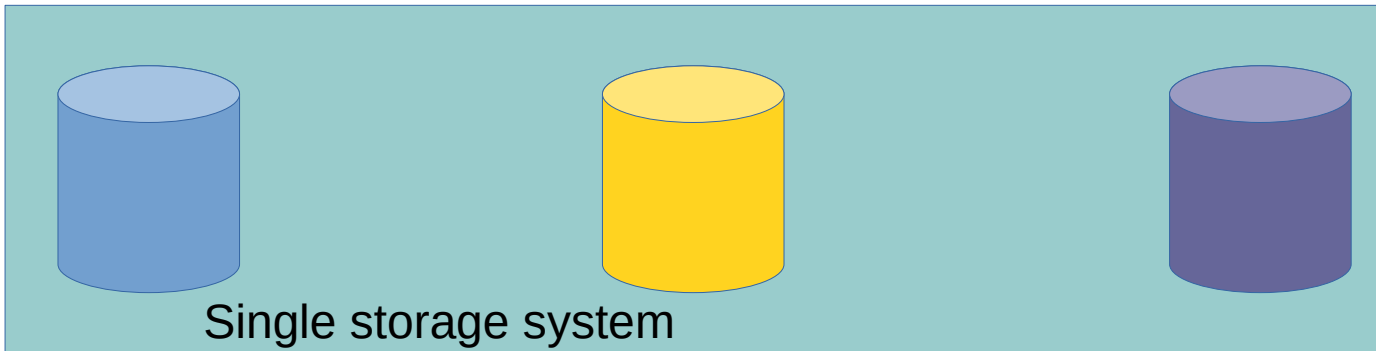


Different backgrounds: common goal

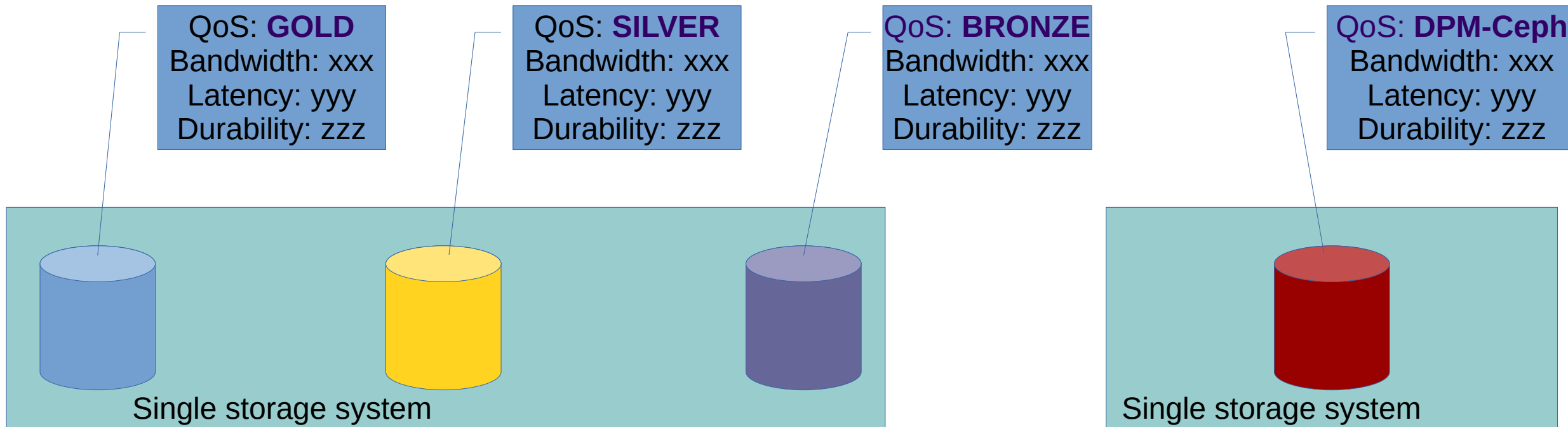
- The Cern/LHC has been in **production** for some ten years:
 - Precursor computation environment predates that, with ~15 years experience.
 - Much of the workflows are well-defined.
- Other ESFRI communities are **yet to take data**: SKA, CTA, FAIR, ...
 - Limited experience, except through pathfinder projects (e.g., LOFAR for SKA).
 - Workflows are a “best guess” and subject to change.
- We want to **build on experiences** from LHC
 - But remain flexible to changes – things are not set in stone.
- Cern/LHC itself is also changing:
 - The pressure from the future High-Lumi run created the DOMA R&D project.
- It's not a one-way street.



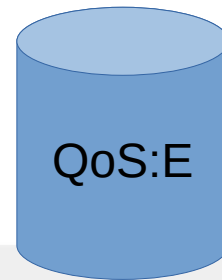
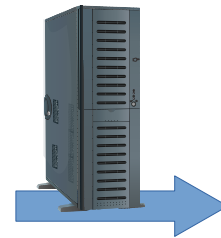
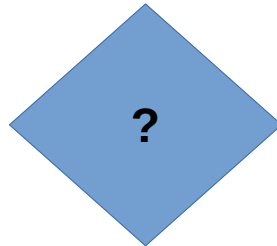
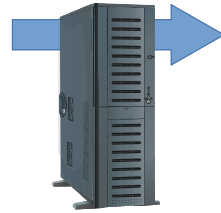
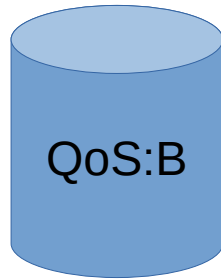
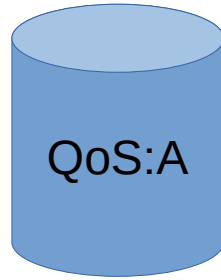
Concepts: Storage QoS classes



Concepts: Storage QoS classes



Concepts: work-flow / data lifecycle





Concepts: VO-QoS-Policies

QoS:A

QoS:B

QoS:C

QoS:D

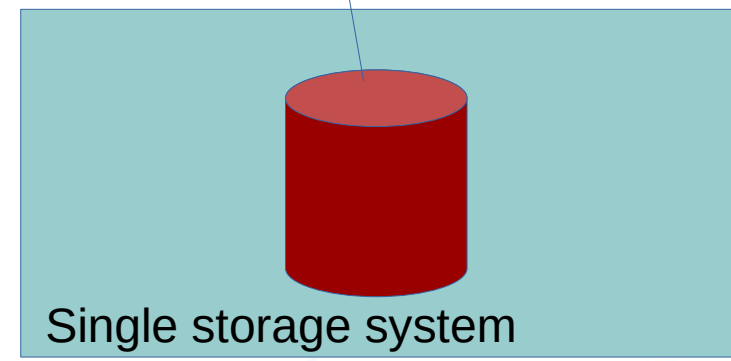
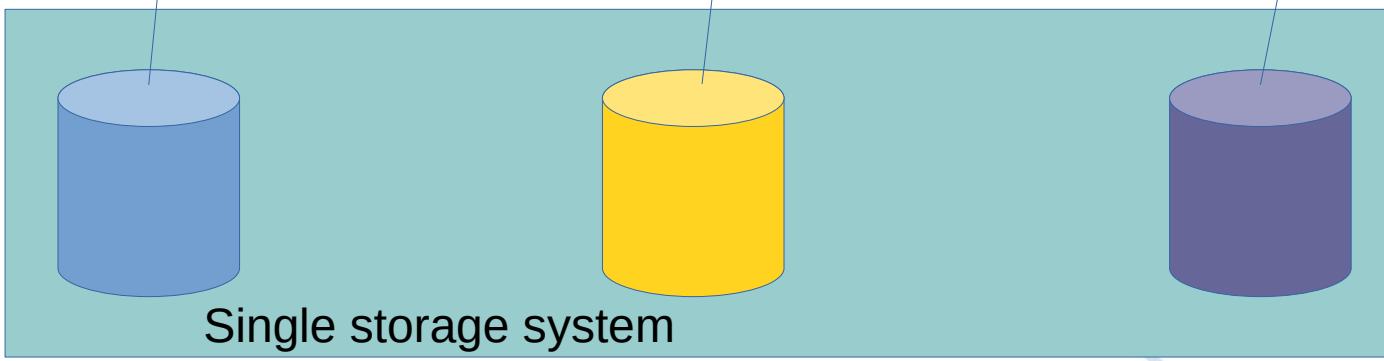
QoS:E

QoS: **GOLD**
Bandwidth: xxx
Latency: yyy
Durability: zzz

QoS: **SILVER**
Bandwidth: xxx
Latency: yyy
Durability: zzz

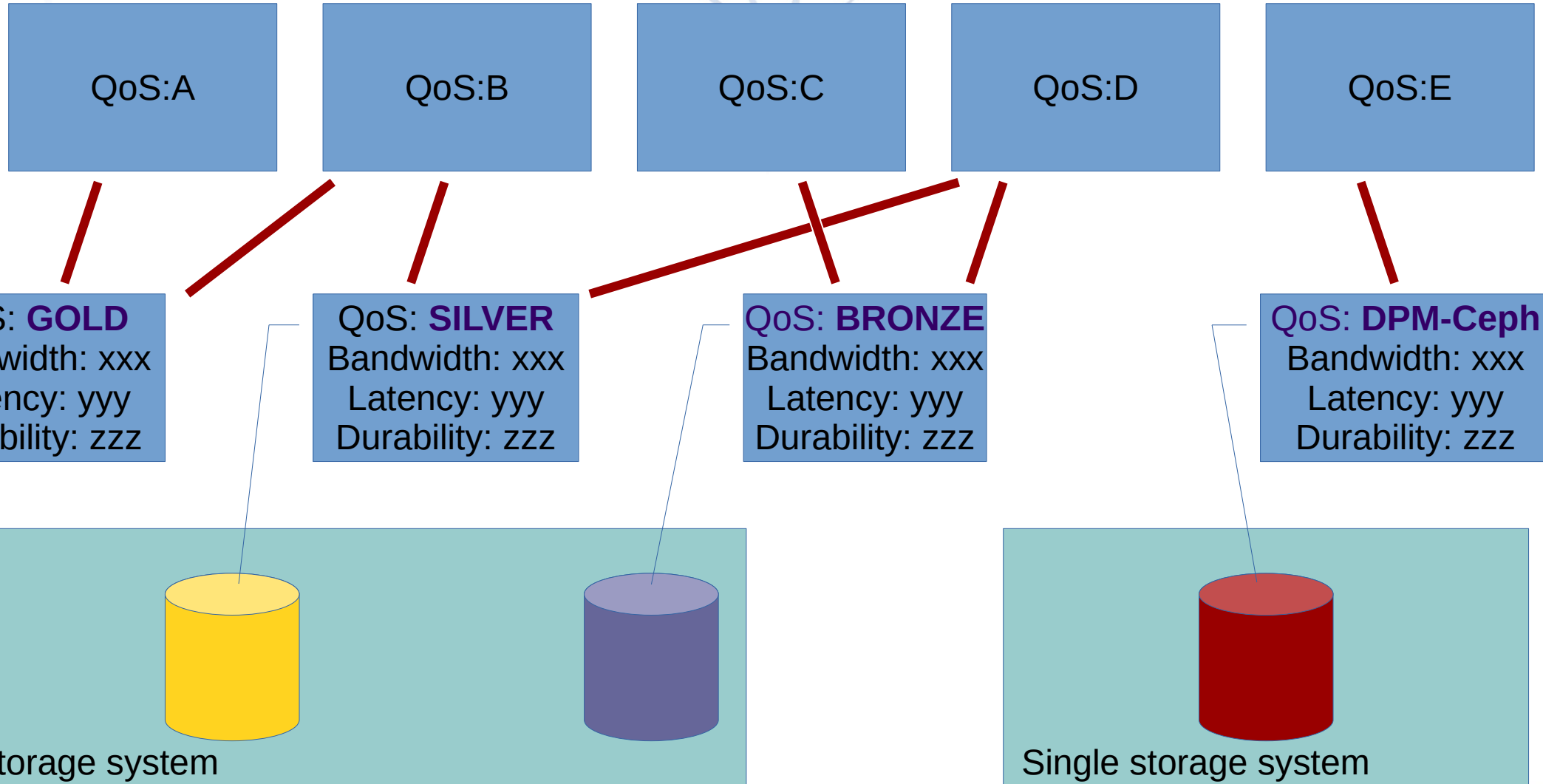
QoS: **BRONZE**
Bandwidth: xxx
Latency: yyy
Durability: zzz

QoS: **DPM-Ceph**
Bandwidth: xxx
Latency: yyy
Durability: zzz





Concepts: VO-QoS-Policies



General approach in Task 2.2

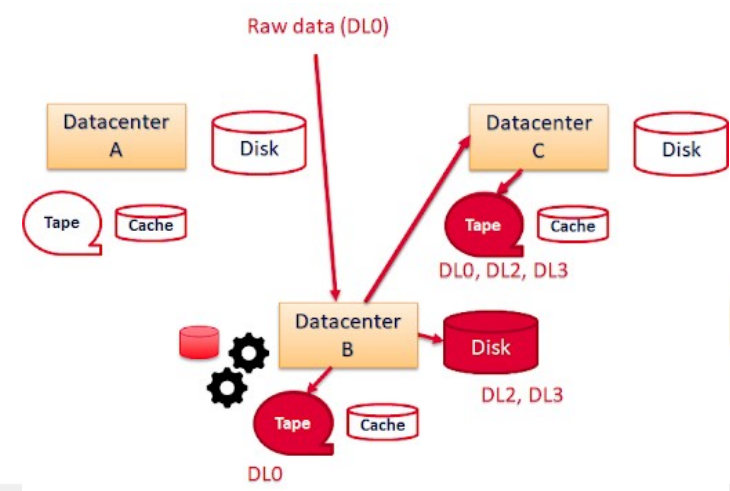
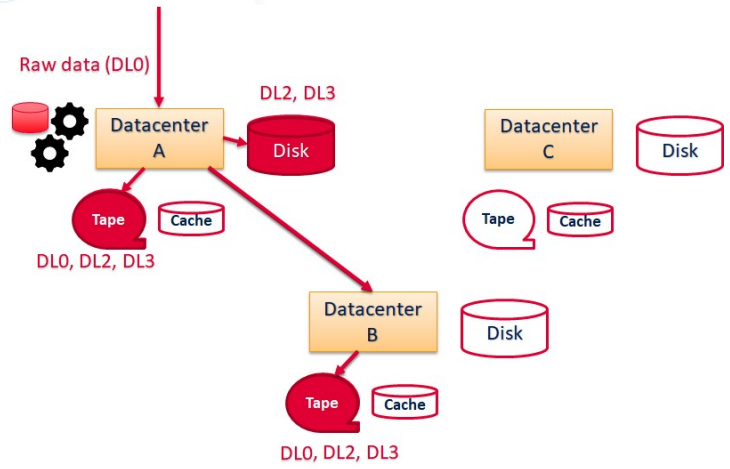
- **Capture** each ESFRI communities QoS requires.
 - Initial interview, to go through the document template, describing what are the different elements and what they mean.
 - The ESFRI representatives discuss the document with other scientists in their community.
 - This might trigger further, ad-hoc meetings to discuss details.
 - Once document is complete, it is presented at the T2.2 regular meetings.
- **Build** a coherent document that takes input from these documents
 - More than just chapters for each community.
 - An initial version, that will contain some open questions
- **Update** the document, based on further discussion.



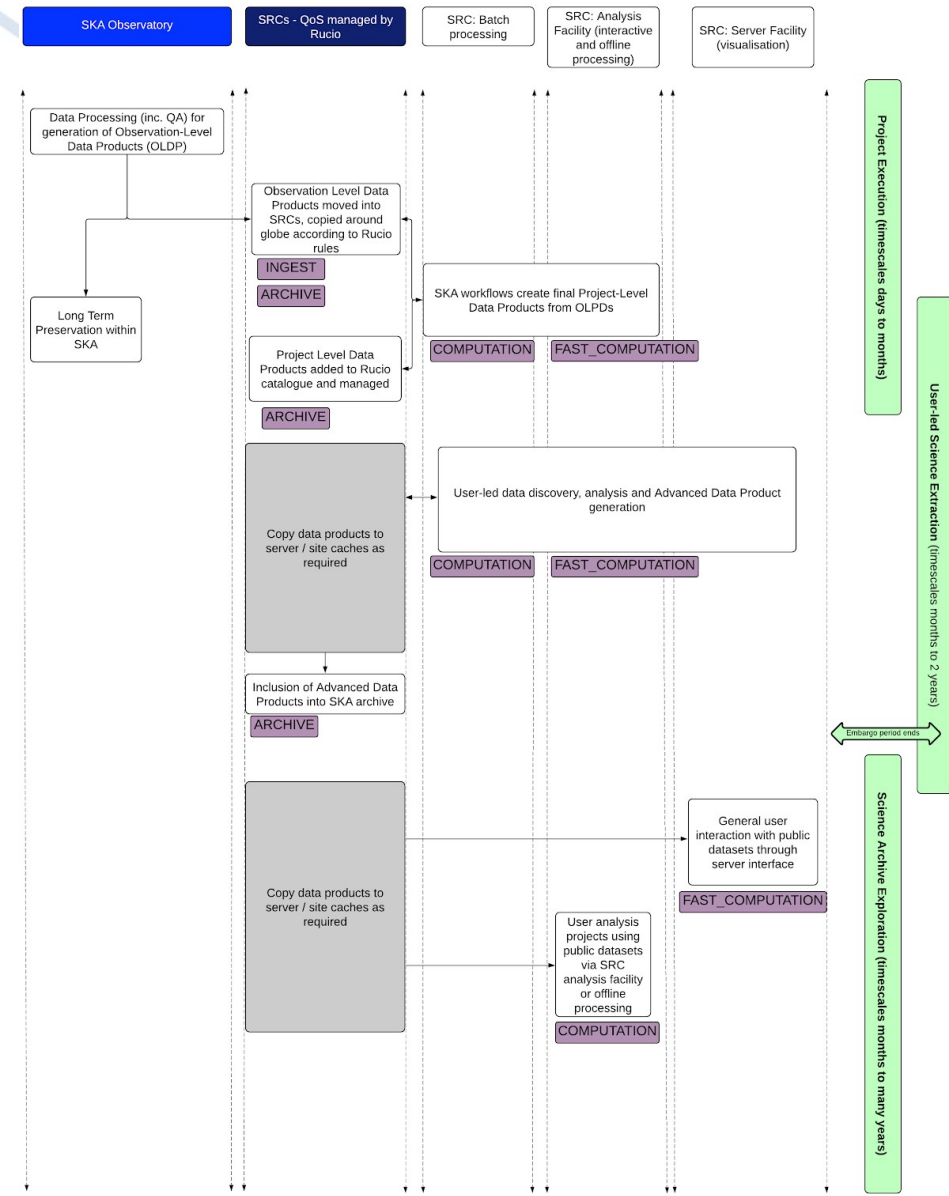
Document format

- The ESFRI-specific QoS document is split into four sections.
- The **QoS policies section** is a table, containing information about how the experiment sees QoS.
 - Name, Where it is used, Important characteristics, Example media.
- The **data life-cycle/work-flows section** is more free-form.
 - Describes the best guess at how data will be handled.
- The **interactions with Rucio** section describes how the experiment framework should achieve the desired QoS
 - List of operations to satisfy life-cycles
- The **use-cases section** provide terse description of user interactions.
 - Bring previous three items together.





CTA workflow (cred. Nadine Neyroud)



SKA data lifecycle (cred. Rohini Joshi)



Current status

- **Received documents** from ATLAS, CTA and SKA
- Document from FAIR is **in preparation**.
 - Killian and Marek for FAIR,
- Next step is to build the **combined document**, providing a single reference for:
 - Driving testbed deployment.
 - Driving software development.
 - Testing QoS support in Rucio.
 - Stimulate further discussion within ESFRI communities.
- This document is (necessarily) an **initial version**:
 - Anticipate an updated version, based on feedback from ESFRI communities



Storage endpoint deployment



Testbed: the storage QoS classes

- Currently, we have **storage endpoints** from (in no particular order):
 - CERN, SurfSARA, CNAF, Napoli, Roma, IN2P3, PIC, LAPP, GSI and DESY.
- There is a **wiki page** that documented what storage technology is being used.
- From this, devised five **storage QoS classes**:
 - JBOD, RAID, EC1, EC2, TAPE.
- Five is a **balance** between
 - too few → difficult to make specific requirements; each class is too general.
 - too many → difficult to work with; each class is not general enough.



The “Aleem” QoS demo

- Answering the **question**: what can we do with what we have now?
- How it **works**:
 - In Rucio, label RSE endpoint, using the attribute **QoS=<storage-class>**
 - When creating rules, use this attribute to describe desired storage QoS
 - E.g., **(QoS==JBOD || QoS== RAID) && Site=IN2P3**
 - Rucio takes care to copy data to specific locations.
 - QoS transitions involve adding or removing rules.
- **Positive**: we can do this right now
- **Negative**: we lose the VO-QoS policy level abstraction
 - This makes the system fragile; changes become difficult.
- Demoed with LOFAR data, ESFRI representatives are replicating this.



Next steps

- FAIR complete their QoS document.
- Complete 1st version of the combined ESCAPE QoS document.
- New version of **Rucio** is released:
 - Once deployed on the ESCAPE testbed, we can start experimenting with the new VO-QoS-policy support.
 - Need to create a mapping between VO QoS Policy and Storage QoS Classes (e.g., in CRIC)
 - Rucio needs to obtain this mapping
- Start **integrating** QoS work with computation workflows:
 - Somehow tag workflows with desired QoS
 - Rules are added as a precursor to running analysis jobs.



Thanks due to ...

- Thanks to everyone for their help so far.
- In particular the ESFRI representatives in their work building the QoS documentation:
 - Mario, Frederic, Nadine, Rosie, Rohini, Kilian and Marek
- ... and the sites for providing storage and QoS information
- ... and Aris in helping establish the QoS attributes.



Thanks for listening!

