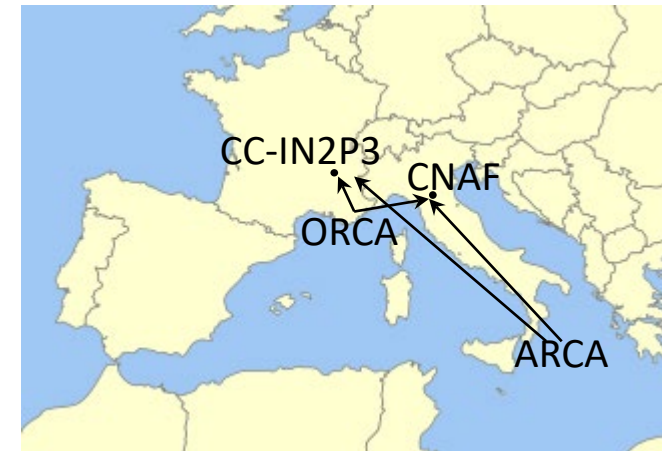


KM3NeT activities during the DAC21 week

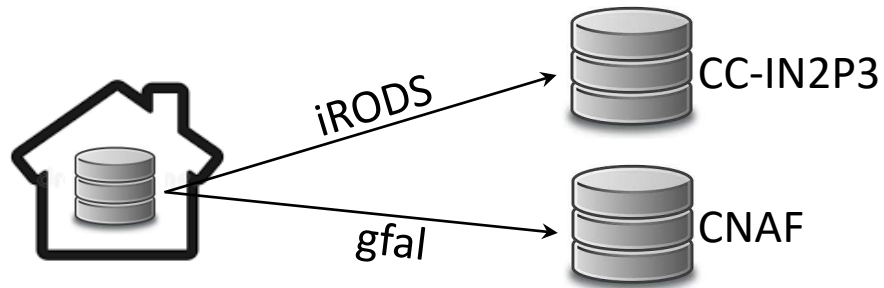
Mieke Bouwhuis (NIKHEF)

- Use cases are similar to regular procedures
- KM3NeT is under construction
- KM3NeT numbers (today)
 - ORCA (6 strings) and ARCA (8 strings)
 - Data rate sea → shore: 2 – 2.5 Gb/s
 - On-shore data filter output: 1 Mb/s
 - 6-hour runs (24/7): 3 – 5 GB/run
 - ORCA shore → CC-IN2P3+CNAF: 12 GB/day
 - ARCA shore → CC-IN2P3+CNAF: 20 GB/day
- Increasing detector size by deploying more strings (115 ORCA; 230 ARCA)
- iRODS current data management tool at CC-IN2P3 (CNAF none)
- ESCAPE Rucio instance for all use cases
- Questions format DAC21 week

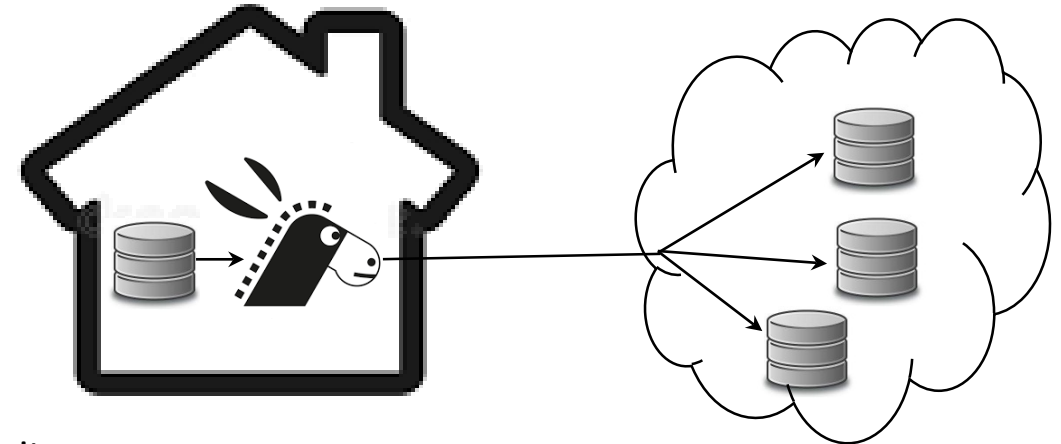


Name	Ingestion from the shore station and replication
ID	KM3NET001
Goal/Aim	Ingestion of KM3NeT raw data from the shore station (remote "RSE") into the data lake, replication of the data onto tape and disk, and deletion of the data at the shore station.

TODAY:



USE CASE KM3NET001:



- Advantages:
 - Only upload and replicate to have the data available at multiple sites
 - Replicate to have SAFE (permanent) and FAST (for calibration) storage
- To verify:
 - Reliability compared to iRODS (however only using current ORCA output)
- Procedure:
 - daily cronjob that checks for data
 - decides on destination (data sets/containers in Rucio)
 - script with rucio commands with the rucio client as a singularity image
 - delete the data from the shore station

Name	PMT calibration
ID	KM3NET002
Goal/Aim	Apply the PMT calibration to a raw data file that is retrieved from the data lake, and ingestion of the calibrated data into the data lake.

- **Advantages:**
 - Replica on disk available for fast access
 - Easier access from other data processing sites
- **Result:**
 - Upload the calibrated (and reconstructed) data for high level analyses
- **Procedure:**
 - Uses the data that were ingested in use case KM3NET001
 - Rucio client running at the data processing site
 - KM3NeT data processing software

Name	ORCA high level analysis - data ingestion
ID	KM3NET003-a
Goal/Aim	Convert processed and calibrated data to DST

- Similar to use case KM3NET002:
 - Fast access
 - Easy access from different sites
- Procedure:
 - Data already in the data lake
 - Usage of ESAP (not using DaaS)

- Other use cases, not completely in DAC21 scope:

Name	ORCA high level analysis - ESAP availability
ID	KM3NET003-b
Goal/Aim	Perform a high level analysis on DST retrieved from the data lake.

Name	Integration of a legacy data archive in the data lake
ID	KM3NET004
Goal/Aim	Linking an iRODS data archive (legacy archive) to the data lake such that the data from the legacy archive are findable in Rucio, and can be replicated.