

Netherlands Institute for Radio Astronomy

# LOFAR Data Management

#### Aras Bozkurt Data Engineer

The 19th International dCache Workshop

#### Overview

- LOFAR & data flow
- Long-Term Archive (LTA)
- Data staging & access
- Storage rebalancing
- Conclusion





### LOFAR - Low Frequency Array



The Low Frequency Array is a multi-purpose radio telescope. Its main application is astronomy at low frequencies (10-250 MHz).

Its heart is in the Northeast of the Netherlands, but LOFAR stations are spread over the whole of Europe



### LTA – Long Term Archive



- Datastreams captured by LOFAR are sent to a central processing (CEP) unit at the University of Groningen.
- On CEP, datastreams are combined and processed further.
- After initial processing to reduce the data volume and assess the quality is completed, the scientific data are transferred and stored in the LTA.
- LTA is used to manage data storage and provide data access to users.



### LTA – Long Term Archive



Growth across LTA sites has caused uneven data distribution due to different storage capacities and placement policies.

1.2 PB of LOFAR data has been migrated from SURF (Amsterdam) to FZJ (Jülich) using the FTS service operated by SURF and the LOFAR Stager service.



### LOFAR Stager

	HOME	SEARCH [	DATA E	ROWSEPR	DJECTS I	HELP												
	PULP Data (total 222) - edit columns stage selected																	
#	Project	Creator	Privileges	Release Date	DataProduct Identifier	dataType	Array Beam	Array Beam ID	Pipeline	DataProduct Type	Ingestion Date	File Format	Filename	Valio Flag	l Dirty Flag	Storage Writer	Storage Writer Version	File Content
1	LT5_004	4 AWTIER0	2	2017-05-13	17608452	CoherentStokes	CoherentStokesBeam	24E51D7E278B3772E053C016A9C37E40	LOTAAS-	PULP Beam	2015-11-19	PULP	L404242 SAP002 B073 P000 bf 313e1250.ta	ur 1	0	unknown	unknown	11
									P0333A/PULP		12:48:29	tarball						
2	✓ LT5_004	4 AWTIER0	2	2017-05-13	17608451	CoherentStokes	CoherentStokesBeam	24E520780CB53FA3E053C416A9C367B2	P0333A/PULP LOTAAS- P0333A/PULP	PULP Beam	12:48:29 2015-11-19 12:49:19	tarball PULP tarball	L404242_SAP002_B072_P000_bf_f6f0a309.ta	r 1	0	unknown	unknown	11

- Stager service moves data from tape to disk for external transfer.
- Users can stage data via the LTA web interface, selecting files from the catalog.
- Once staged, data is pinned on a dCache disk pool and kept for at least one week, ensuring reliable access for users.



### FTS - File Transfer Service

#### **Data Migration Strategy**

- 1.2 PB of LOFAR data divided into 150 TB batches
- Each batch staged and transferred every 2 weeks
- Batch size based on available dCache disk
  pool capacity

#### **Transfer Management**

- Transfers handled by FTS (File Transfer Service) from CERN
- SURF operates FTS3 instance actively used by LOFAR transfers.
- Users submit and track transfers using FTS
  Client

File ID	File State	File Size	Throughput	Remaining	Start Time	Finish Time	Staging Start	Staging End			
+ 7288116	FINISHED	6.15 MiB	30.76 MB/s	-	2021-03-22T09:47:25Z	2021-03-22T09:47:27Z			Log		
🏫 srm://srm.grid.sara.nl:8443/pnfs/grid.sara.nl/data/lofar/user/sksp/archive/AGLOW/logs/scheduler/2018-09-01.tar.gz											

🛓 srm://srm.grid.sara.nl:8443/pnfs/grid.sara.nl/data/lofar/user/sksp/distrib/distrib\_natalie/2018-09-01\_4.tar.gz

(https://doc.grid.surfsara.nl/en/)



#### Conclusion

- LOFAR generates large volumes of scientific data across Europe.
- The Long-Term Archive (LTA) manages this data for long-term storage and user access.
- Distributed infrastructure and coordination across data centers enable scalability and reliability.
- LOFAR Stager and FTS support efficient data access and transfers.



## Thank You for Listening

