Using the REGARDS Framework for the renewal of the CNES archive system

Julien Petiton
CNES, 18 av E. Belin, 31401 Toulouse Cedex 9, France

Context

For the long-term preservation of space mission data, CNES has developed the STAF service. The STAF service was introduced in 1995 and has been steadily improved since then. STAF infrastructure is currently in version 3 (“STAF v3”).

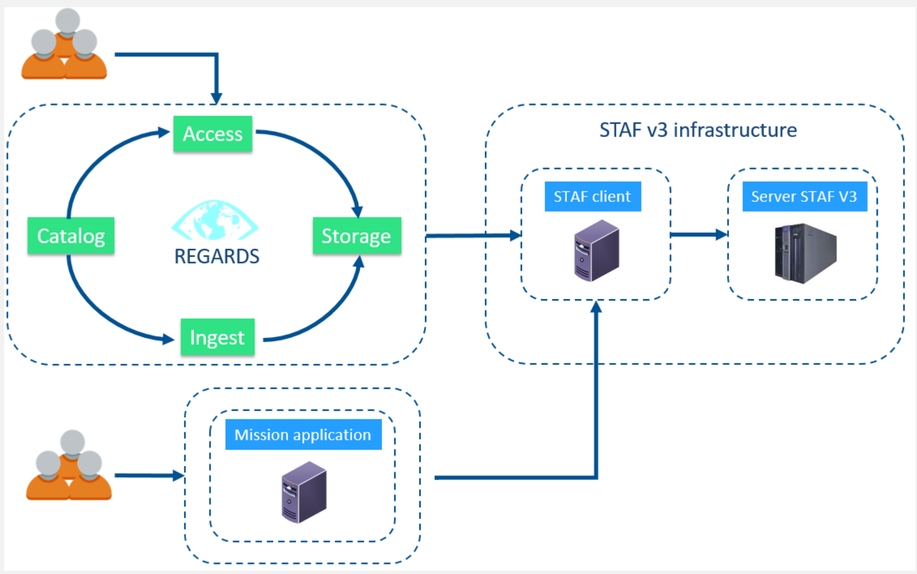
Unlike STAF v3, which uses dedicated infrastructure and software, the STAF redesign ("STAF v4" project) relies on components made available to the Mission Centers or Data Centers, to adapt to their needs in terms of data storage or catalogs. These components are the Datalake object storage infrastructure and the REGARDS access catalog framework (Open Source project available on Github).

STAF v4 project is also an opportunity to improve governance of the CNES archives.

STAF v3 architecture

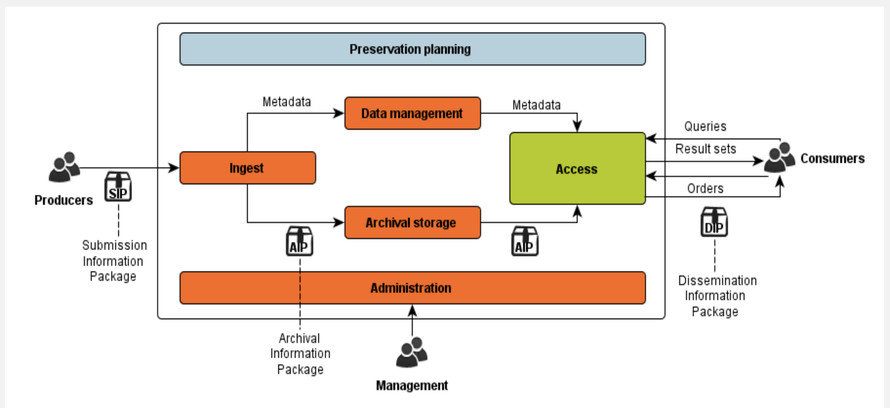
The service needs to evolve for the following reasons:
- To handle a growing archive volume
- End of maintenance for SL8500 libraries
- Production of T10K cassettes SOLARIS stopped
- Software obsolescence (STAR client)
- The STAF is only managed by STAR client in command line (no user interface, no access right managed)

Assessment of 25 years of use:
- 4 petabytes (~57 million files)
- No data loss
- Archive of exclusive data
- Data archived but not referenced in a catalog
- Archive managed by REGARDS and by other applications (resulting unreferenced and non-usable data)



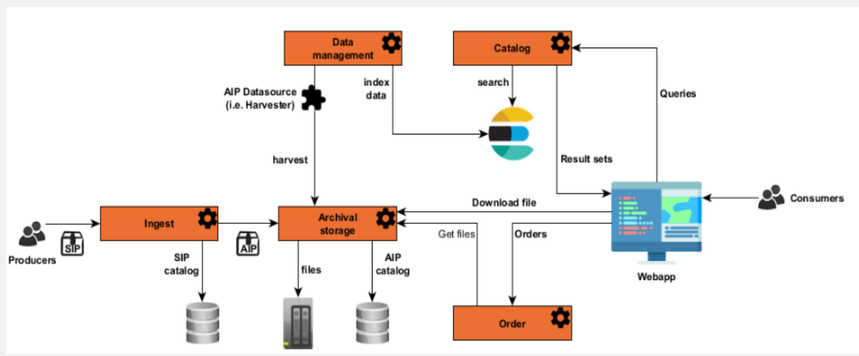
Framework REGARDS

- **Generic** software (Open Source available on Github)
- Development started in 2015
- Using for CNES archives & Mission Center (SWOT)
- Implement the **FAIR** principles
- Implement **OAIS** functionnal model (CCSDS)
- Composed of a back end and a front end
- Easily **adaptable** & **configurable** to various space projects



[Back End]
- Microservices architecture
- Each microservice matches an elementary REGARDS function
- Plugin mechanism to extend functions of microservices and web interface
- Each microservice exposes a REST or AMQP API

[Front End]
- Provides a user and an administration web interface
- Provides enhanced IHM configuration capabilities

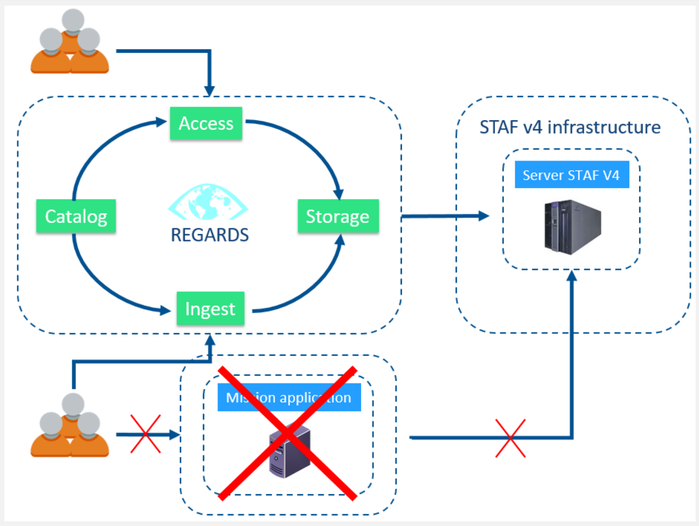


Data type	Filename	Ingestion date	Stored	Filesize	Storage duration (ms)
L0A_HR_Packet	SWOT_IVK_20230406T07...	04/06/2023 08:40:07	true	2147433876	57083
L0A_HR_Packet	SWOT_IVK_20230406T07...	04/06/2023 08:40:07	true	2147433876	58384
L0A_HR_Packet	SWOT_IVK_20230406T07...	04/06/2023 08:30:08	true	2147433876	55900
L0A_HR_Packet	SWOT_IVK_20230406T07...	04/06/2023 08:30:08	true	2147433876	60528
L1_GPSP_RINEX	SWOT_L1_GPSP_RINEX_1...	04/06/2023 08:23:00	true	1290240	1376
L2_RAD_OGDR	SWOT_OPRAD_2PaS000_...	04/06/2023 08:22:27	true	5376000	1816
INERTIAL_GYRO_10	SWOT_INERTIAL_GYRO_1...	04/06/2023 08:21:42	true	56717509	1662
L0A_GPSP_Packet	SWOT_IVK_20230406T07...	04/06/2023 08:21:03	true	1814445	116
ECOM_PFSW_TM_HKAN...	SWOT_STR_1PaS202304...	04/06/2023 08:21:02	true	84721062	2375

Capture of REGARDS IHM with storage value

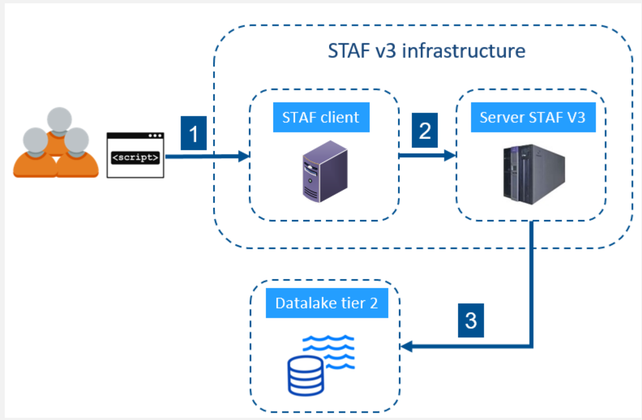
STAF v4 architecture

- STAF v4 is a componant of the CNES **datalake** infrastructure (tier 3)
- Capacity of **25 petabytes** (~400 million files)
- Object storage (**S3**)
- Implement user interface & access right managed by REGARDS
- Only accessible through **REGARDS** (interface & REST API)
- Infrastructure evolution transparent for the business
- Build a metadata catalog "storage" similar to that of the datalake infrastructure
- Implement an overview of the archive by REGARDS catalog
- The renewal of the service allows to make an **inventory of the obsolete archive**

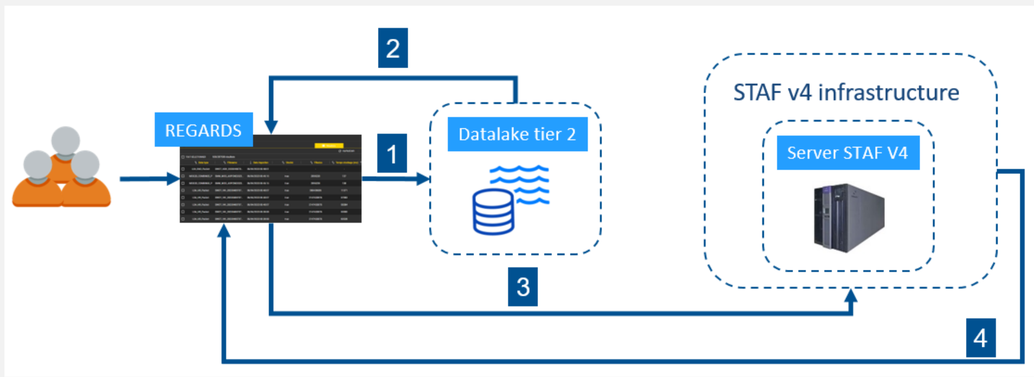


Use case migration

- [1] Request to STAR client to extract files
- [2] Star client sends a request to STAF v3 to extract files from the target tape
- [3] Extraction & copy on the datalake tier 2 (disk)



- [1] REGARDS sends request to the tier 2 datalake (disk)
- [2] Data referenced in REGARDS (building metadata)
- [3] REGARDS request STAF v4 to copy files
- [4] File path updated in the catalog (with integrity control)



Planning

- Migration of only valid files after inventory
- Migration is realized by businesses with the support of SERAD