How are ALMA data products packaged?

Jared Anand Crossley - 2023-04-18 - Archive & Data Retrieval

For Cycles 5 and onwards, raw data and processed data are ingested into the archive as individual files. Upon retrieval - optionally - they are packaged together into tarfiles for better downloads. Data from earlier cycles are directly stored as tar files in the storage system.

A more detailed description of the data products can be found in the document ALMA QA2 Data Products that can be found on the [Documents page in the Science Portal](#).

Raw data

The raw data are written in ALMA Science Data Model (ASDM) format with file names

<ASDM UID>.asdm.sdm.tar
example: `uid___A002_X7d44e7_X956.asdm.sdm.tar`

Processed data

Processed data sets include:

1. README file
2. FITS images
3. Calibration and imaging scripts
4. Quality Assurance (QA) report/plots
5. Calibration tables
6. Log files

The processed data file holding the FITS products have file names

<PROJECT CODE>_<OUS UID>_<X>_of_<Y>.tar
example: `2017.1.01419.S_uid___A001_X1284_X1ef2_001_of_001.tar`

Where the `OUS UID` is the unique identifier for the Observing Unit Set (OUS), X is the file number, and Y is the total number of fits product tar files for this OUS.

The processed data file holding the auxiliary products (scripts, logs, quality assurance, calibration tables, ...) have file names
<PROJECT CODE>_<OUS UID>_auxiliary.tar

A CASA measurement set of calibrated visibilities can be created using the scripts contained in the processed data set by following the instructions outlined in this KB article: How do I obtain a file of calibrated visibilities ("measurement set") for ALMA data?

Cycle 1-4

In Cycles 1-4 the processed data products are packaged (mostly) into the same
<PROJECT CODE>_<OUS UID>_<X>_of_<Y>.tar
tar file.

Cycle 0

In Cycle 0, raw and processed data were ingested within the same tar file (some early deliveries to PIs in Cycle 1 were also made this way, but are not retrievable from the archive). These tar files contained processed products 1-5 listed above, plus:

1. Raw CASA measurement set
2. Calibrated CASA measurement set

The older tar files have file names

<PROJECT CODE>_<DATA DELIVERY DATE>_<X>_of_<Y>.tar

example: 2011.0.00465.S_2012-08-03_009_of_011.tar

As a consequence of including the raw and calibrated measurement sets, these older tar files are much larger than the processed data files currently ingested. The raw ASDM was not included in these older files but can still be obtained by submitting a query to the ALMA Helpdesk.