GBT Ingestion

Meeting on 12 Jun 2019 Rick Lively [X] Thomas Chamberlin

Action Items

- ☑ Bitbucket account for Thomas Rick Lively [X] 14 Jun 2019
- ☑ Confluence account for Thomas Rick Lively [X] 14 Jun 2019
- ☐ Ask about ownership of GBT ingestion code Rick Lively [X] Thomas Chamberlin
- ☐ Update ingestion documentation Rick Lively [X] **14 Jun 2019**
- ☐ Provide latest schema diagram for Thomas Rick Lively [X] **14 Jun 2019**
- ☐ API for GBT ingestion Rick Lively [X] Thomas Chamberlin
- ☐ SQLAlchemy ORM model for the GBT metadata db Thomas Chamberlin

Current Ingestion

Currently ingestion is invoked via the "ingest" command which invokes the "main" function of archive.py.

It's function is to parse the command-line arguments and determine what type of ingestion (EVLA SDM, VLBA IDIFITS, image, calibration) will occur. Other flags are used to control the ingestion process (for executions blocks to be calibrated or not, whether the input should be ingested into the NGAS file system, whether it's being re-ingested, etc.).

For the ingestion of telescope observations (EVLA, VLBA, ALMA), there are 2 steps in to process:

1. Extracting the metadata from the raw input data (SDM: EVLA and ALMA; IDIFITS: VLBA)
2. Loading the metadata into the database

Step 2 consists of 2 steps:

1. Loading the science metadata (project, execution blocks, scans, etc.)
2. Loading the filegroup and file information necessary to link to the data in NGAS

For GBT data, the appropriate set of flags would include:

- A new flag indicating ingestion of GBT data
- -n : Do not calibrate
- -d : Do not ingest into NGAS
- -r : If the data is being reingested

Note: this assumes that the GBT metadata transfer would be initiated by the "ingest" command which may not necessarily be the case.

Database

Science Metadata (in progress)

Project: example GBT/18B-387

```sql
select *
from project
where name = 'AGBT18B_387'
;
```

Given the projectID from the result can be used to explore the entire structure.

Session

Scan

obsParameter

obsProcedure

Observer

file

history ❓

File Location Metadata (in progress)