Application of Pipeline Calibration

Drew Medlin | NRAO Data Reduction Workshop | 2016

What are we talking about?

- What to do after a Pipeline Processing Complete email.

- Extra flagging, re-derive & apply calibration.

- Rerun for known issues that cause problems.

- Apply existing pipeline calibration to raw data.

- Things to consider when running on your own.

- Remote access to NRAO computing.

Extra flagging

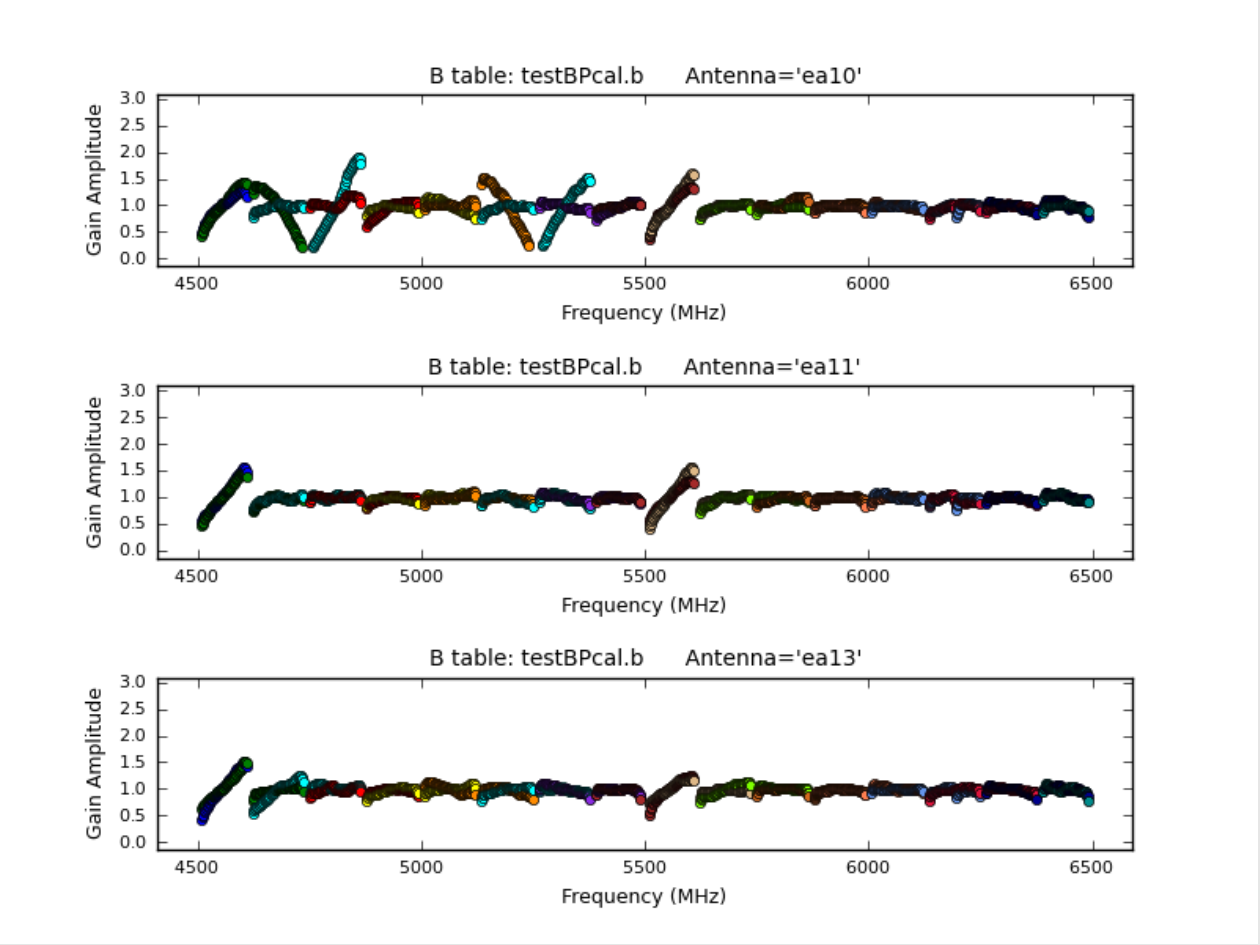
The pipeline may not flag everything needed …

Antenna hardware issues – RFI – may flag good data – ~~Alien signals*~~

*Theoretical only at this time

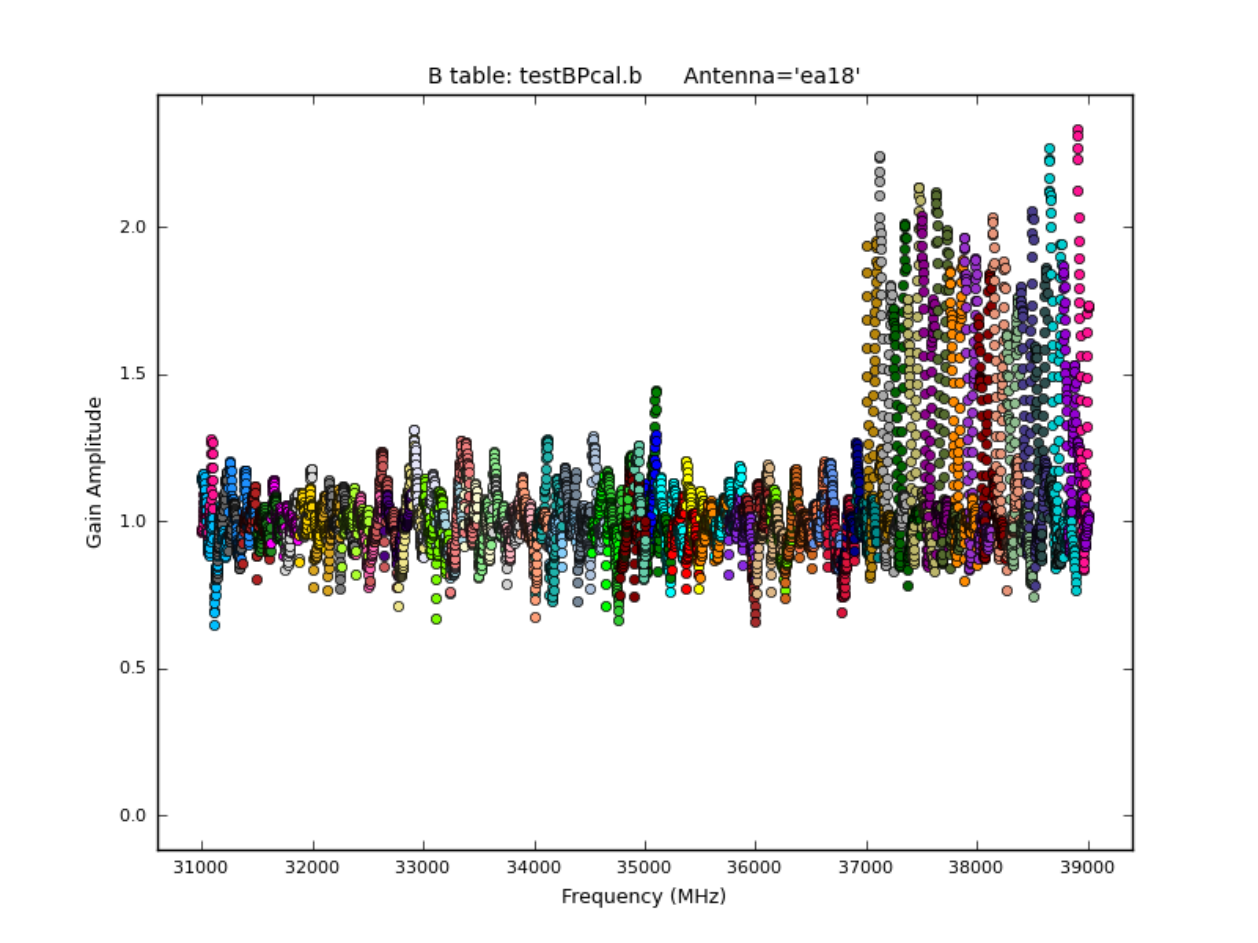
Additional flagging

ea10 bandpass, bad data (DTS issue); ea11, ea12 OK



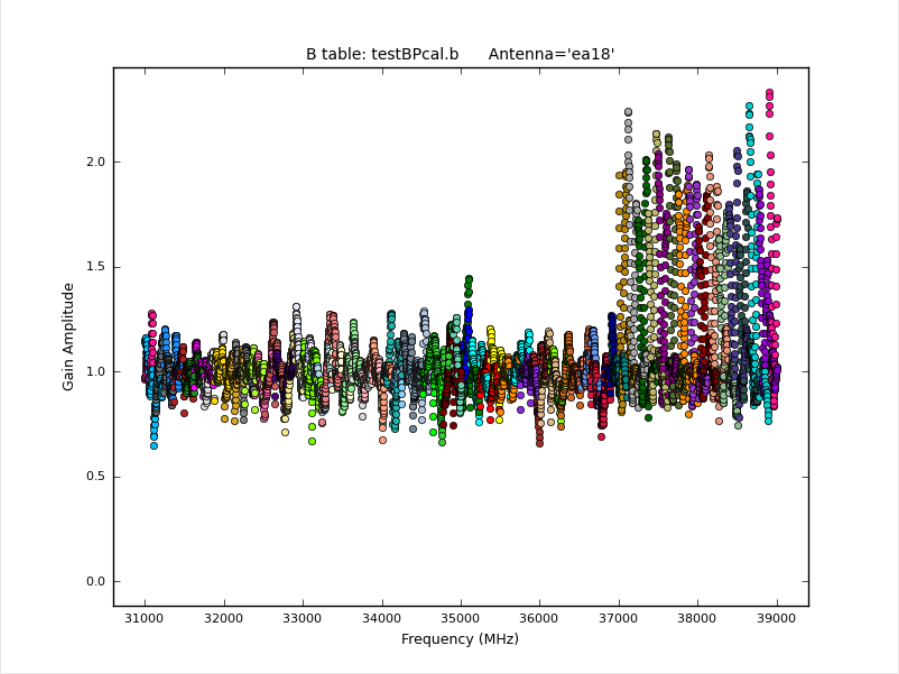
Additional flagging

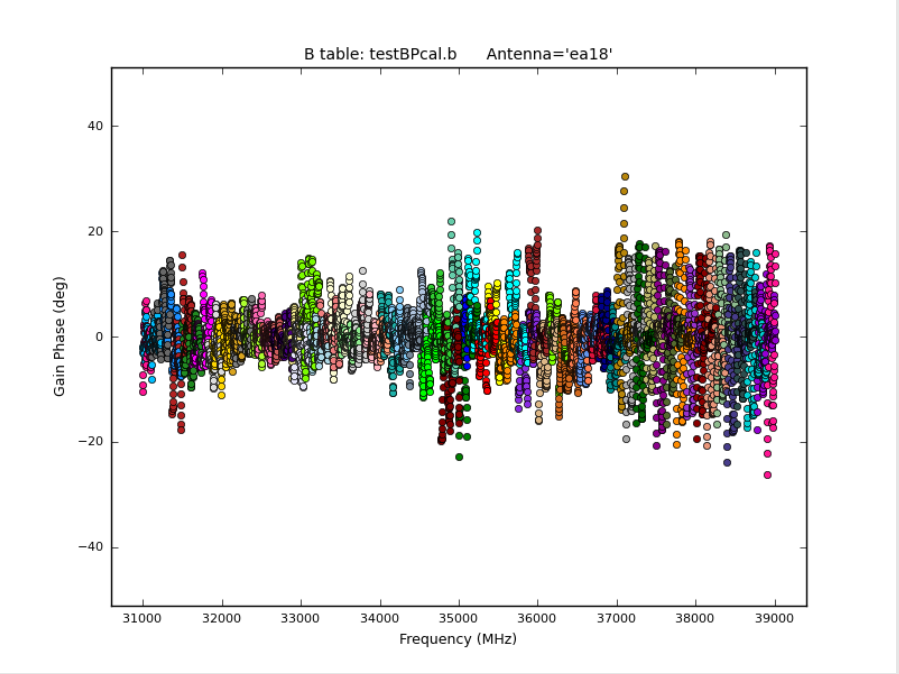
ea18 bandpass, bad data (DTS issue for 37-39GHz)



Additional flagging

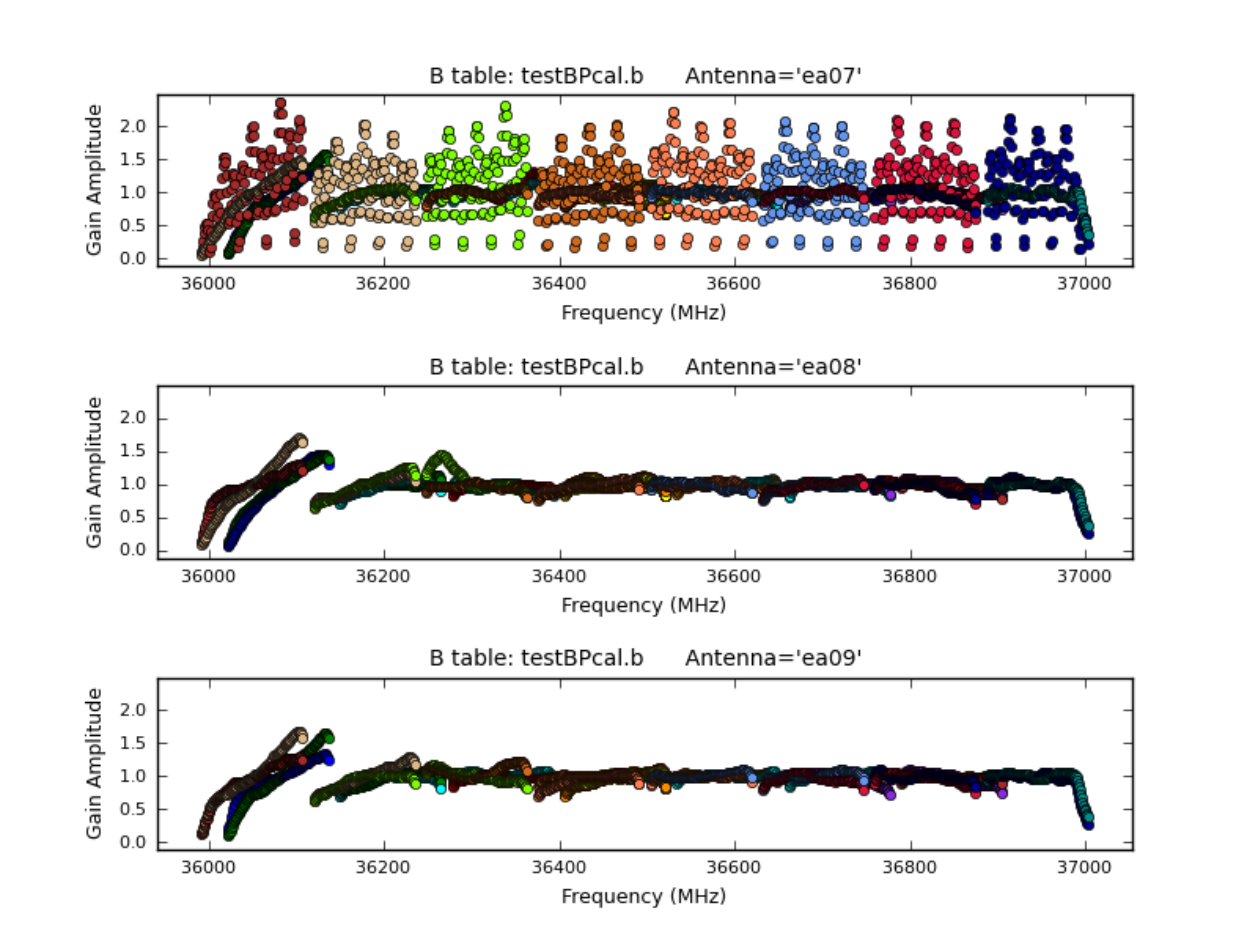
ea18 bandpass **and** phase affected, bad data (DTS issue)





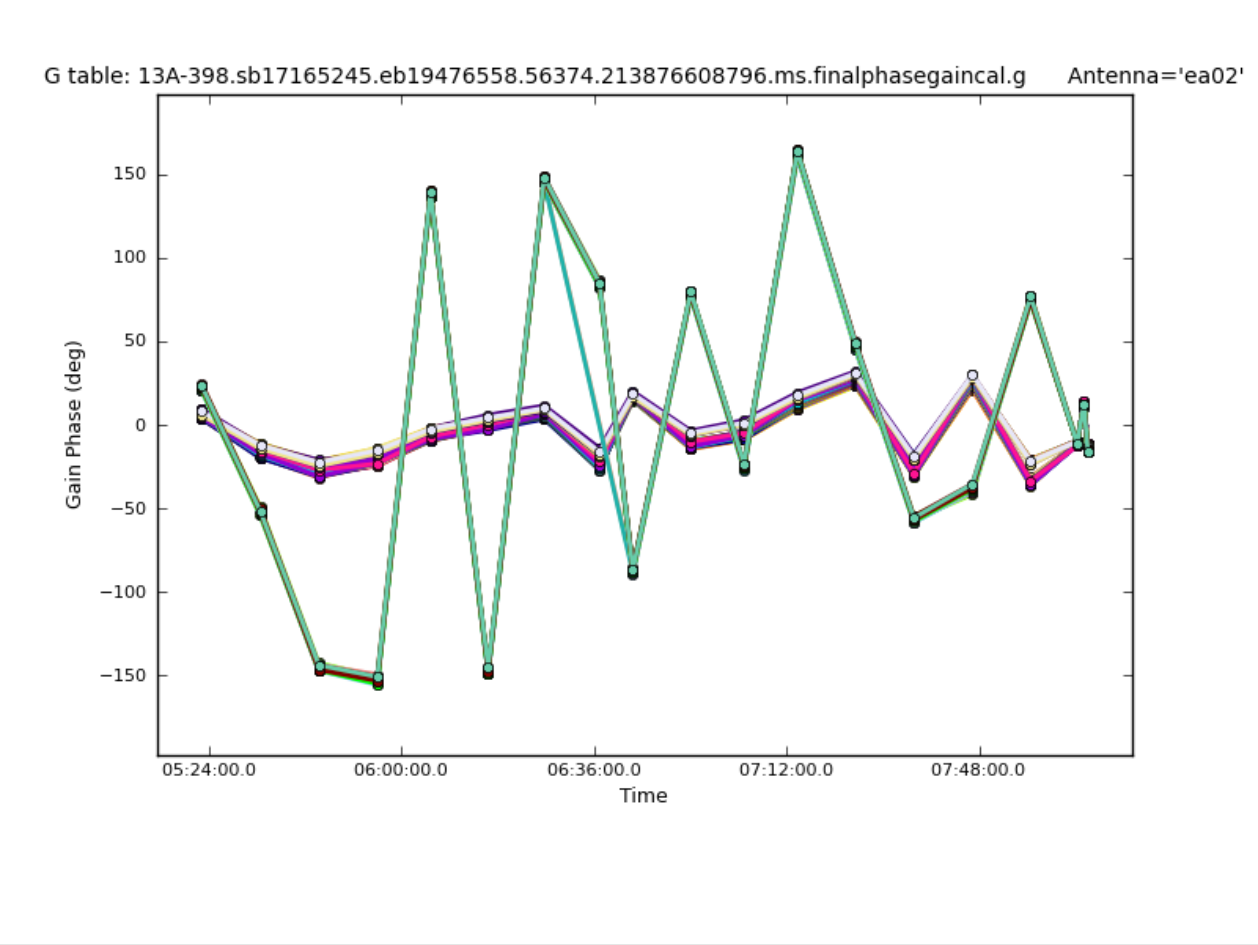
Additional flagging

ea07 bandpass, bad data (DTS issue); ea08, ea09 OK



Additional flagging

ea02 phase jumps for some spws



Additional flagging

- Carefully check your data and the calibration from the pipeline.

- If extra flagging **ONLY** on your science target(s), no recalibration of your data required: use CASA as usual.

- If any of your **calibrator sources** require additional flagging, you should re-derive the calibration with your additional flagging included.

- To recalibrate your data using the VLA Calibration Pipeline, there are two good starting points:

    - Pipeline calibrated MS

    - Raw data (SDM-BDF) >> (see Scripted Pipeline webpage)

Additional flagging: Cal'd MS

1. Create a pipeline recipe **without** Hanning smoothing:

   - Navigate to your CASA installation with the pipeline.

   - Go to /pipeline/pipeline/recipes inside your CASA installation.

   - Copy hifv.py and **rename** to hifv_nohanning.py.

   - Comment out call to hifv_hanning (# symbol)

   - Save changes.

2. Inspect the calibrated MS and flag as needed in CASA.

3. Create a **new directory** and copy **ONLY** the calibrated and fully flagged MS to this new directory. **No** other files should be copied.

Additional flagging: Cal'd MS

4. From your new directory with the flagged MS, start CASA:

   casa --pipeline (be sure to start the same version where you were editing a new recipe!)

5. Clear the *calibration* using **clearcal** with addmodel=False.
   *See the pipeline web page (special topics) for details.

6. Next, run the **clearstat** task in CASA.

5. import pipeline.recipes.hifv_nohanning as hifv_nohanning

6. hifv_nohanning.hifv(['MSname'])

7. Wait again while the pipeline runs.

Rerun for known issues (scripted pipeline)

The pipeline may choose the worst possible option …

- Problems during the flux or delay calibration scan(s)
- Bad reference antenna or setup issues
- Run the Scripted Pipeline (see Scripted Pipeline webpage)

Rerun for known issues (scripted pipeline)

1. From the SMD-BDF, create a MS and apply online flags. You may request the online-flags-applied MS from the NRAO archive.

2. Carefully inspect your data to determine what additional flagging is required.

3. Using the **flagdata** task in CASA, you will create a text file (e.g. additional_flags.txt) for all additional flags needed (no need to apply these flags now!) We recommend:

   flagdata(action='none',mode='manual',savepars=True,outfile='additional_flags.txt', antenna='ea##,ea##', scan='scan#,scan#')

4. Rerun the above until all additional flags have been added. Through this approach, all the flags will be accumulated in a single text file that may be applied on the data through one application once all extra flags have been recorded.

Rerun for known issues (scripted pipeline)

5. Make a copy of the file "EVLA_pipeline.py" and rename it.

6. Edit your renamed copy of EVLA_pipeline.py to include an extra step of the CASA task **flagdata** just after 'flagall.py' is called.

   flagdata(vis=ms_active,mode='list', inpfile='additional_flags.txt', action='apply', flagbackup=True)

   Please take care to **notice indentation** amounts and to set your CASA flagdata command to the same indentation as the other "execfile()" commands.

7. In a new directory with the SDM-BDF, start the pipeline by calling your edited version of EVLA_pipeline.py

   execfile('/path/on/computer/EVLE_pipeline_edit.py')

8. Wait … wait longer … just come back the next day.

Apply to Raw Data

May only have pipeline calibration & flag tables, no MS

- Calibrated MS held by NRAO for only 15 days:
  Calibration tables, flag tables, weblog archived!

- Local storage limitations:
  Reduced storage needs by only keeping the tables.

Apply to Raw Data

1. Download the correct CASA version with the pipeline.

2. You will need the following
   - SDM-BDF
   - Session_1.caltables.tar.gz
   - *.ms.flagversions.tar.gz
   - *.ms.calapply.txt
   - casa_piperestorescript.py

3. Make a directory called “restoration”.

4. cd to restoration, and create three more directories inside:

   rawdata, working, & products << names must be exact!

Apply to Raw Data

5. Put your SDM-BDF into the “rawdata” directory.

6. Put all the *.tar.gz files and *.txt files into the “products” directory.

7. Put casa_piperestorescript.py into the “working” directory.

8. Go to the “working” directory and edit casa_piperestorescript.py:

    - Insert “../rawdata/” before the SDM-BDF name (mySDM) in the call to hifv_restoredata.

    - Save your changes.

Apply to Raw Data

9. From the "working" directory, start CASA with the pipeline

   casa --pipeline

10. Execute the casa_piperestorescript.py file:

    execfile('casa_piperestorescript.py')

11. Enjoy calibrated data once the process completes.

Considerations

Scan intents correct?

Hanning Smoothing?

Computing time?

Disk space – 3-4X raw size(!)

PL version differences.

CASA version differences.

NRAO cluster available for remote Access

Remote Access: Accounts

Use your **visitor account** (what you're using now)

- Remote processing
- Data staging for download
- Short term work, **NOT** long term storage.

Use your account's **"data" directory**:

- Archive deliveries (future archive will use this)
- Pipeline data requests
- Don't change permissions of this directory!

Remote Access: Node Request

Login with your account username

- ssh nm-####@login.aoc.nrao.edu

Go to nmpost-master and request a node

- ssh nm-####@nmpost-master
- nodescheduler --request 14 1

If you get no email, you are probably queued

- Don't keep requesting more nodes

Exit nmpost-master, then ssh to your assigned node

- ssh nm-####@nmpost###

Remote Access

Interact with your data for reduction and analysis

- SSH and VNC available for working with your data.

Download your data:

- RSYNC, SFTP, SCP, LFTP available.

Too much data to download?

- Use hard disk shipping option.

Need help?

- https://science.nrao.edu/facilities/vla/docs/manuals/computing-resources

Questions?

- VLA CASA Calibration Pipeline information at:

https://science.nrao.edu/facilities/vla/data-processing/pipeline

- CASA Integrated Pipeline & Scripted Pipeline available

- Have Questions?
- Need Help?

- Use the **NRAO HelpDesk**: **https://help.nrao.edu/**

- Submit your ticket under the **Pipeline Department.**

- Please include specific details when submitting HelpDesk tickets.
  (Project code, SB number, CASA/PL versions, errors, etc.)