Calibration Enhancement (CE) Archive Requirements

Prepared for the Archive Coordination Meeting, Feb. 24-25, 2003

Anastasia Alexov

ST-ECF Instrument Physical Modeling Group

Background

- Post Operational Archives (POA) project for the re-calibration of the FOS found it vital to have easy access to 100% of all FOS archived data.

- All FOS data were retrieved; 'getref' ('bestref') was used to update the headers; 'calfos' was used as 1st look at calibration.

- POA project functioned within a STATIC framework -- the data, the calibration reference files and pipeline were "frozen" versions. There were no more FOS observations, and therefore essentially no software updates.

- Similarly, the STIS Calibration Enhancement (CE) project has the same set of framework requirements for working with STIS Side I data.

STIS CE Archive Start-Up

- In Feb. 2002, the ECF/ESO archive had missing STIS POD files due to data backlog problems
  - -> Special request was made to Faith Abney for delivery of all the STIS Side I POD files (completed in late-spring 2002)

- STIS OTFR had to be in reliable working order at the ECF/ESO archive; we needed "OPUS 14.1 SHARE 4.1 * 27 Aug. 2002" release
  - -> Installed and tested in Sept. 2002 by the archive
  - -> Our STIS OTFR "request-window" lasted only until the next the release of OPUS OTFR (in order for all our STIS data to be in the same framework)

STIS CE Archive Start-Up (Cont.)

- Disk space was needed for the data as well as easy access; CPU was needed to assist in the massive STIS CE request processing
  - -> We purchased a SunFire 280R machine with 2 X 900MHz CPUs, 4G RAM, attached to a 5T RAID
  - -> This machine was set up for dual use as both the STIS CE data server/workstation as well as an archive operational machine (OTFR and data I/O done locally)

- Pipeline and calibration reference files had to be within the same frame of reference as the output of STIS OTFR for proper calibration
  - -> Made a "frozen" copy of the calibration reference files
  - -> Used same version of STSDAS and 'calstis' (v 2.13b) for all the data processing

OTFR of STIS Side I Data

- There are 66,095 STIS Side I datafiles in the HST archive (from the launch of STIS to the end of STIS Side I operations in May 2001)

- Divided into 66 lists of 1000 rootnames each for archive requests

- From Oct. - Nov. 2002, 5-10 STIS list were requested weekly (8 weeks). Archive efficiency:
  - Submission -> OTFR -> result **"rate"** started at approximately **6,600/week (10%)**
  - Iterative process of fixing problems by archive group members: web request forms timed-out, access to the database was a bottleneck, transient DB-errors occurred, Sybase bugs had to be fixed, multi processing was initiated, errors in DB tables were noticed, corrupted POD files were found...
  - STIS OTFR **steady state** reached at **10,000/week** by the end (**15+%**)
  - In Dec. 2002, follow-up re-requests of ~120 datasets took place on all previous errors (missing POD, OTFR problems, etc.)

- 99.94% of 66,095 STIS datasets OTFRed - gzip - 110G; **43 errors**:
  - **33** are from proposal 863 proprietary data [Markus Dolensky was able to get these from STScI]
  - **10** are due to: missing POD files (**7**), and OTFR 'fitsverify' errors (**3**)
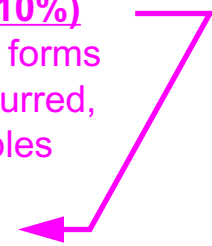
STIS Side I Data Calibration

- Four SunBlade 150 (1G RAM) and 2 CPUs on SunFire 280R, were used to run 'calstis' on all the "raw" STIS data; took 3-5 days per CPU to finish

- 99.11% of 66,052 STIS datasets completed (equivalent of OTFC) - gzip - 835G; **590** errors:

```
num     'calstis' error message
---     -----------------------
320 -> Warning SPTRCTAB `oref$l2j01381o_1dt.fits' not found or can't open.
192 -> Warning PHOTTAB `' not found or can't open.
 75 -> Warning CCDTAB `FILL_ME' not found or can't open.
  1 -> Warning BIASFILE `' not found or can't open.
  1 -> IRAF error: invalid floating point operation (502) [o3ura1070]
  1 -> Output file `o68y010a0_crj.fits' already exists.   [o68y010a0]
```

- 588 errors were actually due to OTFR problems, and/or possible DB replications issues. Re-requested during "secure" times w/o DB replications problems; 74 errors:

```
num     message
---     -------
 72 -> Warning CCDTAB `FILL_ME' not found or can't open.
  2 -> Warning BIASFILE `' not found or can't open.
```

<u>Conclusions</u>

- <u>Enormous</u> task to <u>OTFR 66,095 STIS</u> Side I datasets. This was <u>cumbersome but beneficial</u> for both parties involved - helped the system get better and we got our data.

- Took a lot of <u>fixing and massaging</u> of the system to finish processing the data within a reasonable time frame. Any <u>future requests</u> of STIS Side I data will be "<u>easy</u>".

- We were lucky not to encounter more <u>DB replications problems</u> at the time of the requests; however, there are now <u>uncertainties</u> in our "<u>raw" STIS data</u>.

- There was no easy mechanism of <u>checking</u> which STIS rootnames <u>failed</u> to be <u>OTFRed</u> - one corrupted POD file can botch up several STIS datasets; CE scripts did the checking and reporting to the archive. This could be confusing on the user-end -- <u>both OTFRed files and old "raw"</u> can be given as results.

Conclusions (Cont.)

- Only after running 'calstis' did further clues creep out regarding OTFR problems.

- We thank Faith Abney for providing us with the much needed STIS POD files; Markus Dolensky, Benoit Pirenne and Alberto Micol for the frequent assistance with archive/OTFR issues; and, to Jonas Hasse for watching over every STIS CE archive request - he thought the torture would never end...

- In the end, out of 66,095 STIS Side I datasets, we have 99.94% of the "raw" file and 99.11% of those are fully calibrated. A repository of 65,978 out of 66,095 datasets is solid ground to start the STIS CE project.

- Our work makes the assumption that POD files are fixed and will never to re-ingested or updated. However, if POD updates were to happen in the future, then our paradigm of using a fixed version of data/OTFR will be broken.