

## Using OPUS to Perform HST On-The-Fly Re-Processing (OTFR)

Michael S. Swam, Edwin Hopkins, Daryl A. Swade

*Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218*

**Abstract.** The Hubble Space Telescope (HST) OPUS implementation of On-The-Fly Calibration (OTFC) processing currently provides the benefit of applying the most current calibration algorithms, reference files, and repairs of errant header keyword values (aperture, shutter, etc.) to Level-1b datasets as they are retrieved from the HST archive. While OTFC has performed well, a number of concerns about maintenance and flexibility have resulted in the evolution towards an On-The-Fly Re-Processing (OTFR) System. Also based on the OPUS pipeline architecture, OTFR carries further the notion of creating products for archive users at the time of their request, by completely regenerating calibrated (Level-2) data products for an exposure from the base telemetry files sent from the HST (Level-1a data). By starting processing at this earlier state, and taking advantage of the changes to the data processing software that are made as that software matures, improved, more consistent calibrated (Level-2) data products are produced. Use of OPUS distributed multiprocessing, the relatively small size of HST datasets, and the efficiency of the data processing and calibration software results in a very small impact on the overall time it takes to complete an archive retrieval. There could be an impact to archive research, however, since the archive catalog meta-data will not completely reflect the reprocessed products as they would be delivered to the archive user. This problem will be addressed by performing catalog updates for any major discrepancies. This paper will describe the concerns raised about OTFC, the design of the OTFR pipeline system, and the benefits of using the OPUS architecture in this design.

### 1. The Existing On-The-Fly Calibration (OTFC) System

The Hubble Space Telescope archive at the Space Telescope Science Institute (STScI) has had an on-the-fly calibration system (Swam & Swade 1999) in place for the Space Telescope Imaging Spectrograph (STIS) and the Wide-Field Planetary Camera 2 (WFPC-2) instruments since December 1999. Whenever calibrated STIS or WFPC-2 products are requested from the HST archive, the exposures are automatically recalibrated using the latest reference files, calibration algorithms, and known repairs to the FITS header keyword values. These improved products are then delivered to the user who made the archive request.

## 2. Problems with OTFC and Reasons for On-The-Fly Re-Processing

While the WFPC-2 instrument support group at STScI has been very pleased with the OTFC system, the STIS instrument group has had to do quite a bit more work to get their OTFC output products in optimal form. Because STIS is a more complex and less mature instrument than WFPC-2, over 200,000 header keyword repairs (most only for aesthetic reasons) are specified in the OTFC keyword repair database table (Lubow & Pollizzi 1999) for STIS exposures, compared with around 700 for WFPC-2. This large set of repairs now has to be checked for layering issues as each new STIS repair is implemented.

More importantly, there is separate pre-archive pipeline software that converts HST raw telemetry (Level-1a data) into calibrated data products for loading the archive catalog (Swade, Hopkins, & Swam 2001). Therefore new repairs must be made in both the keyword repair table for data already in the archive, and in the pre-archive pipeline software for any new exposures taken with HST. There are also some cases where repairs to the data format of STIS exposures in the archive are warranted (e.g., time-tag observations of very bright sources, Sahu 2000), and the OTFC system cannot automatically make these repairs, since it currently only fixes header keyword values. The exposures must be pulled out of the archive and reprocessed through the pre-archive pipeline software, with the resulting products then re-ingested into the archive.

For the Near-Infrared Camera and Multi-Object Spectrometer (NICMOS) and the Advanced Camera for Surveys (ACS), the complication of adding multi-file associated exposures is required. This capability already exists in the pre-archive pipeline code.

These problems all point to using the pre-archive pipeline software in the OTFC system, so that each archive request results in reprocessed products using the most complete, up-to-date data processing software. This is the on-the-fly re-processing system, OTFR.

## 3. Design of the OTFR System

The OTFR system makes use of the OPUS pipeline architecture (Rose & Miller 2001, Miller 1999, Swade & Rose 1999) to spread the automated re-processing over a number of nodes in a cluster of CPUs. The pre-archive pipeline software is already implemented using OPUS, so it was straightforward to add OPUS pipeline stages to perform OTFR. Stages were added before the pre-archive pipeline code to collect OTFR requests from the archive and feed them in for processing, and stages were added after the pre-archive pipeline code to collect processed products and feed them back to the archive for distribution.

The stages before the pre-archive pipeline have to:

- poll for OTFR requests from the archive,
- convert requested exposure names into the names of the corresponding raw telemetry files (a many to many relationship, held in database tables),
- request the raw telemetry files from the archive,
- segment raw telemetry files into one exposure per file,
- pass the segmented raw telemetry files to the pre-archive pipeline.

The stages after the pre-archive pipeline have to:

- poll for completed products,
- convert all products to FITS format, if not already,
- return the completed products to the archive,
- clean out completed products from the pre-archive pipeline.

#### 4. Using OPUS in OTFR

The OPUS system provides a number of features beneficial to the implementation of OTFR:

- distributed multi-processing keeps throughput high;
- integration of scripts and programs into pipeline stages is straightforward;
- support for file arrival, time trigger, and blackboard events provides the flexibility for integrating the OTFR pipeline with the archive system;
- a single executable or script can be configured multiple ways using ASCII resource files, allowing one set of pre-archive code to meet many needs (Boyer & Choo 1997).

This last point is very important in meeting one of the goals of OTFR; to allow the pre-archive pipeline software to be used to both populate the archive catalog, and to reprocess exposures for archive users. This reduces maintenance and promotes code reuse, allowing fixes and enhancements to one set of software to benefit both tasks.

#### 5. Issues for OTFR

The use of on-the-fly re-processing does have implications for the state of the archive catalog meta-data. This meta-data, consisting of astronomical target information, instrument configuration, and other ancillary parameters, is populated when the exposure is initially received from HST and processed through the pre-archive pipeline. This implies that the meta-data reflects the state of processing at that instant in time. When an archive user later requests this exposure through OTFR, the exposure is again processed through the pre-archive pipeline code, but this code could have evolved since the original population of the catalog meta-data. This means that the catalog meta-data may not accurately reflect the exposure data at retrieval time. This has implications for archive users, as they search the catalog to find observations that meet certain criteria. To address this issue, the HST archive is committed to updating the catalog meta-data when significant differences exist between the original catalog values, and new values coming out of OTFR. The intent is to eventually automate this process, using special processing flags for the OTFR system.

#### 6. The Future of OTFR at STScI

OTFR is currently being tested, and has been scheduled for deployment at STScI in early 2001 for STIS and WFPC-2. The NICMOS will be added shortly thereafter, while the ACS will be an OTFR instrument from the moment it is added to HST during servicing mission 3B, currently scheduled for fall of 2001. It is projected that the future Cosmic Origins Spectrograph (COS) and

the Wide Field Camera 3 (WF3) will also be OTFR instruments as soon as they are installed in the HST.

**Acknowledgments.** Many thanks to those who reviewed the design and code for the OTFR system at STScI: Dorothy Fraquelli, Chris Heller, Warren Miller, Jim Rose, John Scott, Lisa Sherbert, and Steve Slowinski.

## References

- Boyer, C., & Choo, T. H. 1997, in ASP Conf. Ser., Vol. 145, *Astronomical Data Analysis Software and Systems VII*, ed. R. Albrecht, R. N. Hook, & H. A. Bushouse (San Francisco: ASP), 42
- Lubow, S., & Pollizzi, J. 1999, in ASP Conf. Ser., Vol. 172, *Astronomical Data Analysis Software and Systems VIII*, ed. David M. Mehringer, Raymond L. Plante, & Douglas A. Roberts (San Francisco: ASP), 187
- Miller, W. 1999, in ASP Conf. Ser., Vol. 172, *Astronomical Data Analysis Software and Systems VIII*, ed. David M. Mehringer, Raymond L. Plante, & Douglas A. Roberts (San Francisco: ASP), 195
- Rose, J. F. & Miller, W. 2001, this volume, 325
- Sahu, K. 2000, private communication
- Swade, D. A. & Rose, J. F. 1999, in ASP Conf. Ser., Vol. 172, *Astronomical Data Analysis Software and Systems VIII*, ed. David M. Mehringer, Raymond L. Plante, & Douglas A. Roberts (San Francisco: ASP), 111
- Swade, D. A., Hopkins, E., & Swam, M. 2001, this volume, 295
- Swam, M., & Swade, D. A. 1999, in ASP Conf. Ser., Vol. 172, *Astronomical Data Analysis Software and Systems VIII*, ed. David M. Mehringer, Raymond L. Plante, & Douglas A. Roberts (San Francisco: ASP), 203