Instrument Science Report WFC3-2002-08

WFC3 Science Calibration Plan

Part 2: An overview of the pre-flight data handling system

1 September, 2001

by

I. Neill Reid (STScI) and Frank Bartko (Bartko Science and Technology Inc.)

ABSTRACT

This document outlines requirements for the pre-flight data handling system for science calibration. We discuss both the computer hardware and the software packages necessary for acquisition, formatting, logging, transfer, archiving and analysis of test images taken as part of the WFC3 thermal/vacuum calibration campaign. This discussion document outlines options for implementation of the data handling system, and identifies trade studies that should be completed so that a cost-effective, technically optimal solution can be recommended.

1 Introduction

The Wide-Field Camera 3 (WFC3) is a two-channel imaging camera, designed to replace WFPC2 on the Hubble Space Telescope during the fourth servicing mission. Extensive campaigns will be undertaken to calibrate the instrument before launch, including tests at the sub-component and component level, as well as systems level tests with the fully-assembled optical bench. The WFC3 Science Calibration Plan (SciCal Plan) defines a series of pre-flight systems-level measurements that are designed to verify that instrument performance meets the requirements relevant to scientific observations given in the CEI specifications. A companion plan will be designed for on-orbit calibration.

Since the WFC3 optical bench operates at 0 C and the infrared cold enclosure at –30 C, all tests that require the detectors to be at the on-orbit operating temperature are possible only when the instrument is under thermal/vacuum conditions. As a result, the majority of the measurements described in the SciCal Plan can be carried out *only* in a thermal/vacuum environment. This is in sharp contrast to previous HST instruments, including ACS and STIS, where many science calibration measurements were possible under ambient conditions.

The present document considers and defines the requirements for the overall data handling system architecture that will provide the support needed to accomplish the tasks described in the SciCal Plan. The Calibration Reference files will be placed in the STScI DADS, and consideration is being given to placing other ground-calibration data in the STScI archive. However, those data must also be available for quick analysis, both by WFC3 ICAL and Science IPT members at GSFC/STScI, and by interested members of the WFC3 Science Oversight Committee (SOC). We identify and discuss several software options for achieving these goals and satisfying the calibration requirements.

Finally, since Thermal/Vacuum testing is schedule and cost intensive, we describe several trade studies to evaluate the relative merits of the various options. The results of those studies would identify a technically optimal, cost effective calibration strategy.

2 Background

2.1 The Thermal/Vacuum Campaign

WFC3 optics will be assembled at Ball Aerospace, Boulder, Colorado. Delivery of the fully-assembled optical bench to Goddard Space-Flight Center is scheduled for October 2002 (predicated on a November, 2003 launch date). The thermal/vacuum testing is currently planned for March 2003, and includes tests of instrumental stability (including thermal balance) besides the science calibration tasks described in the SciCal Plan.

2.2 Configuring WFC3 and the Optical Stimulus

During testing, WFC3 will be mounted in the Radial Image Alignment Facility (RIAF) so that the instrument can be interfaced with the Optical Stimulus (OS). The latter provides stable NIST-referenced illuminating sources for the wide variety of tests outlined in the Science Calibration Plan. These include broad-band, wavelength-selectable and laser diode light sources, which can simulate point sources or provide uniform illumination. The OS configuration will be computer controlled (see WFC3 Optical Stimulus Functional Requirements).

During thermal/vacuum testing and calibration, WFC3 will be configured and controlled by the Science Instrument Test System (SITS). The SITS can be commanded in real-time from a command console, or run from stored command SMS files. Current intentions are that both WFC3 and OS operations will be automated, as far as is possible, with the OS configuration specified through uniquely-identified keywords in SMS routines. Implementing this approach requires that the individual measurements outlined in the SciCal Plan are converted first to written procedures, and then to SMS routines.

Procedures are currently being written by ICAL members for a number of calibration test measurements. The mechanism by which these procedures are converted to SMS routines is under evaluation.

2.3 Data handling

WFC3 produces Standard Header Packet (SHP) and Science Data Image (SDI) files. In addition, the OS will produce ascii files of parameters which describe its configuration (Optical Stimulus Configuration files – OSC files). Those telemetry data are interpreted and converted to keyword/value pairs that will be written into the header, along with the raw image data in a Flexible Image Transport System (FITS) format file.

We expect that several thousand 32 Mbyte images will be generated during the T/V science calibration campaign, including data from both WFC3 channels. Additional images will be acquired in pre-thermal/vacuum detector characterization, and in the optimization, optical alignment and functional testing that is required before and after major environmental tests. Thus, depending on the need for diagnostic measurements, problem resolution and stability verification, we may acquire as many as 20,000 images, with a total data volume approaching 1 Terabyte. In addition, the OS will provide OSC files at regular intervals, monitoring the stability of relevant parameters, such as flux level and stability, image location and stability and wavefront errors. Those data must be easily accessible to the calibration team to allow evaluation of OS performance and stability.

A subset of the data obtained from WFC3 during the thermal/vacuum calibration, including at least the Calibration Reference Files, will be placed in the STScI data archive (DADS) via the OSS-PODPS Unified System (OPUS). OPUS will eventually interface with CALWFC3, the WFC3 data-analysis pipeline, to produce fully-calibrated images once WFC3 is fully operational in orbit. During thermal/vacuum testing, OPUS may be used to process incoming telemetry to produce FITS files. A memorandum of understanding is currently being developed between the STScI WFC3 and Archive teams to identify issues and define responsibilities related to possibly incorporating WFC3-specific features in OPUS.

2.4 The pre-flight data system

The data system used during the thermal/vacuum campaign must format, process and catalogue the pre-flight calibration files, and route the data to an appropriate archive. These activities fall under three distinct categories:

1. Data acquisition and processing, including merging information from SHP, SDI and OSC files
2. Data analysis, providing timely assessment of instrumental performance
3. Data archiving, allowing appropriate access to relevant datafiles

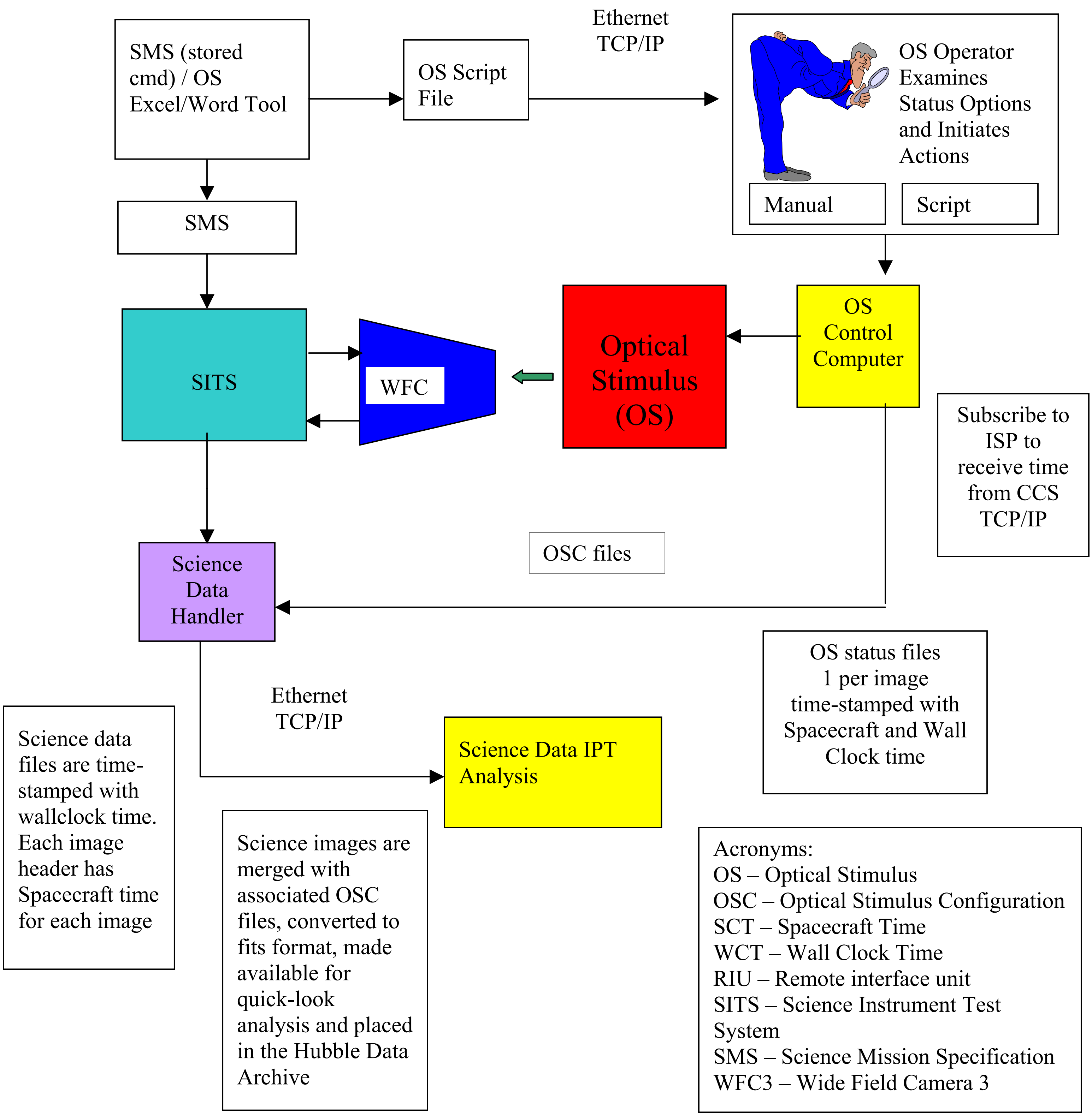

Figure 1: Overall dataflow within the system planned for data handling during the WFC3 thermal/vacuum test and calibration campaign

Figure 1 outlines the overall dataflow that is planned to support the WFC3 calibration program. The data system can be based on the OPUS system and the STScI DADS archive, used for in-flight HST systems; alternatively, heritage software from previous HST instrument calibration campaigns, in particular that developed by ACS, may be used. It has not yet been established which approach is more cost-effective and efficient. By considering the system functional requirements, we aim to lay the basis for making this decision.

3 Functional requirement

3.1 Description of required functionality

Our goal is to acquire and process WFC3 pre-flight calibration data with full automation, under the supervision of the Calibrating Scientist and an analysis team. This methodology permits the calibration program to run efficiently, while conserving personnel resources, schedule and testing costs. The architecture of the data handling system may be patterned after legacy software from the ACS and/or COS programs, and/or may use selected elements from existing STScI processing capability. We briefly describe each of these possible approaches, and suggest trade studies that would result in the definition of the optimal system. In either case, the requirements for the data handling and processing system must be established clearly.

3.2 Calibration tasks

The software architecture adopted for ground calibration should perform the following operations:

1. **Data acquisition.** Acquire raw image (SDI format) and SHP data from SITS, acquire snapshot OSC files from the OS control computer, and merge and convert those data into a single FITS file.
2. **Data logging.** Automatically log all calibration activities. Provide a capability for calibration scientists to record all relevant notes on calibration setup, anomalies and progress.
3. **OS monitoring.** Store all OSC configuration and telemetry data not associated with individual exposures in the form of ancillary OSC files. These are expected to be similar to snapshot OSC files, but may be taken at regular intervals or on command.
4. **Data storage and real-time access.** Store and provide a high-level view of all of the above items (1-3) throughout the T/V campaign. Images taken during T/V shall be available immediately for display on-site. For comparison purposes during T/V, it shall also be possible for calibration scientists at GSFC to access any additional datafile through the query facility within 3 minutes. All datafiles shall be assigned unique identifiers so that they are not overwritten when new data are generated.
5. **Data archiving.** Transfer all of the above data to STScI for users at STScI, permanent storage and archive retrieval. The archived data shall be available within 2 days (TBD) of generation and shall remain accessible throughout the

operational life of WFC3. This requirement does not imply that the Hubble Data Archive (HDA) must be used for all the data (see(7) below), only that a record be preserved.

6. **Exposure traceability.** Provide a capability for tracing FITS files back to the SMS scripts used to generate the WFC3 commands that produced the data. Each SMS script shall have a unique programmatic designation. Programmatic information from SMS scripts shall be included in the FITS keywords of all images generated using the scripts. Both during the T/V campaign and afterwards, users shall have the ability to query a database to identify the SMS script corresponding to a FITS file, and also to identify the set of images that correspond to a given SMS script or collection of scripts.
7. **Archive retrieval.** Most thermal/vacuum images shall be stored in the HDA for efficient access and retrieval. Images that are poorly suited for HDA storage may be excluded from that requirement. The data management system shall permit efficient data retrieval via special FITS keywords designated by the science IPT.
8. **Automation.** The data management system shall provide a sufficient level of automation so that calibration scientists are not obliged to divert a significant amount of time during T/V to manual data management tasks (e.g. file transfer).
9. **Reuse.** To minimize software development costs, existing legacy software shall be used in the data management system wherever possible.
10. **Compatibility.** For efficiency, the software/hardware system selected should be capable of processing images from both the UVIS and the IR channels.

Expanding on these issues,

1. Data describing the OS configuration during a particular exposure must be associated with the relevant image frame for each measurement. Snapshot files capture the configuration immediately before and/or after each WFC3 calibration image is taken. Most of the relevant information can be incorporated automatically in the FITS header of each calibration image, using a unique set of keywords defined by the ICAL. This requires merging OS data with the SITS SDI raw image files, and modifying the science header to accept the additional keywords.
2. Calibration tests should be logged electronically in an automated fashion. This is essential to allow the Calibration Scientist to document changes or anomalies in test activities, to assess progress in the calibration campaign, and to allow rapid, efficient data access.
3. OSC files can be acquired during an exposure to provide continuous monitoring of wavefront quality, positional stability, temperature and flux stability. Those data should be archived in a short-term database and time-tagged using a method that allows ready association with contemporary WFC3 calibration images
4. All data files obtained during the thermal/vacuum must be accessible for immediate inspection after image readout is complete for verification of exposure levels and quick-look analysis. The system should allow comparison on timescales of a few minutes with any relevant previous-taken images, from both T/V and component analyses, and with associated OSC files. This requires the adoption of a naming convention, permitting unique identification of each file.

Suitable options are currently being explored. This is essential to allow rapid identification and resolution of potential anomalies, and hence efficient use of thermal/vacuum calibration time. The STScI DADS database can support those activities to some extent. However, data archived in DADS are not usually available until 24-48 hours after deposition. Moreover, it is not clear whether all non-science calibration imaging (e.g. images taken for the purpose of optical alignment, OSC files) are suitable for DADS archiving. It may therefore be necessary to maintain a second, short-term database to provide rapid access to all calibration files for real-time inspection. All FITS-formatted images deemed scientifically acceptable by the Calibration Scientist can then be transferred to the DADS at STScI, as well as preserved in the short-term database.

5. Given the data volume, the calibration files must be archived in databases that provide both a permanent record and rapid, flexible access for inspection and analysis. If a second database, in addition to the HDA, is required for rapid data access, that database would *not* be maintained as a long-term archive; the HDA will provide that function.
6. Linking the calibration files to the generating SMS scripts will simplify the identification of relevant images for a particular test, both during the T/V campaign and throughout the operational lifetime of WFC3.
7. The calibration files will be analysed, both on-site and off-site by the calibration Team and by interested SOC members, to establish that WFC3 performance meets the CEI specifications. It is anticipated that most SOC members will use the DADS to inspect and analyse calibration data. This activity requires suitable software packages at both locations.
8. Process automation is critical for efficiency and maintaining tight schedules with limited resources. All steps of the data acquisition and handling process should be integrated seamlessly with legacy software for smooth and un-interrupted flow of calibration activities. Intervention by the Calibration Scientist for real-time commanding will be necessary when anomalies occur and for measurements that are deemed necessary in his or her judgment. The calibration team will be stationed to provide the necessary quick-look review of image files for science quality and verification of instrument and OS configurations, exposure levels and adherence to the calibration procedures. Additional team members will be available for more detailed analysis of the data, providing feedback on data quality to the Calibration Scientist. Real time communications between all team members will be vital to assure efficient work-flow and completion of all calibration tasks in a timely manner.
9. The WFC3 team should avoid unnecessary duplication of effort.
10. While the WFC3 UVIS camera will produce frames in ACCUM format, the IR camera data are in MULTIACCUM format – a series of up to 15 sequential non-destructive reads in addition to the initial zero-level read. UVIS and IR images will be acquired closely spaced in time; processing those data using separate reduction packages would involve unnecessary duplication of effort, and result in loss of efficiency.

4 Data Flow Architecture

We describe two possible approaches to designing the pre-flight data-handling system used in the WFC3 thermal/vacuum campaign. Figure 2 provides a modular summary of one approach to structuring the dataflow during thermal/vacuum calibration and test. The second option, based on the COS ground calibration plan, is outlined in section 4.3 and Figure 3.

The hardware and analysis software for implementation of the structure illustrated in Figure 2 are discussed in sections 5 and 6. Summarising the architecture, the data output from WFC3 via SITS is combined with OS parameters in Merge. At this juncture, both Merge and Fitscon should be regarded as conceptual stages of analysis, rather than necessarily as separate software programs. The output files from Merge follow two possible paths. One path leads to OPUS and subsequent archiving in the DADS at STScI. OPUS ingests the data, producing FITS files that are transferred to DADS and are available for subsequent analysis. The PMDB module represents the Proposal Management Database, accessed by OPUS to populate a subset of the FITS header. Thermal/vacuum calibration tests will not be commanded from proposals, but a similar database will be required to feed OPUS. It is possible that OS parameters may have to be accessed via the PMDB.

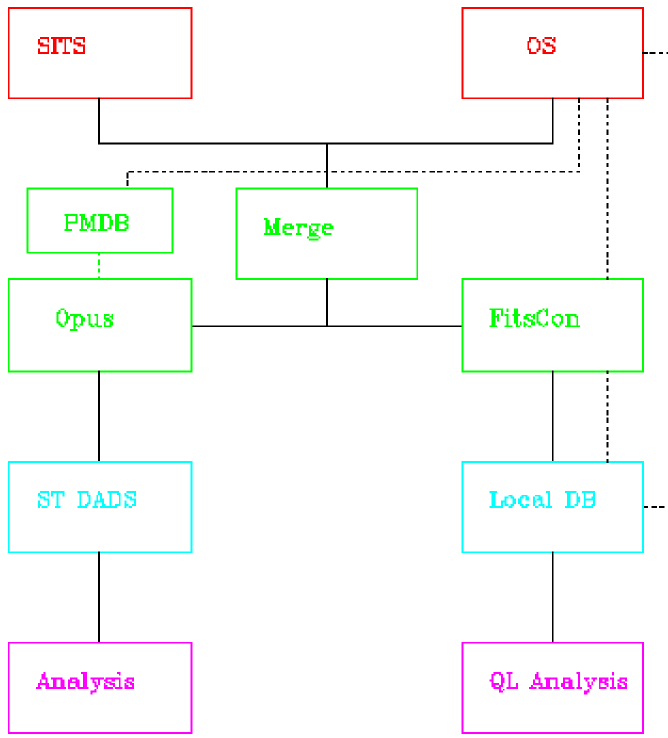

Figure2: WFC3/OS dataflow for the thermal/vacuum campaign – model 1

A second path, based on STIS/ACS experience, leads to archiving in a local database in which additional OSC engineering and monitoring files are maintained, with subsequent transfer of image data to an institutional database. In this scenario, the combined image data files are fed to Fitscon, which merges and converts those data to FITS format. The output files from this stage are stored in a local database (Local DB in Figure 2), where they are available for real-time quick-look analysis. Additional monitoring data from the OS are also fed into the local database, either as ascii files, or, via Fitscon, as FITS files. Deciding the appropriate architecture for T/V is clearly of high priority.

4.1 Architecture based on using OPUS

This architecture follows the path shown on the left side of Figure 2. Several variations can be envisaged:

- Data from WFC3 and the OS can be combined directly in a modified version of OPUS i.e., rather than a standalone piece of software, Merge is part of the OPUS format conversion. A Merge program would still be required to feed the local database via Fitscon.

- Fitscon and Merge could be combined, with the output FITS files directed to both DADS and the local database (i.e. no requirement for OPUS on-site)
- An on-site version of OPUS could be adapted to combine OSC files and SITS data, producing FITS files which feed both the DADS and the local database.

We do not consider it feasible to rely on the DADS as the sole database for calibration activities because of inherent time delays for data retrieval. The STScI archive will provide access for SOC members with casual interest in the calibration results. It will also provide reliable long-term storage. However, there are significant limitations in its performance in at least four areas:

1. Real-time access: turnaround time of at least 24, and probably 48, hours between data in and data available.
2. Limited storage ability: there is no mechanism at present for archiving and retrieving OSC monitoring files.
3. Limited data access: Data access and retrieval only through specific keywords, with consequent limited ability to associate relevant data files.
4. User traffic: Competition with in-flight STScI instruments and other STScI users.

The DADS was not designed for use as a real-time calibration database.

4.2 Architecture based on ACS heritage

ACS IDL calibration software has been developed to acquire, catalogue and analyse pre-flight instrument test data in order to meet the primary goals of the calibration team. This software follows the path shown on the right hand side of Figure 2.

The software has been used successfully to acquire raw data, usually SHP and SDI files taken directly from the instrument. The telemetry data are interpreted and converted into keyword/value pairs that are written, along with the image data, to FITS format files.

A preselected subset of the available telemetry items were catalogued in an IDL database. In many cases these items correspond to the Space Telescope Science Header definition. However, the ACS database and headers predate the ST definition, so that not all keywords and values correspond exactly. A copy of the catalogue, with on-line search capacity, is available at the ACS website (http://adcam.pha.jhu.edu/).
The software for acquisition and cataloguing was based originally on similar software developed for the STIS HST program. However, the ACS calibration team made great improvements in code stability and process automation, such that by mid-1999 the software was automated completely. This includes seamless integration with the legacy ACSVU engineering (GSE) archive.

This process automation allows the calibration team to analyse data rapidly, providing the necessary feedback to engineering groups involved in the design of the instrument, in addition to information used to verify instrumental performance. Like the acquisition and cataloguing software, some specific analysis routines have been developed that can proceed without user intervention. For example, the DINO (Detector Induced NOise analysis) program was run for many weeks during the ACS/NCS compatibility tests during the summer and fall of 2000.

The ability to process ACS data automatically has allowed the calibration team to operate more speedily and consistently while conserving personnel resources. Much of the software, written in IDL, is available as a library of routines (summarised in Appendix I).

The ACS approach has already addressed the issues discussed in section 3. Some modifications and extensions will be required, to deal with WFC3-IR Multiaccum frames, for example. However, the software is available for adoption by WFC3.

4.3 The data system for the COS thermal/vacuum campaign

The ground calibration plan for the Cosmic Origins Spectrograph (COS) is outlined as part of the COS AV-03 document (COS-01-003). COS will undergo thermal/vacuum testing in July, 2002, either at Ball Aerospace or at GSFC. Besides using inboard lamps for calibration, the instrument will be integrated with a modified version of the RAS/Cal external stimulus used in ACS calibration. Figure 3 illustrates the data system architecture adopted for the COS campaign.

COS will use an on-site version of OPUS, running on a Sun workstation. Only the OPUS code for the generic conversion routines will be run to generate FITS files from the SDI and SHP files produced by the COS SITS. As with WFC3, the headers in COS calibration files must include keywords which identify the configuration of both the instrument and the external stimulus. The Sybase database supports translation of the instrumental parameter values to keyword values; however, RAS/Cal is not computer controlled, and information on the stimulus configuration will be input manually by the calibration team, using software provided by the COS IDT.

After conversion to FITS format, the COS calibration images are transferred to the STScI DADS, which will serve as the sole, long-term archive for COS calibration data. The DADS will not be used, however, for the essential quick-look and first-depth analysis undertaken *during* the thermal/vacuum calibration campaign. COS calibration files will also be accessed and archived in a local database set up by the COS IDT, which may also hold pre-thermal/vacuum calibration files. Analysis of these data will be carried out using CEDAR, a suite of calibration routines developed by the COS IDT from STIS heritage software. The local database will not be maintained as a duplicate, long-term archive.

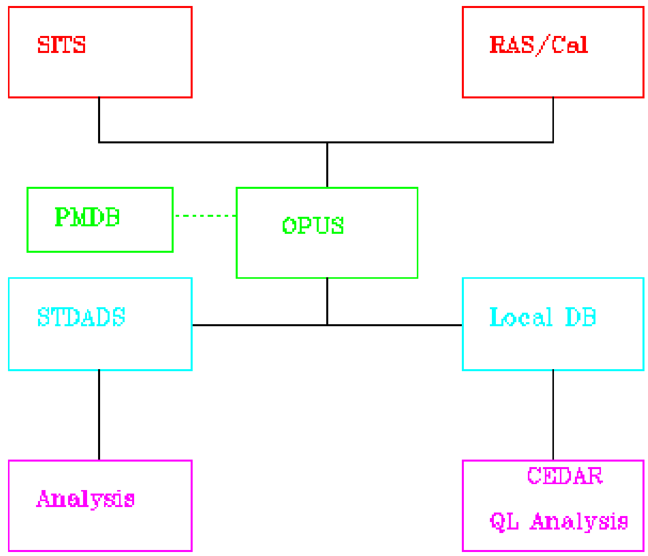

Figure 3: Dataflow structure for the COS thermal/vacuum campaign

The principal justification for the use of a separate local database during COS thermal/vacuum testing is the lack of real-time access to data archived in the DADS. The system architecture adopted by COS is conceptually almost identical with the ACS-based system described in section 4.2.

5 Data Analysis Packages

The software used to analyse WFC3 calibration data must be installed as a robust, flexible package, capable of applying individual analysis routines and displaying the results for inspection. The results from calibration studies should be logged individually and archived as reference data for subsequent studies. Standard software tools developed for this purpose include extensive libraries of IDL routines, and the STSDAS/IRAF collection of data-handling, processing and analysis tasks.

5.1 IDL and IRAF routines from the STScI DCL group

The STScI DCL group has compiled a number of standalone image analysis routines to determine detector characteristics and perform basic operations. These include:

PRAXIS: IDL code for the following procedures,

- Bias subtraction
- Mean frame calculation (pixel statistics)

- Read noise calculation
- Hot pixel identification
- Event selection, for CTE and gain measurement (see TARA)

TARA (Tools for ACIS Real-time Analysis)

- Event browser – a routine for analysing X-ray images, determining gain, energy resolution and CTE from 55Fe exposures

IDL code

- Eper.pro, fpr.pro – these routines, originally written by Mike Jones for ACS, determine CTE from flatfield images, and calculate the signal level, bias level used for bias correction, and readnoise in the prescan region

IRAF tasks

- Split_ccd, eper_cc, fixint – these codes are designed to deal with DCL unsigned integer format, and create FITS files with the appropriate header values so that the images can be read and displayed by SAOIMAGE and IDL.
- Split_ccd also separate images produced by detectors with dual amplifiers, to give 2 separate files, with the amplifier located in the lower-left corner. These data can be accessed by Praxis routines.

5.2 IDL routines used by the GSFC DCL

The GSFC DCL has also compiled test procedures and routines for characterizing detector properties. These are described in DCL document numbers DCL-WFC3-002 (UVIS detector characterisation) and DCL-WFC3-003 (IR detector characterisation) . They are capable of performing the following measurements

- Absolute quantum efficiency
- Readout noise and bias stability
- Dark current
- Full well depth, linearity and dynamic range
- Charge transfer efficiency – including degradation with radiation
- Flat field uniformity, stability and cosmetics
- Overexposure and ghost images
- Transient response to radiation
- Sub-pixel response
- Out of band quantum efficiency (IR only)
- Exposure time control (IR only
- Detector temperature sensor calibration (IR only)

5.3 ACS IDL software package

The ACS team has compiled an extensive package of software routines for use during their thermal/vacuum and ambient-temperature calibration campaigns. Those routines are based to a large extent on routines used to calibrate previous HST instruments, notably STIS and WFPC2. In the spirit of re-use of STScI heritage, these programs can be adapted for use by WFC3. The ACS software falls under four major headings:

1. Data acquisition, preflight cataloguing, and database archiving tools
2. Data analysis, image manipulation and reduction tools
3. External stimulus related tools
4. General plotting and fitting routines, spreadsheet manipulation and graphical window tools.

Since all of these routines have been applied to ACS data, they are already configured to deal with the relevant formats, and extensive work has gone into providing the necessary user interfaces. Furthermore, the entire process has been integrated with automated SMS procedures, including the development of a software tool (SMSTOOL) that can be used to generate SMS procedures automatically. The latter can be adapted for use with WFC3 and the OS (see section 2.2). The full suite of routines is outlined in Appendix 1.

5.4 NICMOS analysis software

The WFC3-IR camera will use non-destructive readout to produce images in Multiaccum mode. The data format, both produced by the SITS and after transformation to FITS, is designed to match NICMOS images. A number of software tools have been constructed specifically for analysis of NICMOS data, and these should also be suitable for used with WFC3-IR data. In particular, the document ISR NICMOS97-30 describes a number of IRAF tasks designed for mathematical, statistical and image structure analysis, together with calibration-specific routines. These routines are part of the STSDAS/IRAF package, and are identified in Appendix 2.

5.5 STSDAS/IRAF analysis software

The Space Telescope Science Data Analysis System (STSDAS) consists of software for calibrating and analysing data from the Hubble Space Telescope. STSDAS includes the same calibration routines as are used in the routine data processing pipeline, as well as general-purpose tools and enhancements to the Image Reduction and Analysis Facility (IRAF).

IRAF consists of a wide range of image processing tools that may be useful for display and analysis of calibration images. IRAF is a command language that has been used to develop a large number of image processing, reduction and analysis tasks that can be employed for converting data into scientifically useful information for the ICAL calibration scientists, the Science IPT and SOC members.

IRAF provides the user interface (command language, or CL) and general purpose graphic and image display facilities. The CL provides a wide range of functions typically provided by an operating system. Some of the CL features include input/output redirection and piping, command buffers and history editing, minimum matching of commands, host-independent file naming, parameter checking , background and batch-processing modes, an integrated online help system, and a script authoring environment.

Because STSDAS is layered on IRAF, the system runs on any computing platform for which an IRAF port is available. Appendix 2 lists potentially useful routines currently available in STSDAS.

6 Hardware requirements

It is essential that sufficient resources will be provided to permit timely analysis of WFC3/OS calibration data. This is vital for interpretation both of science-related imaging and of data taken for technical purposes, such as detector characterisation and optical alignment. Any unnecessary delays in identifying and correcting problems will result in an immediate impact on the efficiency of the thermal-vacuum program, with potential cost implications.

Our suggested model for the WFC3 T/V hardware system is based on proven results from ACS testing at Ball Aerospace. The data system employs two PCs and four workstations at the testing location; these provide system control and quick-look/first-depth analysis. Quick-look analysis implies simple inspection as data are taken, checking for gross problems; first-depth indicates a first cut at in-depth analysis of a subset of the frames (comparison with previous datasets, for example). Between three and four additional workstations were available at a remote location (2nd floor office at Ball) for more detailed analysis. Following that model, a possible baseline for the WFC3 thermal/vacuum campaign is as follows:

1. PC to control OS and provide OS data output (in hand as part of OS system)
2. PC to merge SITS and OS data to provide FITS images
3. Workstation 1 to log automatically all WFC3 image information
4. Workstation 2 (at test location) for quick look analysis (Scientist 1)
5. Workstation 3 (at test location) for first-depth analysis (Scientist 2)
6. Workstation 4 (at test location) for first-depth analysis (Scientist 3)
7. Sufficient on-site memory and disk-storage to cope with 0.5 to 1 Terabyte of calibration data

We anticipate that at least 3 members of the ICAL/IPT team will be on-site to ensure efficient use of thermal/vacuum time. Since thermal/vacuum testing is an around-the-clock operation, a significant team of support scientists will be required to support the total calibration effort. In addition to the above hardware, located on-site at GSFC, we anticipate use of 3-4 workstations at a remote location, either at GSFC or at STScI. Each workstation will require access to the GSFC local database to allow full analysis flexibility.

7 Summary

The WFC3 T/V calibration activities require a software architecture that can

- Log calibration activities
- Acquire SITS calibration images
- Combine SITS and OS datafiles
- Produce FITS images for archiving
- Archive OS monitoring data
- Allow real-time access to calibration data and rapid access to 20,000 images in a data archive

We recommend that a trade study, to consider the technical, schedule and cost merits of the two concepts described above, should be undertaken by the ICAL group. This should consider the following issues:

1. What system should be used to produce FITS files?
2. What is the optimum data storage strategy?
3. What software should be used for quick-look analysis?
4. What software should be used for off-site T/V data analysis?
5. What overall architecture should be adopted?

We also recommend that the desired system concept, identified by the trade study, should be adopted as a prototype, assembled, and made available during the verification and testing of the WFC3 Optical Stimulus. The current schedule places OS testing in mid-2002. A decision should be made on the architecture of the data-handling system by 30 September 2001 to permit the development of a prototype system to meet that schedule. Additional computer hardware needed for software development, testing and verification, at both STScI and at GSFC, remains to be discussed in detail.

APPENDIX 1: ACS CALIBRATION SOFTWARE

This appendix lists the routines constructed for the ACS calibration campaign. All are IDL routines. Unless noted, the author is Jon McCann (JHU); [L*] indicates co-authored with Don Lindler.; [L] written by Lindler; [H/McC] written by George Hartig and John McCann.
All of these routines can be downloaded from
http://adcam.pha.jhu.edu/instrument/calibration/software/idl/
Some modifications will be required to deal with WFC3 data; some routines may not be useful for WFC3 analysis.
RAS/HOMS is the Refractive Aberrated Simulator/Hubble Opto-Mechanical Simulator, the ACS equivalent of the optical stimulus.

Acquisition software

ACS_ACQ: ACS data acquisition tool
ACS_ACQUIRE: Routine to acquire raw ACS data (packet files) [L*]
ACS_ACQUIRE_COMMON: includes file to define ACS_ACQUIRE common block [L*]
ACS_ACQUIRE_EDS: populates engineering data snapshot headers [L*]
ACS_ACQUIRE_HEADER: creates science data header [L*]
ACS_ACQUIRE_LOG: adds an observation to acs_log catalogue [L*]
ACS_ACQUIRE_RASCAL: acquires RASCAL header information [L*]
ACS_ACQUIRE_SCI: reads science data packets and creates data array [L*]
ACS_ACQUIRE_SETUP: sets up common block for ACS_ACQUIRE [L*]
ACS_ACQUIRE_SHP: gets detector temperatures from standard header packets [L*]
ACS_ACQUIRE_STUFF: acquires STUFF header information [L*]
ACS_ACQUIRE_UDL: processes internal UDL packet [L*]
ACS_ACQUIRE_UNCOMPRESS: uncompresses WFC data [L*]
ACS_DATA: selective acquisition of IDL data [L*]
ACS_FIX_SMS_TIMES: fix observation time, date in database entry that used SMS time rather than current time [L*]
ACS_READ: read ACS raw data files [L*]
ACS_READIMAGE: reads iraf, fits, MAMA, CC200, SDAS or RAW images files [Beck]
ACS_TIME: converts ACS time to UTC [L*]
ACS_UNSCRAMBLE: changes orientation of frame [L*]
ACSVU_ARCHIVE_FILE: logs science image to ACSVU database
LOG_ACSVU: logs science image to ACSVU database server via RPC

ACS Logging/database software

ACS_FIND: widget tool to search preflight database [L*]
ACS_GET_FILE: retrieve datafile associated with ACS_LOG entry number

ACS_TEXT_LOG: retrieves information from database and formats as text file
ACS_ARCDISK: adds archive disk item
EVAL_DOC: print image evaluation document
LATEST: notify user when database has been updated
LATEST_ENTRY: returns latest data entry number
RASHOMS_LOG: print out commonly used information for a list of database entries

Analysis software

ACS_CMFSCAN: reduces focus scan data obtained by moving corrector focus [Hartig/McCann]
ACS_CMFSCAN_MP: reduces multi-field focus scan data [H/McC]
ACS_DETALIGN_HRC: computes detector alignment from HRC focus sweep [H/McC]
ACS_DETALIGN_WFC: computes detector alignment from WFC focus sweep[H/McC]
ACS_DUST_SEARCH: produces ratio images for analysis of filter dust search
ACS_LEDGE: leading edge ramp plots
ACS_STATS: CCD S/N levels for darks and biases [L*]
ACSGAIN: calculate detector gain
DINO: detector noise analysis
DINO_LOAD: load new image into active DINO
FITGAIN: fit line to given mean and variance data
LEDGE: leading edge ramp plots
RENO: detector read noise monitor
STAMMRE: stability and repeatability testing
STATBOX: mean and variance statistics in boxes over pairs of images

Archiving

ACS_GET_FILE: retrieve data for particular log entry
ACS_TBA: shows amount of data left to be archived
PARCH: gather data for preflight archive cdrom

Image manipulation software

ACS_CR: combines ACS images with cosmic ray removal [L]
ACS_EE: calculates encircled energy within specified aperture [H/McC]
ACS_EER: computes encircled energy vs radius from specified center [H/McC]
ACS_IMAGE_STATS: statistics on image and overscan region
ACS_OVERSCAN: subtract overscan bias level and remove overscan pixels
ANAMORPH: corrects for anamorphic projects [L]
AZIMUTHAL_AVG: computes azimuthal average

Image viewers

ACSLOOK: quick look image tool [Beck]

BIGTV: routine for viewing big images [L]
MVIEW: graphical image viewer with analysis tools
MVIEW_LOAD: load new image into active MVIEW

Plot software

ACS_IMAGE_PS: generate postscript file for specific image
ACS_PS: generate postscript image file of specified dataset
PPLOT: interactive plotting tool
RASHOMS_PS: ACS_PS driver routine for RASHOMS spots and darks

IDL windows

WIN_DRAW_BOX: draw a box on the graphical display
WIN_HILIGHT_REGION: highlight a region
WIN_TRACE: draw a crosshair on the graphical device
WIN_VALID: determine if a graphical window is open

General purpose GUI tools

X_DISPLAY_LIST: create a widget to display a list
X_DRAW_BOX: draw a box on a graphical display
X_GET_INPUT: create a widget to get input from user
X_SELECT_BOX: allow user to interactively select a box
X_SELECT_LIST: select item from list
X_TRACE:

Mathematical procedures

Gaussian fitting
Plane fitting
Cross-correlation

Additional procedures are available for manipulating spreadsheets, SMS-related commands, audio features and miscellaneous other issues.

APPENDIX 2: IRAF/STSDAS ROUTINES

The IRAF STSDAS routines are grouped under eight categories: analysis, contrib., fitsio, graphics, hst_calib, playpen, sobsolete and toolbox.
This appendix lists the subset of the available routines which are likely to be most relevant to the acquisition and analysis of thermal/vacuum calibration data.

Analysis
- Fitting: curve fitting tools, including
  - Function – generate functions as images, tables or lists
  - Gfit1d – interactive 1-D linear curve fit to images, tables or lists
  - I2gaussfit – interactive fitting to noisy images (script)
  - Nfit1d – interactive 1-D non-linear curve fit
  - Ngaussfit – interactive 1-D multiple Gaussian fitting
  - N2gaussfit – 2-D Gaussian fit to images
- Fourier: fourier analysis routines
  - Autocorr – autocorrelate an image
  - Carith – multiply or divide complex images
  - Crosscor – cross-correlate two images
  - Factor – display prime factors of a number
  - Fconvolve – convolve two images
  - Forward – compute forward fourier transform of an image
  - Frompolar – convert amplitude and phase to real and imaginary parts
  - Inverse – compute inverse fourier transform of an image
  - Listprimes – display first N prime numbers
  - Powerspec – compute the power spectrum of an image
  - Shift – shift an image by a specified number of pixels
  - Taperedge – apply a cosine bell or linear taper to image edges
  - Topolar – convert real and imaginary parts to amplitude and phase

Contrib
- Acoadd – image addition using Lucy method
- Gaussfit – least squares and robust estimation program

Fitsio
- Catfits – produce a catalog of fits files on tape or disk
- Fitscopy – copy a fits file
- Strfits – convert a fits file to tables or to a GEIS files

Graphics
- Sdisplay – image display package for SAOImage display devices
- Stplot – general plotting utilities

Hst_calib
- Ctools: general calibration tools, including
  - Chcalpar – modify calibration keyword value
  - Eng2tab – extract information from SHP, UDL files
  - Getcal – populate a pset with calibration with calibration parameter values

        - Msbadpix – detects bad pixels in STIS and NICMOS data
        - Msreadnoise – measures read noise in STIS and NICMOS images
        - Pprofile – plot or print radial profile or encircled energy
        - Putcal – place parameter values from a calibration pset in a data file
        - Rcombine – average or sum rapid readout data with propagation of errors
    - Nicmos: tasks for calibrating NICMOS data, including
        - Biaseq – equalize bias levels in Multiaccum readouts
        - Ndark – build NISMOS DARKFILE calibration reference file
        - Pstack – plot a stack of pixels from a Multiaccum image
        - Pstats – plot a stack of Multiaccum statistics
        - Sampcum – accumulate a Multiaccum image from a set of sample differences
        - Sampdiff – compute first differences of samples in a Multiaccum image
        - Sampinfo – print sample information from a Multiaccum image
    - Wfpc: tasks for calibrating WFPC and WFPC2 data, including
        - Bjdetect – detect bias jump
        - Checkwfpc – perform consistency checks on WF/PC reference file
        - Combine – combine images using various algorithms and pixel rejection
        - Crrej – combine images to make a cosmic-ray free image
        - Engextr – extract information from WF/PC engineering data
        - Mkdark – create a dark by combining images to get rid of cosmic rays
        - Noisemodel – determine noise model parameters from a CCD frame
        - W_calib – tasks for deriving the WFPC instrument calibration,
            - Flagflat – flag flat-field pixels with anomalously high or low values
            - Mka2d – make or correct a-to-d correction reference file
            - Sharp – measure sharpness of star(point-source) images
        - Wstatistics – compute and print WFPC image statistics
- Playpen (experimental tasks)
    - Bwfilter – 1-D Fourier filtering
    - Edge – 2-D image edge massaging
    - Hsubtract – Baade-Lucy background subtraction algorithm
    - Immean – compute mean of image and save as task parameter
- Toolbox
    - Headers: tools for modifying image headers
        - Eheader – interactive edit an image header
        - Groupmod – add/delete group parameters from a multigroup dataset
        - Hcheck – check image header keywords
        - Hdiff – display the difference between two image headers
        - Upreffile – update calibration reference file names in image headers
    - Imgtools: tools for manipulating and examining images
        - Addmasks – combine several masks or bad pixel lists
        - Gcombine – combine a set of GEIS images into one image
        - Gcopy – generic multi-group copy utility
        - Gstatistics – compute and print image pixel statistics for all groups
        - Imcalc – perform generic arithmetic operations on an image
        - Mstools – tasks to handle STIS/NICMOS IMSETS
        - Pixedit – screen editor for image pixels

Rbinary – create an image from a binary file
Stack – stack images to form a new image with one more dimension
Xyztoim – interpolate table values, writing results to an image

Tools: generic data handling and utility tools

Ttools: table manipulation tools

APPENDIX 3: ACRONYM LIST

ACS: Advanced Camera for Surveys
ACIS: Advanced Camera Imaging System (Chandra)
ACSVU: Database for archiving ACS engineering data
CALWFC3: WFC3 calibration pipeline
CEDAR: Data analysis system for COS
CEI: Contract End Item
COS: Cosmic Origins Spectrograph
DADS: Data Archiving and Distribution System
DCL: Detector Characterisation Laboratory
DINO: Detector Induced Noise analysis
FITS: Flexible Image Transport System
GEIS: Generic Edited Information Set (HST image file format)
GSE: Ground Support Equipment
GSFC: Goddard Space Flight Center
HDA: Hubble Data Archive
HST: Hubble Space Telescope
ICAL: Image Calibration Group
IDL: Interactive Data Language
IDT: Instrument Development Team
IPT: Integrated Product Team
IRAF: Image Reduction and Analysis Facility
NCS: NICMOS Cooling Shroud
NICMOS: Near-Infrared Camera and Multi-Object Spectrograph
NIST: National Institute of Standards and Technology
OPUS: OSS-PODPS Unified System
OS: Optical Stimulus
OSC: Optical Stimulus Configuration
OSS: Office of Space Sciences
PMDB: Proposal Management Database (OPUS)
PODPS: Post-Operation Data Processing System
RAS/Cal: Reflected Aberrated Simulations/Calibration
RIAF: Radial Instrument Alignment Fixture
SAOImage: Image display device
SDI: Science Data Image
SHP: Standard Header Packet
SITS: Science Instrument Test System
SMS: Science Mission Specifications
SOC: Science Oversight Committee
STScI: Space Telescope Science Institute
STSDAS: Space Telescope Science Data Analysis System
TARA: Tools for ACIS Real-time Analysis
UDL: Unique Data Log (file type)

WFC3: Wide-Field Camera 3
WFPC: Wide-Field Planetary Camera
WFPC2: Wide-Field Planetary Camera 2

Appendix 4: WFC3 Thermal-Vac Data-handing system:

Timestamp procedure requirements and interface with OPUS

version 1.7 2 May, 2002

Thermal-vacuum calibration of the Wide-Field Camera 3 instrument is currently scheduled for February-April, 2003. The data-handling system must collect files from a variety of sources, collate associated files, rename them in an appropriate manner and place them on staging disks for transfer to the analysis and archiving modules. This process may involve a number of separate software routines/programs. The procedure must meet the following requirements:

- The Science Data Handler in the SITS Flight Simulation machine packages science data into files with the following naming convention:

  CSI*nyydddhhmmss.ext*

  Where *n* is the Science Instrument letter (I for WFC3)

  yydddhhmmss is the ground receipt time (wallclock time) of the first science datapacket

  ext is the type of science data. ".sdi" and ".shp" files are required by OPUS

  These files are transferred by ftp to a specific destination.

  The timestamp procedure shall be triggered by the transfer of new image data to that destination.
- The timestamp procedure shall identify and correlate the .sdi and .shp files which refer to a particular image. The timestamp procedure shall transfer those files to two staging disks: disk 1 will be accessed from STScI by the OPUS system; disk 2 will be accessed by the local quick-look reduction system.
- The CASTLE Optical Stimulus will provide snapshot files characterising its configuration. These Optical Stimulus Configuration (OSC) files can be generated on command from either the SITS (via an RIU line) or from the OS operator. The OSC files are named according to the format

  CSI*iyydddhhmmss*.oss

  where *yydddhhmmss* is a wallclock timestamp.

  OS wallclock time and SITS wallclock time are taken from independent sources.
- The timestamp procedure shall identify the OSC file or files corresponding to a particular .sdi file.
- The timestamp procedure shall use the OSC file or files to generate a file including parameters which will be placed in the FITS header of the final image (as processed by OPUS and the quick-look system). This file is ascii format, 80 characters per line, with the following format:

KKKKKKKK= VVVVVVVVVVVVVVVVVVVV /CCCCCCCCC

where

columns 1-8 : keyword (capitalised)
column 9 : =
column 10 : blank
columns 11-30: value
column 31 : blank
column 32 : /
columns 33-80: comment

Character strings can extend beyond column 30, terminated by " /"; non-character values (integer, real, boolean) are usually right-justified. The FITS keywords for the OSC files are specified in appendix I. In some cases it may be necessary to combine parameters from several OSC files, for example, if a series of point sources are generated by taking separate exposures, moving the location of the fibre between exposures.

- The file with the OSC parameters will be named using the format

  CSI*iyydddmmss*.osc

  where the time is identical to the names of the corresponding .sdi and .shp files.

This file will be copied to both staging disks for access by the quick-look analysis and archiving systems.

- Additional files with comments on idiosyncracies of a particular image will be copied to the staging disk with the format CSI*iyydddmmss*.com
- The OPUS system shall combine the .sdi, .osc and (if present) .com files and generate a fits file for each exposure, named using the following prescription:

  IPPPSSOOT_CSI*iyydddmmss.ext*

  where I is the instrument (I for WFC3)
  P is the program ID
  S is the observation set ID
  O is the observation ID
  T is source of transmission or association product number (probably R=real time)
  i is the instrument (I for WFC3)
  ext is .fits in this case
  and the remaining parameters represent the date/time of the exposure and are identical to the timestamp on the input .sdi, .shp, .osc files. The IPPPSSOOT is recovered from the internal header of each image file.