1. SCOPE

This document specifies the interfaces between the *New Horizons* Science Operations Center (SOC) and the instrument pipeline, which process data from Level 1 to Level 2. The purpose is to define the various aspects of the interfaces in sufficient detail to establish a clear understanding between the SOC and the instrument team to allow for a parallel pipeline development.

2. SIGNATURES, AUTHORSHIP, AND RESPONSIBILITY

This document is unusual in that many parties took part in its writing. Specifically, sections 1 through 6 were written by the document author, whereas the bulk of the instrument-specific sections (7 through 14) were written by representatives of the instrument described. Because of this, a few words will be said regarding signatures.

Each instrument team will have a person or persons responsible for their section. If changes are made to that section, only the person(s) responsible need to sign the new revision. If, however, changes are made to sections 1 through 6, all parties need to sign. The title(s) of the person(s) responsible for each instrument section are given in the signature section above.

The *New* Horizons project maintains a copy of the signed version of this document under Configuration Management (CM); from 2008 through 2014, a fork ofthis document has been edited many times as a support document for PDS data set deliveries and the signatures no longer apply.

3. APPLICABLE DOCUMENTS

3.1 *New Horizons* Documents

- *New Horizons* SOC Data Pipeline Guide (SwRI Doc. No. 05310-SOCPL-G-01)
- *New Horizons* SOC Level 1 Data Formats (SwRI Doc. No. 05310-SOCL1DATA-01)
- *New Horizons* SOC Pipeline User Manual (SwRI Doc. No. 05310-SOCPLUM-01)
- *New Horizons* Data Management and Archiving Plan (SwRI Doc. No. 05310-DMAP-01)
- *New Horizons* Longevity Plan (APL Doc. No. 7399-9009)
- Definition of the Flexible Image Transport System (FITS) (ftp://legacy.gsfc.nasa.gov/fits_info/fits_office/fits_standard.pdf)

4. INTRODUCTION

The *New Horizons* SOC is part of the ground system that processes data returned from the *New Horizons* planetary spacecraft. Data downlinked from the spacecraft in raw packetized form is retrieved by the SOC from the Mission Operations Center (MOC) along with navigation and related ancillary data. The SOC generates the higher level (more refined) data products used by the instrument teams and science teams. In addition, the SOC performs archiving of data (but not data products, such as maps) to the Planetary Data System (PDS). The science data processing component of the SOC is called the SOC pipeline. The Level 2 pipeline (called the “instrument pipeline” and also known as the “calibration pipeline”) software is a segment of the SOC Pipeline. The instrument pipeline for each instrument is written by the appropriate instrument team. It is run on the SOC processing station to transform Level 1 decommutated data into Level 2 calibrated science data. The instrument pipeline creates PDS standard, Level 2 provisional products in Flexible Image Transport System (FITS) format, which is described in the document referenced herein (Definition of the Flexible Image Transport System).

The SOC pipeline is divided into two main parts: the Level 1 pipeline segment and the Level 2 (instrument) pipeline segment. Pipeline processing is carried out sequentially. Results of the Level 1 pipeline are provided as inputs to the instrument Level 2 pipeline segment. More information about the formats of Level 1 data can be found in the "*New Horizons* SOC Level 1 Data Formats" document (SwRI Doc. No. 05310-SOCL1DATA-01). The instrument pipeline generates Level 2 results that the SOC forwards to the PDS archiving process. Level 1 data is also archived to the PDS.

The SOC obtains science data from the Mission Operations Center (MOC) through automated network transfers. Low-speed housekeeping and high-speed science data from the spacecraft are provided in packetized Supplemented Telemetry Packet (STP) form. Navigation data, ephemerides and spacecraft pointing information, is provided in SPICE format from the MOC and is used in Level 1 processing. Upon getting packets (housekeeping and science data) from the MOC, the data is decommutated in the SOC and written to an SQL database. Housekeeping from the database and science data are associated by MET time and other methods, such as by using meta data inserted in the high-speed telemetry. Data for an observation are combined to create the Level 1 uncalibrated data file. A PDS detached header file (a separate file containing the PDS header with a pointer to the data file itself) is also created for each file. The header of the Level 1 file contains most of the necessary information about the particular observation needed by the instrument pipeline (an exception is detailed pointing, which will be calculated during calibration). The instrument pipeline segment creates the Level 2 calibrated data file from the contents of the Level 1 file and calibration data provided by the instrument team. Level 1 and Level 2 science data files are stored in FITS format. Lastly, the SOC archives pipeline data products to the PDS.

SOC pipeline processing is automated under the control of the Master Data Manager (MDM), which is a collection of scripts that control the flow of the pipeline. While manual execution of the program is permitted, normal operation of the SOC pipeline is not directed by manual requests or user inputs. Pipeline segments are called by the MDM when data from the MOC or from a previous pipeline step is available.

The hardware platform used for the SOC as implemented for launch and early mission is an Intel Pentium 4 processor running at 3.2GHz with 4GB RAM and a 146GB SCSI hard disk. In the case of the primary SOC (TSOC), located in Boulder, Colorado, two of these machines are used (one for pipeline processing and the other for data storage). At the secondary (backup) SOC (CSOC), only one machine is used for both purposes. The operating system used in all cases is Linux. Migration to newer machines compatible with SOC requirements is a possibility during the long flight missions.

5. INTERFACE DESCRIPTION

The SOC pipeline code will call the Level 2 pipeline code by executing a separate process. The name of the executable or script shall be:

`[instrument]_level2_pipeline`

(where "[instrument]" is the full instrument name (i.e. alice, leisa, lorri, mvic, pepssi, rex, sdc, or swap) in lower case).

The parameters (all are character strings) passed to the Level 2 code will follow the executable name and will be in the following order (note that "full path," when stated below, means a file specification containing the filename and all directories in the file's path):

- ***`in_file`*** **- Location of input (Level 1) file (in_file)**

The SOC will provide the full path of the Level 1 input file to the Level 2 pipeline code.

- ***in_pds_header*** **- Location of input (Level 1) detached PDS header**
  The SOC will provide the full path of the Level 1 PDS header to the Level 2 pipeline code.

- ***calibration_dir*** **- Location of calibration data and temporary file storage**
  Data provided by the instrument team that is needed for calibration shall be supplied by the instrument team. The SOC will provide the root directory containing these files (and potentially, subdirectories) to the Level 2 pipeline code so it references the correct location. The structure of the directories under this directory is up to the instrument team.

- ***temp_dir*** **– Location for temporary storage used by code**
  This is a place where the instrument pipeline code may write files for temporary use. The contents of this directory will be erased upon completion of the instrument pipeline.

- ***out_status*** **- Location of status file**
  The Level 2 pipeline, upon completion, may write a short machine readable status file used to communicate the results of the processing to the main SOC pipeline. The location (full path) of this file will be supplied by the SOC.

- ***out_file*** **- Location of output (Level 2) file**
  This is the file (full path) to which the output will be written. The SOC will provide this to the Level 2 pipeline code.

- ***out_pds_header*** **- Location of output (Level 2) detached PDS header**
  This is the file (full path) to which the Level 2 PDS header will be written. The SOC will provide this to the Level 2 pipeline code.

6. REQUIREMENTS

This section describes the various requirements that the instrument team shall follow with regard to the Level 2 pipelines.

6.1 Level 2 (output) Files

The Level 2 data files shall be FITS files, and there shall be exactly one Level 2 file produced for each Level 1 file (in any given run of the instrument pipeline). The headers will be mostly the same, except for optional additional keywords inserted in the Level 2 files (this could include, for example, refined pointing). In other words, typically, the Level 2 header keywords will be a superset of the Level 1 header keywords.

The filename of the Level 2 file will be supplied by the SOC, and it will be similar to the Level 1 filename. PDS requires a “27.3” ASCII character limit on the filenames, and the format of the names shall be as follows:

- Level 2: ***[instrument]_[MET]_[ApID]_sci_[version].fit***

(where "[instrument]" is the first three letters of the instrument name (i.e. ali, lei, lor, mvi, pep, rex, sdc, or swa) in lower case, "[MET]" is the 10-digit MET time, either from the instrument itself (low-speed data) or from the instrument interface card (high-speed data), "[ApID]" is the hexadecimal value of the ApID (Application ID of the main packet used in the data product; also abbreviated as APID in this document and elsewhere) for this observation, and "[version]" is the version stamped on this file based in existing previous versions produced).

The instrument/mode strings above will be derived from the ApID of the data, and these filenames will be supplied to the instrument pipeline (see the interface description above).

- Level 1: ***[instrument]_[MET]_[ApID]_eng_[version].fit***

Whereas the Level 1 files will be in the same "units" as the data coming from the spacecraft/instrument (i.e. same binary representation - this is partly to avoid any round-off or conversion loss), the Level 2 files shall express values in physical units useful for scientific interpretation.

Whereas the Level 1 files only contain the header and data itself, the Level 2 files shall contain (when appropriate) two additional "panes" (FITS extensions):

- Error (specifies error bars on the numbers – defined by the instrument team)
- Quality (Indicates the quality of the data – defined by the instrument team)

If it is desired to re-run the Level 2 pipeline (because of new Level 2 code or calibration data), a new version of the output file will be named as specified above when calling the Level 2 pipeline.

6.2 Pointing Information

The pointing information included in the Level 1 files will be mostly non-instrument specific (except for bore-sight vector where applicable). It also may not cover the time granularity needed for calibration in the Level 2 pipelines (see the "*New Horizons* SOC Level 1 Data Formats" document (SwRI Doc. No. 05310-SOCL1DATA-01) for specifics). Therefore it is expected that the Level 2 pipelines may have to make use of SPICE. It is therefore the responsibility of the Level 2 pipelines to provide this functionality. SPICE kernels will be available on the SOC.

6.3 The Code

The SOC defines the interface the code uses to access the required data. This includes the directory structure on the disk where the Level 1 data file can be found as well as the path (specific to each instrument) where instrument-team-supplied calibration files and other data will be stored and accessed. Also, the filename of the output file is supplied.

The code shall be written in a language that meets the SOC's longevity requirements (see section 6.11). More information on this can be found in the "*New Horizons* SOC Pipeline Guide" (SwRI Doc. No. 05310-SOCPL-G-01). The languages allowed are as follows:

- C
- C++
- Fortran
- IDL

- Python
- Java
- Perl

The code written by the instrument team shall not implement very time consuming iterative numerical processing to the extent that it has an impact on the daily processing done by the SOC. In other words, if the time to compute Level 2 data is so extreme that it jeopardizes the completion of each daily run of the pipeline (so ample idle time is not available between runs), the situation will need to be re-evaluated. It is expected that a daily run of the entire pipeline be complete within a few hours of its start. This gives most of the day for users to access new data before the next run is initiated. The maximum time allowed for execution of an instrument's Level 2 pipeline shall be **5 minutes** (for each input file processed). The predicted actual maximum time is negotiated and specified in the instrument-specific sections.

Any needed 3rd party libraries also shall meet the longevity requirements. Specifically, source code should be available and must be provided with the code unless already available to the SOC.

6.4 Calibration Data

The code will most likely need calibration data in addition to the Level 1 data files themselves. This data can be anything required. The SOC will provide a directory where these files will be placed on the SOC pipeline system, and the instrument pipeline code will be able to access them there.

The combination of the Level 1 file (and detached PDS label) and the data provided must be sufficient to produce each Level 2 file. If housekeeping information (instrument or spacecraft) is needed, these must be already in the header (or tables) of the Level 1 file. If continuously varying values (e.g. temperature over many seconds, etc.) are needed, a FITS table will be written into the Level 1 file with this information.

6.5 Documentation

All code and data files shall be accompanied by thorough documentation. The code itself should have clear and appropriate comments throughout. Error conditions shall be documented in the code as well (see section 6.6 for more on this topic).

Documentation and code shall be written assuming that it will be read by someone years from now who is unfamiliar with the system. Understanding of the documentation should not rely on special scientific knowledge or specific domain knowledge.

6.6 Error Conditions and Integration Troubleshooting

If there are any reasons the code might abort processing, these shall be defined, and the resulting action should be specified. Also, if such an abort happens, the reason should be noted in the status file written ("out_status" file described above).

A contact person shall be specified who will be responsible for helping the SOC operators when unexpected errors occur. This person should be able to help with debugging and should also be available to respond and help in two days or less for consultation during the pipeline integration process.

6.7 PDS Archiving

The Level 2 (output) files, as well as Level 1 files, will be archived to PDS. Therefore the format of the files shall meet PDS requirements. This includes size requirements set forth in the *New Horizons* Data Management and Archiving Plan (#05310-DMAP-01).

PDS detached labels for the Level 2 files shall be generated by either the instrument pipeline code or by the SOC using a translation table (from FITS to PDS keywords). Which method is appropriate will be determined on an instrument by instrument basis. If generated by the SOC, "in_pds_header" and "out_pds_header" can be ignored in the instrument pipeline code.

6.8 Configuration Management

All code, documentation, and calibration files will be put under configuration management at the SOC. Also, the necessary keywords shall be inserted into the Level 2 headers by the Level 2 pipeline code to specify the version of the code and data used to produce the Level 2 files. This ensures that data is traceable back to the correct code version.

In addition, the Level 2 pipeline code shall insert, using header keywords, the calibration files used and the versions thereof (if applicable).

6.9 Pipeline Updates

Updates to the instrument pipeline (including code, documentation, and calibration data) are to be delivered to SOC personnel for integration; all such updates will require appropriate documentation and will fall under SOC CM. The code will be checked in to the SOC configuration management after regression tests are run. Any special instructions or changes should be communicated to the SOC personnel, and a file containing release notes (called "RELEASE_NOTES") should accompany the update. The SOC personnel will notify the instrument team when the new update is in place and active.

6.10 Acceptance Review

The instrument pipeline (including code, documentation, and calibration data) shall be subject to an acceptance review.

6.11 Longevity

The code and all third party libraries and data files used shall meet the longevity requirements as specified in the *New Horizons* Longevity Plan (#7399-9009). Also, development, documentation, and testing of the pipeline shall adhere to these requirements.