Scientific Data Archives in Hayabusa2 Mission

By Yukio YAMAMOTO,1)2) Naoki KOBAYASHI,1)2) Masahiko HAYAKAWA,1)2) and Yoshiaki ISHIHARA1)

1)Institute of Space and Astronautical Science, JAXA, Sagamihara, Japan
2)Department of Space and Astronautical Science, SOKENDAI (The Graduate University for Advanced Studies), Sagamihara, Japan

(Received August 7th, 2015)

The Hayabusa2 project will produce scientific data archives in PDS4, which is newly developed by NASA. The Hayabusa2 project defines the DAC team under the science team that dedicates scientific data archives to manage data archiving tasks. The DAC team starts from the basic design to produce PDS4 on ISAS common infrastructure: SIRIUS, Reformatter, and DARTS. In this report, we will introduce the common infrastructure and how to use it in Hayabusa2.

Key Words: Hayabusa2, DAC, PDS4, SPICE

1. Introduction

Hayabusa21) is a subsequent mission of the original Hayabusa to obtain sample materials from C-type asteroid 162173 (1999 JU3). It was launched on December 3rd, 2014 by H-IIA launch vehicle from Tanegashima Space Center.

For lunar and planetary data archives, Planetary Data System (PDS)2) developed by NASA has been de-fact standard in the world. The PDS version 4 (PDS4) is recently released, and Hayabusa2 is the first mission to adopt PDS4 in Japan. Therefore, Hayabusa2 must leave the experience and know-how to future missions such as BepiColombo MMO, which is already determined to use PDS4.

Also, the collaboration with OSIRIS-REx, a sample return mission from an asteroid by NASA, must be considered in data preparation. These two missions are in the same concept, and obtained data are expected to be compared everywhere. Therefore, the contents of the PDS4 label should be the same if the keyword has the same meaning, and the negotiation for PDS4 implementation should be required in advance.

2. Organization of Data Archives

2.1. DAC team and scope

Data archive team (DAC) is defined under the science team as well as the instrument teams to promote scientific data archives(Fig. 1). The DAC team handles all scientific data-related tasks except for data policy and returned samples. The science data policy is discussed in the Hayabusa2 Joint Science Team (HJST) and returned samples data are scoped in Hayabusa2 Sample Allocation Committee (HSAC). The main scope of the DAC team is of processing and documentation considering standards.

2.2. Supervisors

The DAC team supports the activities of data production in the instrument teams to follow the standard in planetary science and keep consistency in the project. Therefore, the DAC team defines supervisors in all instrument teams and some subgroups, and holds meetings regularly. The subgroups

in the data archives are determined considering the importance of the data production, such as shape model creation, PDS4 preparation, ancillary archives, and documentation.

Fig. 1. Structure of the Hayabusa2 science team. Data Archive team is defined under the Hayabusa2 Science Team.

2.3. Data archive meetings and camps

There are several types of meetings for data archives. The DAC team has weekly meeting to manage the schedule with work breakdown structure and action items. The working group for data archives is held once in several months, where archiving policy in technical part is discussed, and the rules are determined. The data archive camps are held to promote the implementation of data related tasks, such as software development and documentation.

3. Instruments

The Hayabusa2 spacecraft has remote sensing instruments and several separate modules. The scientific data obtained by these instruments and modules will be archived and published after a protection period. Table 1 shows the brief summary of their data formats. The data format FITS is often used in astronomical satellites, and ASCII text will be used. PDS4 allows to use these formats, which means FITS and ASCII with PDS4 label can be recognized as PDS4 dataset. Some instruments are used both for science and engineering analysis,
and the publication of engineering data is not planned so far.

ONC\(^3\) is an optical imager system with three frame cameras: a telescopic camera (ONC-T) and two wide-angle cameras (ONC-W1 and ONC-W2). These data are archived in FITS that is often used in astronomical satellites.

NIRS\(^3\) is a near-infrared spectrometer to observe 3μm absorption band for hydrated minerals. As well as NIRS onboard the original Hayabusa, the data are currently archived in FITS.

TIR\(^3\) is a thermal infrared imager to determine thermo-physical properties of the asteroid surface. TIR is designed on Longwave Infrared Camera (LIR\(^6\)) onboard Akatsuki, Venus Climate Orbiter by JAXA. The data is currently archived in FITS as well as LIR.

LIDAR\(^7\) is Laser Altimeter, which is a part of attitude and orbit control system (AOCS). Scientific measurements are scheduled such as the determination of shape, mass, surface properties, etc. These data will be archived as scientific data.

DCAM\(^3\,8\) is a very small optical camera system for the observation of the artificial impact experiment by Small Carry-on Impactor (SCI\(^9\)). DCAM\(^3\) is composed of two categories: DCAM\(^3\)-A for monitoring SCI operation by an analog camera and DCAM\(^3\)-D for scientific observations by a digital camera with wide-angle and high-resolution.

MASCOT (Mobile Asteroid Surface Scout\(^10\)) is a hopping lander developed by DLR and CNES. It has four instruments: CAM (wide-angle imager), MARA(6-band thermal radiator), MAG(3-axis magnetometer), and MicroOmega (hyper spectral infrared microscope). The publication of MASCOT data will be discussed under HJST.

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Instrument Type</th>
<th>Format</th>
</tr>
</thead>
<tbody>
<tr>
<td>ONC-T/W1/W2</td>
<td>Imager</td>
<td>FITS</td>
</tr>
<tr>
<td>NIRS(^3)</td>
<td>Spectrometer</td>
<td>FITS</td>
</tr>
<tr>
<td>TIR</td>
<td>Imager</td>
<td>FITS</td>
</tr>
<tr>
<td>LIDAR</td>
<td>Altimeter</td>
<td>ASCII</td>
</tr>
<tr>
<td>DCAM(^3)-D</td>
<td>Imager</td>
<td>(T.B.D.)</td>
</tr>
<tr>
<td>MASCOT</td>
<td>Lander with 4 instruments</td>
<td>(T.B.D.)</td>
</tr>
</tbody>
</table>

4. Ground System for Data Archives

4.1. Data flow

Figure 1 shows generic data flow to operate spacecraft and build data archives in the Institute of Space and Astronautical Science (ISAS). The ground system is divided into two parts for the different purposes: (1) spacecraft operation to maintain spacecraft functions, (2) data processing for analysis and archives.

The spacecraft operations are performed at Sagamihara Space Operation Center (SSOC). The telemetry data obtained by antennas are transferred to SSOC. While the spacecraft operations place importance on the real-time monitoring, the processing for data analysis and archives place its importance on data integrity. Therefore, there are two different concepts to relay telemetry data in the ground system: one is streaming transfer implemented by a unique TCP protocol: Space Data Transfer Protocol (SDTP)\(^11\) real transfer mode and the other is file-based transfer using FTP or SDTP non-real transfer mode.

SDTP is essential to deliver telemetry data in ISAS. SDTP handles CCSDS compliant frames and packets, and additional information is attached on each CCSDS packet. SDTP users can obtain telemetry data specifying antenna identifier, virtual channel identifier (VCID), application process identifier (APID), date-time, etc. VCID and APID are defined in CCSDS. VCID is often used to identify real-time telemetry and repro telemetry, and APID is used to identify the different type of data.

There are four major sub-systems to build data archives: SIRIUS, Reformatter, Science Data Storage, and DARTS.\(^12\) The boundary point in the ground system between spacecraft operations and data processing for archives is SIRIUS.

4.2. SIRIUS

SIRIUS is a telemetry database that stores CCSDS frames and packets. The telemetry received on the ground system may include duplicate packets, therefore SIRIUS removes the duplicates, and sorts the CCSDS packets by the combination of Application Process Identifier (APID), Packet Sequence Count, and packet generation time, where the packet generation time is converted from TIME CODE FIELD,\(^13\) in the packet secondary header using a unique Time-Calibration System\(^14\) in ISAS.
The instrument teams can receive CCSDS packets from SIRIUS using SDTP. Therefore, each packet is attached with additional SDTP information that includes converted time in UTC.

Because packets are merged and sorted after SIRIUS registration, there is no necessity to consider which antenna is used, or whether telemetry data are in real-time telemetry or repro telemetry. Therefore, a unique antenna identifier is assigned to SIRIUS although SIRIUS is not an actual antenna, and VCID is also assigned to merged and sorted telemetry that are not included in the actual telemetry transferred by the spacecraft.

4.3. Reformatter
Reformatter is a server where each project develops software to handle CCSDS space packets for higher-level products. Using the reformatter servers, instrument teams can connect to SIRIUS directly to fetch CCSDS packets via SDTP. Most of the processing are pipelined to produce data automatically. The reformatter server also mounts external storage to store the processed data. This external storage is shared in all projects, therefore, available volumes are allocated following the projects’ requirements.

Recent reformatter servers are prepared as Redhat Linux on virtual machines if there are no special requirements. Applying virtual machines, the long-term preservation of the processing environment can be achieved for long-term missions.

4.4. Science Data Storage
The processed data on the reformatter server are stored into high quality storage system with huge volumes. We call this storage system Science Data Storage. The stored data in the storage are also accessible from DARTS servers. The Science Data Storage is mounted as rewritable from the reformatter server, but it is mounted as read-only from DARTS servers to keep security.

4.5. DARTS
DARTS is an acronym of Data ARchives and Transmission System. It is a publication system to open the processed data via the Internet. The processed data are normally generated on the reformatter, but DARTS sometimes open other dataset such as data obtained at International Space Station, or laboratory analysis of Hayabusa returned samples. However, it is the basic philosophy that DARTS provides science data for scientists in the world.

Inside DARTS, the management of website contents and publication data are different. The website contents are registered to our repository system to control histories.

5. Common Software
Preparing the final outputs for science, ISAS has started to provide common software for recent spacecraft. This common software is designed to support developers beyond missions. The Hayabusa2 data must be archived in the PDS4 utilizing these tools. Besides, ancillary data such as orbit and attitude of the spacecraft is archived in SPICE.15)

5.1. SDTP Library
SDTP library is essential to obtain telemetry data. It is a C-language library. All applications in SSOC and reformatter servers using SDTP use it to receive telemetry data.

5.2. Time-Calibration System
Time-Calibration System is a common system after Hisaki (SPRINT-A) mission. The system is composed of a server to produce time calibration table, and C-language library to calibrate the packet time. Most of the scientific spacecraft in ISAS produce the time calibration packet periodically, and the calibration table is generated on the server using the time calibration packets.

The time calibration table is intrinsically equivalent to SPICE SCLK, which includes spacecraft clock and UTC pair in each line but includes much more information to validate the calculation from the source.

5.3. GSTOS
GSTOS16) is an acronym of Generic Spacecraft Test and Operations Software. The design of command and telemetry is stored into Satellite Information Base Version 2 (SIB2) in XML. GSTOS also provides SIB2 related software: design tools with Graphical User Interface, and conversion program to binary object file which can used in a C-language program. Wrapped C-language library, GTAPI, is useful to obtain engineering conversion data.

5.4. L1TSD
L1TSD is a utility software to deal with Level-1 Time Series Data. The current target is the conversion from House Keeping (HK) telemetry to other formats. L1TSD Suite software provides HK telemetry without users needing to handle engineering conversion.

For example, L1TSD CSV provides HK telemetry in Comma-Separated Values (CSV) format. L1TSD FITS also provides specific items of HK telemetry in FITS format. L1TSD SPICE converts the correspondence table of Time Calibration System to SPICE SCLK and quaternions in HK telemetry into SPICE CK. This L1TSD software is based on GSTOS.

5.5. SPICE
The Hayabusa2 project provides SPICE kernels as well as other planetary missions. For example, SPK is generated converting an ISAS’s original orbit format. SCLK and CK are created by L1TSD SPICE. FK and IK must be written manually. After the Hayabusa2 project reveals the features of the asteroid 1999 JU3, PCK and DSK are generated for the rotation properties and shape model respectively.

5.6. PDS4
JAXA’s experience of planetary data archives is Hayabusa and Kaguya. Both missions provide their data in PDS version 3 (PDS3). Hayabusa2 is the first mission to prepare scientific data in PDS4. PDS4 is designed using XML technologies to describe metadata.

PDS4 website provides three files in each release. These three files’ extensions are xsd, sch, and xml respectively. The extension .xsd is an XML schema file to produce a PDS4 label. The extension .sch is an XML schematron file to validate a PDS4 label. The extension .xml is an XML file to describe these PDS4 XML-related files. These three core files are
under the Planetary Data System (PDS) namespace.

Besides the PDS namespace, each PDS node is preparing XML schema and schematron files under each namespace. Also, a mission can prepare XML-related files under the mission-specific namespace. In the Hayabusa2 mission, XML files can be considered under the namespace PDS, the Small Bodies, and the Hayabusa2. Technically, JAXA or ISAS can have the namespace for cross-cutting perspective, but the well-processed data after calibration may dilute the meaning of such keywords. Such an agency dependent namespace can be used only in ISAS internally even if it is created.

6. Quality Control

6.1. PDS compatible level

The quality control of data depends on the PDS mechanism. PDS is not only a name of data format, but also a framework to implement long-term preservation. PDS requires the peer-review for the PDS compliant data. JAXA has tried to define the PDS compatible level to clarify the data quality (Table 2). The level 1 is that the data are checked by PDS validation provided by NASA. This tool guarantees the data follows the rules of the PDS format. The level 2 means that the data are provided together with enough documents to support science. The level 3 is that the data are checked by peer-review. The level 4 means that the data are published via NASA PDS website.

<table>
<thead>
<tr>
<th>Level</th>
<th>Meanings</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The data are validated by PDS validation tool.</td>
</tr>
<tr>
<td>2</td>
<td>The data are well documented.</td>
</tr>
<tr>
<td>3</td>
<td>The data are peer-reviewed.</td>
</tr>
<tr>
<td>4</td>
<td>The data are released via NASA PDS website.</td>
</tr>
</tbody>
</table>

6.2. Software Interface Specification

There is a document called as Software Interface Specification (SIS) in planetary explorations. This document is used to describe the data-related information: instrument overview, processing level, the list of dataset, etc. In the Hayabusa2 project, the template of the SIS document is created comparing with that of several missions. The instrument teams use this template to explain each dataset.

7. Conclusion

The Hayabusa2 project defined the DAC team to consider the data archiving activities is worthwhile. The ISAS infrastructure composed of sub-systems and common software is utilized to build PDS4 archives and SPICE kernels. Especially, PDS4 activity in the Hayabusa2 project should be enhanced to other missions inside and outside JAXA.

References