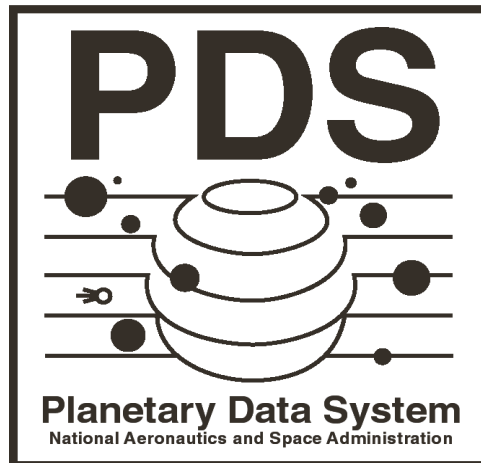


# Planetary Data System

## Data Integrity Use Cases

**DRAFT**

October 11, 2006  
Version 0.10060926



Jet Propulsion Laboratory  
Pasadena, California

JPL D-xxxxx

# CHANGE LOG

Revision	Date	Description	Author
Start Draft	2006-09-05	First Draft	S. Hughes
	2006-09-07	Modified Draft	D. Crichton
0.10060914	2006-09-14	Updated draft from WG telecon	D. Crichton E. Guinness S. Hughes
0.10060926	2006-09-26	Updated draft from 9/26 WG telecon	D. Crichton
0.10061011	2006-10-11	Finalize updates from WG members	D. Crichton

# Table of Contents

<b>1</b>	<b>INTRODUCTION .....</b>	<b>4</b>
<b>2</b>	<b>ACTORS.....</b>	<b>5</b>
<b>3</b>	<b>DEFINITIONS .....</b>	<b>6</b>
<b>4</b>	<b>REQUIREMENTS .....</b>	<b>7</b>
<b>5</b>	<b>USE CASES .....</b>	<b>8</b>
5.1	UC-1 Data Delivery.....	8
5.2	UC-2 Data Distribution .....	8
5.3	UC-3 Data Transfer to Deep Archive .....	9
5.4	UC-4 Data Node Termination .....	9
5.5	UC-5 Archive Integrity .....	9
5.6	UC-6 Media Migration.....	10
5.7	UC-7 Recover Data from Deep Archive .....	10
5.8	UC-8 Data Transfer between PDS Nodes.....	11

# 1 Introduction

The purpose of this document is to capture use case scenarios for PDS data integrity as part of the discussion from the PDS Data Integrity Working Group. This set of use cases are derived from the PDS level 3 requirements. For this document, data integrity will have the following broad meaning.

1. The condition in which data are identically maintained during any operation, such as transfer, storage, and retrieval.
2. The preservation of data for their intended use.

The intention of the use cases will be to drive level 4 and 5 requirements along with an implementation plan for data integrity across the PDS. It is also the intent that these data integrity use cases be media independent.

While, the scope of this document is on identifying use cases that address file corruption during the preservation and transfer of data to and from PDS, it is important to note that accounting for a set of files within a collection is also a critical function necessary to verify the integrity of collections such as volumes. PDS currently plans to address integrity of collections as part of a later phase of data integrity which integrates the end-to-end tracking subsystem use cases and requirements.

## 2 Actors

An actor is a user who is involved in any step of the life cycle of a PDS data product from data ingestion to data usage. The following actors are referenced or implied in the PDS data integrity use cases.

- PDS Nodes
  1. Discipline Nodes
  2. Data Nodes
  3. Engineering Node
  
- Deep Archive Manager
  1. National Space Science Data Center (NSSDC)
  
- Data Provider
  1. Planetary Scientist
  2. Campaign
  3. Experiment
  4. Mission Proposal Team
  5. Mission Flight Project
  
- Data Consumer
  1. Planetary Scientist
  2. Mission Flight Project members
  3. Mission Operations
  4. Educator
  5. General Public

# 3 Definitions

The following definitions are used in the use case scenarios.

1. **Actors.** An actor is a person, organization, or external system that plays a role in one or more interactions with your system
2. **Corrupt file** – Any file that contains at least one bit error.
3. **Data Consumer** - Entities that receive data from PDS
4. **Data Product** – A data product label and one or more data objects.
5. **Data Product Label** – One or more data object descriptions.
6. **Data Provider** - Entities that submit data to PDS
7. **Data Transfer Package** – Files organized into a collection for the purposes of moving between two points
8. **PDS Node** – Any PDS node including Discipline Nodes, Data Nodes, and the Engineering Node. The Discipline Nodes include both science and support nodes.
9. **Physical Media** – Any computer system device used for short or long term storage of data including but not limited to optical media, tape, and magnetic disk.
10. **Use cases.** A use case describes a sequence of actions that provide something of measurable value to an actor.
11. **Volume** – Any organized collection files that reside on physical media for the purpose of near term storage, online access, data submission, electronic distribution, or long-term archive. Note that this definition includes the PDS archive volume.

# 4 Requirements

The following existing level three requirements relate to PDS data integrity and are referenced in the use cases.

- 2.5.1 PDS will develop and publish procedures for accepting archival data
- 2.5.2 PDS will implement procedures for accepting archival data
- 2.5.3 PDS will inform a data provider why a rejected archival product does not meet archiving standards
- 2.6.1 PDS will develop and publish procedures for storing archival data
- 2.6.2 PDS will maintain appropriate storage for the PDS archive
- 2.6.4 PDS will maintain appropriate storage for non-archived data managed by the PDS
- 2.8.3 PDS will meet U.S. federal regulations for the preservation and management of data.
- 3.2.3 PDS will provide mechanisms to ensure that data have been transferred intact. (Note that 3.2.1 and 3.2.2 specify that mechanisms must be available to transfer data to users both online and offline.)
- 4.1.1 PDS will define and maintain a set of quality, quantity and continuity (QQC) requirements for ensuring long term preservation of the archive
- 4.1.2 PDS will develop and implement procedures for periodically ensuring the integrity of the data.
- 4.1.3 PDS will develop and implement procedures for periodically refreshing the data by updating the underlying storage technology.
- 4.1.4 PDS will develop and implement a disaster recovery plan for the archive
- 4.1.5 PDS will meet U.S. federal regulations for preservation and management of the data through its Memorandum of Understanding (MOU) with the National Space Science Data Center (NSSDC)

# 5 Use Cases

Each use case describes one or more high level scenarios where one actor (user) or a group of actors impact data integrity. These use cases are suggested by several level one, two, and three PDS requirements. These use cases will subsequently aid in the formulation of more detailed data integrity requirements.

## 5.1 UC-1 Data Delivery

**Description:** A Data Provider produces and transfers the data to the PDS via a pipe-line or assembly-line production of data files to a PDS Node. (2.5.1, 2.5.2, 2.5.3, 3.2.3). The PDS Node prepares and receives the data verifying the integrity of the files.

**Actors:** Data Provider, PDS Node

**Scenarios:**

1. A Data Provider ensures that the data file to be transferred has not been corrupted.
2. A Data Provider readies the data file for transfer.
3. The PDS Node ensures that the data file received has not been corrupted during transfer.
4. The PDS Node determines that a data file is corrupted and rejects the data file.
5. The Data Provider investigates the cause of the corruption and submits a new data file.

## 5.2 UC-2 Data Distribution

**Description:** The PDS makes the data products / volumes available (i.e. distributed) electronically to Data Consumers. (2.7.3, 3.2.3)

**Actors:** PDS Node, Data Consumer

**Scenarios:**

1. The PDS Node ensures that the data file(s) to be distributed has not been corrupted.
2. The PDS Node readies the data file(s) for transfer.
3. A Data Consumer checks that the data file(s) received has not been corrupted during transfer.
4. A Data Consumer requests replacements for those files corrupted during the transfer.



## 5.3 UC-3 Data Transfer to Deep Archive

**Description:** The PDS preserves data products / volumes in long-term storage at the NSSDC deep archive. (2.8.3, 3.2.3, 4.1.5)

**Actors:** PDS Node, NSSDC

**Scenarios:**

1. The PDS Node ensures that the data files, to be transferred as a data transfer package, have not been corrupted.
2. The PDS Node ensures that the created data transfer package file is not corrupt.
3. PDS transfers data products / volumes captured in the data transfer package to long-term storage at the NSSDC.
4. The NSSDC rejects and requests retransmission of a corrupted data transfer package.

## 5.4 UC-4 Data Node Termination

**Description:** A Data Node is dissolved and the responsibility for archiving and distributing the data holdings passes to the permanent PDS Nodes. (2.5.1, 2.5.2, 2.5.3, 2.8.3, 3.2.3)

**Actors:** PDS Data Node, PDS Discipline Node

**Scenarios:**

1. The PDS Data Node ensures that the files in its data holdings to be transferred have not been corrupted.
2. The PDS Data Node transfers its data files to a PDS Discipline Node.
3. The PDS Discipline Node rejects and requests retransmission of corrupted data files.

## 5.5 UC-5 Archive Integrity

**Description:** PDS Nodes maintain the integrity of their archives by periodically accounting for submitted data and verifying that no files have been corrupted (2.6.1, 2.6.2, 2.6.3, 2.8.3, 4.1.1, 4.1.2, 4.1.4)

**Actors:** PDS Nodes

**Scenarios:**

1. The PDS Nodes ensure that all data files received from the Data Providers are accounted for.
2. The PDS Nodes ensure that the data files on existing media have not been corrupted.
3. The PDS Nodes ensure that they have an accessible backup copy of their archive holdings.
4. The PDS Nodes develop a plan for recovery of missing or corrupted files on existing media.

## 5.6 UC-6 Media Migration

**Description:** The PDS Nodes and the NSSDC periodically verify the integrity of data files stored on a physical media in the repository. Transfer data files to new media to ensure data integrity because a) existing media is failing, b) reduce physical space needs by moving to denser media, or c) reduce cost by moving to denser media. (2.8.3, 3.2.3, 4.1.2, 4.1.3, 4.1.4)

**Actors:** PDS Node, NSSDC,

**Scenarios:**

1. The PDS Nodes and NSSDC ensure that the data files on the existing media have not been corrupted.
2. The PDS Nodes and NSSDC transfer data files from one physical media to another.
3. The PDS Nodes and NSSDC ensure that the data files transferred have not been corrupted.

## 5.7 UC-7 Recover Data from Deep Archive

**Description:** A PDS Node determines that integrity of an existing set of archival data may have been compromised. The PDS Node recovers data from the NSSDC (4.1.5).

**Actors:** PDS Node, NSSDC

**Scenarios:**

1. The PDS Node requests a copy of the archival data from the NSSDC
2. The NSSDC ensures that the data files to be transferred have not been corrupted.
3. The NSSDC prepares and transfers data files to the PDS Node
4. The PDS Node rejects and requests retransmission of data files corrupted during transfer.

## 5.8 UC-8 Data Transfer between PDS Nodes

**Description:** A PDS Node transfers data holdings to another PDS Node. (2.7.3, 3.2.3)

**Actors:** PDS Node

**Scenarios:**

1. A PDS Node ensures that the data files to be transferred have not been corrupted.
2. A PDS Node readies the data files for transfer.
3. The data files are transferred from the source Node to the destination Node.
4. The destination PDS Node rejects and requests retransmission of data files corrupted during transfer.