

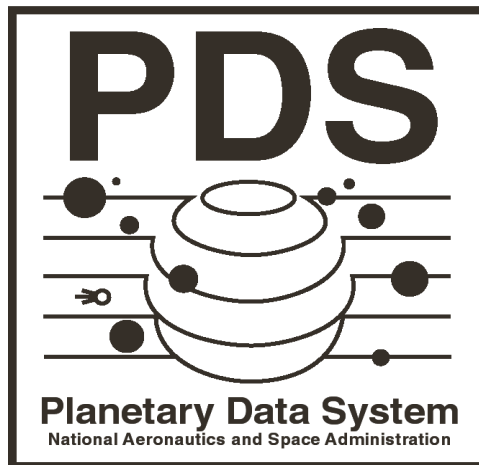
Planetary Data System

PDS-NSSDC Interface

Beta Test Plan/Procedure

March 13, 2007

Version 1.0



Jet Propulsion Laboratory
Pasadena, California

CHANGE LOG

Revision	Date	Description	Author
0.20070304	Mar 04, 2007	Initial cut based on the EN Test Plan	E. Law
0.20070305	Mar 05, 2007	Added specific procedure steps.	J. Wang
1.0	Mar 13, 2007	Cleaned up some stuff here and there.	S. Hardman

TABLE OF CONTENTS

Section Title	Page
1.0 INTRODUCTION	4
1.1 Purpose	4
1.2 Scope	4
1.3 Audience.....	4
1.4 Controlling Documents	4
1.5 Applicable Documents.....	4
2.0 MPGA SOFTWARE CONSTRAINTS	5
3.0 TEST PLAN / PROCEDURE	6
3.1 Select Volume and Request NSSDC Identifier.....	6
3.2 Create Checksum.....	6
3.3 Create AIP at the Node.....	7
3.4 Transfer AIP to NSSDC	8
3.5 Notify NSSDC.....	8
3.6 Retrieve and Store AIP at NSSDC	8
3.7 Retrieve and Repack the Volume and Notify Node	9
3.8 Transfer AIP to Node.....	9
3.9 Unpack and Verify AIP.....	9
APPENDIX A ACRONYMS	11

1.0 INTRODUCTION

The National Space Science Data Center (NSSDC) has defined an Archival Information Package (AIP) and is in the process of tailoring it to support the packaging and transfer of a Planetary Data System (PDS) volume. The NSSDC has provided the Multi-File Package Generator and Analyzer (MPGA) software, which is to be run at for the PDS Node for generating AIPs using the Consultative Committee for Space Data Systems (CCSDS) Standard Formatted Data Unit (SFDU) container.

1.1 Purpose

The purpose of this document is to provide a plan, procedures and guidelines for beta testing the PDS-NSSDC interface at the Nodes.

1.2 Scope

This document is scoped according to the following capabilities NSSDC has delivered to Engineering Node (EN):

- Generate AIP using MPGA
- Deliver AIP to NSSDC via FTP
- Retrieve AIP after deep archived by NSSDC
- Unpack AIP using Splitter

1.3 Audience

This document is written primarily for the Node Staff who will participate in the beta testing of the PDS-NSSDC interface. The expected audience includes:

- ATMOS Node staff
- IMG Node staff
- SBN Node staff

1.4 Controlling Documents

[1] Planetary Data System (PDS) Level 1, 2 and 3 Requirements, August, 2006.

[2] PDS-NSSDC Memorandum of Understanding (MOU).

1.5 Applicable Documents

[3] NSSDC MPGA Installation Guide, TBD.

2.0 MPGA SOFTWARE CONSTRAINTS

The beta test focuses on testing Node specific PDS data sets and/or various volume sizes, within the current MPGA software constraints as follows:

This round of testing will be limited to Nodes with Linux computing environments. The tests of the MPGA software at the Engineering Node were conducted on a system running Red Hat Linux Enterprise Workstation Release 3. There are plans to make the MPGA software available on Solaris and Windows platforms but those are not available and/or have not been tested by the Engineering Node at this time.

Although this is not a constraint of the software, it is suggested that volumes of 4 gigabytes or less be used during this round of testing due to the limitations of using FTP for the transfer mechanism. Even with this size constraint the Engineering Node has experienced transfer times upwards of 3 hours when transferring an AIP of approximately 4 gigabytes. We have also encountered a few cases where the transfer failed and had to be restarted.

3.0 TEST PLAN / PROCEDURE

This section of the document provides the step-by-step procedures as well as helpful guidelines/information for packaging and transferring a single AIP. It is recommended that the Nodes exercise the set of procedures repeatedly for multiple data sets and various sizes of volumes. The procedures assume that the MPGA software, provided by the NSSDC, has been installed in the Node environments and that the Node specific test volumes are accessible in those environments.

3.1 Select Volume and Request NSSDC Identifier

The test will begin with the selection of appropriate Node specific PDS compliant data set(s) and/or volume(s) from which to create the AIP. In order to request an NSSDC identifier, send e-mail to Dave Williams (dwilliam@mail630.gsfc.nasa.gov) and attach the AAREADME.TXT and VOLDESC.CAT files from the selected volume. Information from these files is used to determine the identifier. The identifier returned from Dave will need to be specified in the MPGA configuration file prior to creation of the AIP.

This step routinely takes a day or two to complete. While you are waiting, you can perform the next step.

3.2 Create Checksum

While you are waiting for Dave to get back to you with the NSSDC identifier, you can create the checksum logs for the selected volume. The EN has provided a very rudimentary method for generating checksum logs for the files contained in a volume directory structure. If someone else has a better method for accomplishing this, please feel free to volunteer your method.

This method assumes that the *md5sum* program is accessible in the Node's target environment and is in the executable path. The examples below assume the selected volume is located in the */data/target/* directory and that the checksum log files will be written to the users home directory. The following commands should be performed to generate checksum values for the files that make up the selected data set(s) and/or volume(s):

```
# cd /data/target/  
# md5sum * > $HOME/target_level1_orig.out  
# md5sum */* > $HOME/target_level2_orig.out  
# md5sum */** > $HOME/target_level3_orig.out
```

Continue the above commands based on the number of levels in the selected volume directory tree. If the user encounters the "Argument list too long" error

when running the command on a given level, they may need to run the command for specific directories under that level. So, if this error was occurring at level 2, and level 2 consisted of three subdirectories, then the following commands should be executed instead:

```
# md5sum dir1/* > $HOME/target_level2_dir1_orig.out
# md5sum dir2/* > $HOME/target_level2_dir2_orig.out
# md5sum dir3/* > $HOME/target_level2_dir3_orig.out
```

3.3 Create AIP at the Node

This procedure creates an AIP of the selected data set(s) and/or volume(s) using the MPGA software.

The first step is to modify the configuration file. The configuration file *stage_in_directory_v6_0.lst* is located in the *MPGA/run/* directory. The file contains two tab-delimited lines. The content of the first line is fixed. The second line consists of the following fields:

- Input Directory
The directory specification of the selected volume to package.
- Output Package
The file specification for the AIP file. The example below shows a relative path indicating that the script was executed from the *MPGA/run/* directory.
- Primary Collection ID / NSSDC ID
The NSSDC identifier returned by Dave.
- Grouping Recommended Name
A name given to the volume. Doesn't matter for this test.
- Begin Datetime
- End Datetime
These values are not necessary for this test. Default: N/A
- Volume ID
This value taken from the VOLDESC.CAT file.
- Volume Version ID
This value taken from the VOLDESC.CAT file.

An example configuration file is as follows:

```
PDS_RUN Maintain directory structure / view      DIRECTORY_CONTENT
/data/target/      ../AIP_OUTPUT/output_AIP_target.aip      PSPG-00730
TARGET            N/A      N/A      TARGET      VERSION 1
```

The script to initiate the AIP packaging process is *START-JOB--MPGA-package_generator_v6_0alpha.com* and is located in the *MPGA/run/* directory. Change the location for “output” and “log” parameters to create those files in the

same location with the AIP file. Content for the modified *START-JOB--MPGA-package_generator_v6_0alpha.com* is:

```
cd ../sys/
./MPGA-launch \
  -workdir=../MPGA_WORKING_DIR/ \
  -setup=../PDS-packager.v6_0.in_development.2.setup \
  -input=../run/stage_in_directory_v6_0.lst \
  -mlst_tspec=../TEMPLATE_SPEC/PDS-maplist.v6_0.in_development.2.tspec \
  -one_AIP_per_Row \
  -agen_tspec=../TEMPLATE_SPEC/PDS-aipgen.v6_0.in_development.2.tspec \
  -output=../AIP_OUTPUT/target/out-MPGA-package_generator-
msgs_target.diag \
  -log=../AIP_OUTPUT/target/out-MPGA-package_generator-msgs_target.log
&
cd ../run/
```

The second step is to execute the script. The script should be run from within the MPGA/run/ directory as follows:

```
# ./START-JOB--MPGA-package_generator_v6_0alpha.com
```

3.4 Transfer AIP to NSSDC

The AIP is transferred from the Node to the NSSDC via the File Transfer Protocol (FTP) mechanism. The IP address, user name, password and target directory for accessing an NSSDC FTP server are provided by NSSDC as follows:

Generate a tar file for the selected volume to group the AIP, diagnosis and log files. Due to the way that outputs are organized, the *tar* command can be run at the upper level where all the outputs are store. Generate the tar file as follows:

```
# cd MPGA/AIP_OUTPUT
# tar cvf target.tar target
```

The tar file is transferred to NSSDC via FTP as follows:

```
# ftp nssdcftp.gsfc.nasa.gov
  username: pds
  password: *****
ftp> send target.tar
ftp> quit
```

3.5 Notify NSSDC

Once the transfer of the AIP is completed, the Node will notify NSSDC of the content and name of the tar file. Notification consists of sending e-mail to Pat McCaslin (mccaslin@mail630.gsfc.nasa.gov).

3.6 Retrieve and Store AIP at NSSDC

NSSDC is responsible for retrieving the AIP from their FTP server and placing into deep archive at the NSSDC. The process for ingesting the AIP into deep archive is not part of this test and will not be elaborated on further. This step along with the next step routinely takes a few days.

3.7 Retrieve and Repack the Volume and Notify Node

Upon completion of ingestion of the AIP into the NSSDC deep archive, the package will be retrieved and packaged into an AIP again to be sent back to the sending Node. The AIP will be placed out on the NSSDC FTP site and the Node will receive e-mail from Pat McCaslin notifying them that the AIP is ready to be transferred back to the Node.

3.8 Transfer AIP to Node

Upon notification from Pat McCaslin that the AIP is ready for retrieval, the Node will transfer the AIP from the NSSDC FTP server back to the Node. The splitter software expects the AIPs to be located in the *MPGA/stage_out/* directory. Prior to transferring the AIP from NSSDC, change the local directory to this location.

```
# cd MPGA/stage_out
# ftp nssdcftp.gsfc.nasa.gov
  username: pds
  password: *****
ftp> get target.tar
ftp> quit
```

3.9 Unpack and Verify AIP

Upon transferring the AIP into the staging area, it can now be unpacked and verified. The following steps should be performed to unpack the AIP:

```
# cd MPGA/run
# ./START-JOB--MPGA-reconstitute_packages_v6_0alpha.com
```

The splitter generates a directory with a sequentially generated name and an associated attribute file. For example, file *TEST0000000036.att* and directory *TEST0000000036.dat.dir* could be the result of running the splitter against a target AIP. The content of directory *MPGA/stage_out/TEST0000000036.dat.dir/* would be the same layout as the original target volume.

Now that the AIP is unpacked, the checksum logs can be generated on the retrieved data for comparison with the original checksum logs. The following commands should be performed:

```
# cd MPGA/stage_out/TEST0000000036.data.dir/
# md5sum * > $HOME/target_level1_new.out
```

```
# md5sum */* > $HOME/target_level2_new.out  
# md5sum */** > $HOME/target_level3_new.out
```

Continue the above commands based on the number of levels in the selected volume directory tree.

Run the *diff* command on each pair of original and new checksum log files to examine the difference. A successful restore should result with no difference between the two sets of files.

```
# diff $HOME/target_level1_orig.out $HOME/target_level1_new.out  
# diff $HOME/target_level1_orig.out $HOME/target_level1_new.out  
# diff $HOME/target_level1_orig.out $HOME/target_level1_new.out
```

Continue the above commands based on the number of levels in the selected volume directory tree.

APPENDIX A ACRONYMS

Acronyms pertaining to this document:

AIP	Archival Information Package
CCSDS	Consultative Committee for Space Data Systems
EN	Engineering Node
FTP	File Transfer Protocol
JPL	Jet Propulsion Laboratory
MOU	Memorandum of Understanding
MPGA	Multi-File Package Generator and Analyzer
NASA	National Aeronautics and Space Administration
NSSDC	National Space Science Data Center
PDS	Planetary Data System
SFDU	Standard Formatted Data Unit