10. Metadata Administration

Every science data product generated and archived by the system must be described to the system by metadata that are put into an inventory and then used to retrieve and distribute the data to users of the system. The Earth Science Data Model, described in document 420-EMD-001, Release 7 Implementation Earth Science Data Model, organizes the metadata into groups of related attributes and services to be performed on the data products. These "core" attributes are necessary to identify, interpret and perform services on granules and collections. The Data Model also provides for "product-specific" attributes (PSAs), i.e., attributes which are unique to a specific data product.

The smallest aggregation of data that is independently described and inventoried in the system is referred to as a data granule. Granules are organized into logical groupings called collections in which the granule metadata varies principally by time or location, called single-type collections.

Every collection is described by an Earth Science Data Type (ESDT) and is made known to the system by adding the type to ECS. This means that the parameter values in the ESDT descriptor file must be added to the appropriate databases in the ECS system

Metadata administration includes creating and updating ESDTs. Collections may be modified and updated over time. Collection-level metadata can be updated by updating the ESDT. Granule-level metadata can be updated manually (i.e., not as a result of an operation such as subsetting, which modifies the science data content of a granule) by setting the Quality Assurance flags and explanations. Procedures for updating these flags are provided in Chapter 15, Quality Assurance.

10.1 ESDT Descriptor Files

The primary task in establishing a collection is providing the core and product-specific metadata attribute values. This is done by creating an Earth Science Data Type (ESDT) descriptor file. The descriptor file is also used to specify the data services that are available for granules that belong to the collection. The descriptor file and the DLL are the means by which a collection is made known to the ESDT Maintenance Service.

The ESDT descriptor is composed of the following information:

- Collection level metadata attributes with values contained in the descriptor.
- Granule level metadata attributes whose values are supplied primarily by the Product Generation Executives (PGEs) during runtime.
- Valid values and permitted ranges for all product-specific attributes.
- List of services for all the granules in the collection and events that trigger responses throughout the system.

The services that apply to a collection are specified in the ESDT descriptor file. Product-specific services, such as subsetting or a product-specific acquire, require executable code to enact those

services. This code is contained in the Ingest and Order Manager Server software. After the ESDT (both descriptor file and DLL) has been generated it must be installed using the ESDT Maintenance Service before the first data granule can be inserted. During this installation process, information from the ESDT Descriptor File is propagated to the Inventory Database and the Spatial Subscription Server Database, all of which must be operating during the ESDT installation process.

10.1.1 Steps in Generating a Descriptor File

ESDTs for Distributable Product

These are the typical steps used in generating a descriptor file:

1 Identify desired collection-level metadata attributes.
- For permanent and interim files use only the minimum attributes.
- For distributable products identify all applicable attributes. This will involve reading appropriate documentation and interacting with the data provider.

2 Identify granule-level attributes.
- If a sample metadata configuration file is available from the data provider, use this.

3 Check “valids” (allowable metadata values) for core attributes (write CCR if new valids are required).

4 Check PSAs (register PSAs if new).

5 Use custom built scripts and a text editor to generate the descriptor file.

6 Verify the descriptor file as outlined in Section 10.1.2.

7 Check descriptor files into ClearCase.

10.1.2 Verifying Descriptor Files

1 Run the PERL script "update.pl", following the instructions in the script prologue.
- This script makes sure that the inventory metadata attributes are all listed as event qualifiers in the EVENT group.

2 Run the PERL script esdtQC.pl following the instructions in the script prologue.
- This script checks for more than 30 common descriptor file errors.

3 Make any necessary corrections in response to errors issued.

4 Rerun the PERL script esdtQC.pl.

5 Repeat Steps 3 and 4 until there are no errors.

6 Run the testodl utility to ensure that there are no errors in the ODL structure for the descriptor file.

7 Make any necessary corrections in response to errors issued.

8 Rerun the testodl utility.

9 Repeat Steps 9 and 10 until there are no errors.

10.2 Preparation of Earth Science Data Types

An ESDT goes through pre-operational life cycle steps starting with an analysis of the collection's need and continuing through development and operational installation. This process involves actions by the Data Provider or User in addition to EMD. The procedures are detailed in Software Development (SD) Project Instruction SD-038 ESDT Creation, Testing, Maintenance and Integration at http://dmserver.hitc.com/EMD_PAL/index.html).

10.2.1 Definitions

Archive - A File Type indicating granules will be inserted to Data Server for long-term storage and acquisition for distribution.

Full - A level of metadata coverage intended for data products that are produced within the system.

Collection - A related group of data granules.

Granule - The smallest data element that is identified in the inventory tables.

Interim - A File Type indicating granules are temporarily stored in support of product generation.

Intermediate - A level of metadata coverage intended for contemporaneous data products that are not produced within the system.

Limited - A level of metadata coverage intended for heritage data products brought into the system for distribution

Minimal -A level of metadata coverage sufficient to uniquely identify a collection or granule.

Permanent - A File Type indicating static or semi-static granules that are used only as inputs in product generation.

Product Specific Attributes - Attributes that are defined by the data provider in support of searching for specific granules

Valid - An allowable metadata value.

10.2.2 Process

1 Need Analysis

- The baseline list of science ESDTs and their services is controlled by the ESDIS CCB. This baseline was established through an analysis of the system Functional and Performance Requirements Specification, the Technical Baseline established from inputs from the Ad Hoc Working Group on Production, and meetings with the individual data providers to define the basic requirements of each ESDT.
- The basic requirements are:
  - Data Provider File Designation,
  - File Type (Permanent, Interim, Archive)

  - Level of Metadata Coverage (Minimal, Limited, Intermediate, Full)
- For new ESDTs not currently in the development baseline, the result of the Need Analysis forms the basis for approving the inclusion of the ESDT into the system. This is accomplished through the CCR process.

2 ESDT Specification

- This step results in a set of specifications extending the results of the needed analysis and providing the information needed to implement an ESDT. This step is executed only if the ESDT has been included in the baseline. The roles and responsibilities for developing the specification are as above.
- The specifications must include:
  - ShortName and VersionID of the ESDT
  - A list of the metadata attributes needed, valids, and any constraints on attributes.
  - A list and specification of the services needed (e.g., specification of the INSERT, SEARCH, ACQUIRE and SUBSCRIPTION semantics).

3 ESDT Generation

- Once the ESDT Specification has been developed and the applicable attributes identified, the necessary metadata has to be gathered, the metadata values checked against the valid values and the product-specific attributes (PSA) need to be checked against the list of PSAs that are already defined (see Figure 10.2-1).
- Once the collection-level metadata and granule-level attributes have been checked, then the descriptor file is generated and testing and validation of the ESDT performed. This process is further elaborated in the sections below.
- For a one-of-a-kind, distributable product with Full metadata coverage, this process can take up to six weeks to accomplish. For a related group of products with identical services, much of the Descriptor File of the first ESDT can be reused, and the cycle time for preparing subsequent ESDTs in the related group is much less.

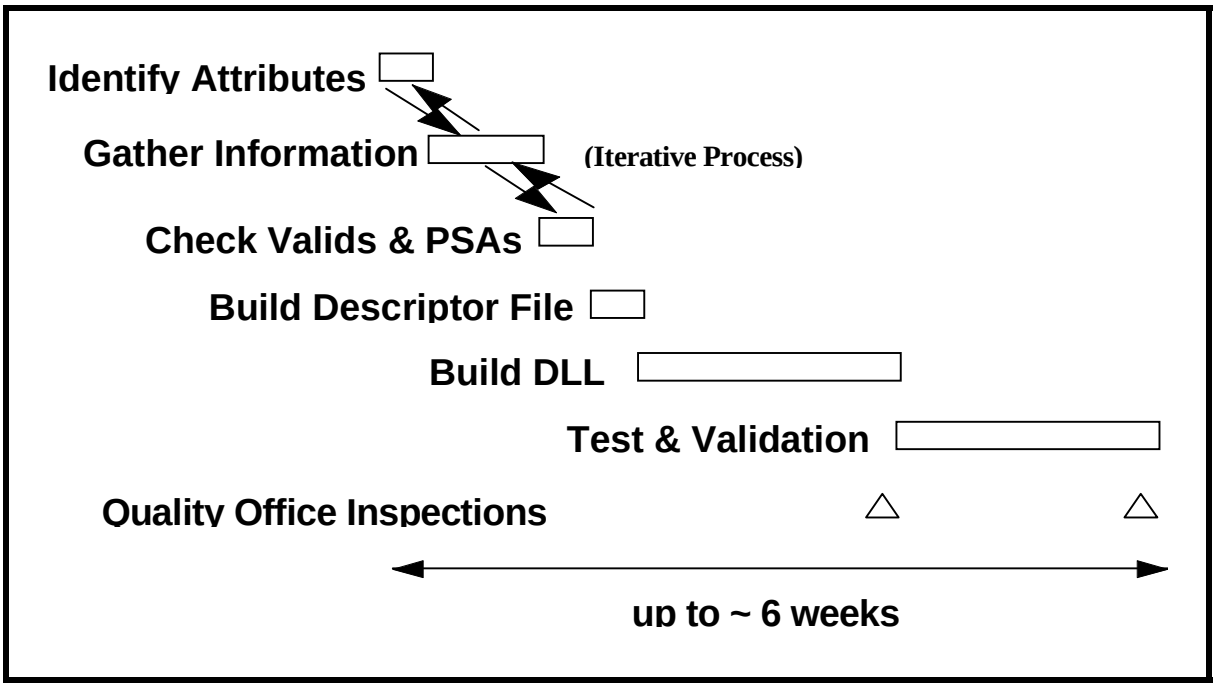


Figure 10.2-1. Steps in ESDT Development

10.3 Metadata Population

10.3.1 Collection-Level Metadata

A majority of the attributes in the Data Model apply to all the granules in the collection. These are known as collection-level attributes. There can be both core and product-specific collection-level attributes, defined once prior to establishing the collection.

Collection-level metadata is input either a text editor or a custom built script.

10.3.2 Granule-Level Metadata

The attributes in the Data Model that can vary on a granule-by-granule basis are known as granule-level attributes. There can be both core and product-specific granule-level attributes.

Granule-level metadata are specified and populated using the Metadata Configuration File (MCF). The MCF is derived from information contained in the ESDT descriptor file and is delivered by ESDT Maintenance Service for use by the Ingest Subsystem. The MCF specifies how the searchable metadata attributes will be populated in the Inventory database. For data products generated within the system, the science software or Product Generation Executive (PGE) interacts with the MCF using metadata tools contained in the Science Data Processing Toolkit. Through this process values are set for metadata attributes specified in the "source" MCF, such as the temporal or spatial coverage of each granule. These values are then inserted

into a "target" MCF at PGE run time. The MCF is used in a similar manner for data entering the system through the Data Pool Ingest.

Procedures for entering data into the system through Data Pool Ingest are described in Chapter 16, Ingest Each data granule consists of one or more physical files. Accompanying each granule is a metadata record; i.e., an ASCII file containing the granule-level attributes and their values in ODL. Only one metadata record is allowed per granule, i.e., no sub-granule records are allowed, and no metadata records are shared between granules.

10.3.3 Product-Specific Metadata

Product-specific metadata can be at both the granule level and the collection level. Product-specific metadata may (at the data provider's election) be contained in the Inventory Database tables, in which case it will be searchable by the system. There is also a provision to store product-specific metadata within granules that is available only when the granule has been ordered and delivered. This is termed archive metadata and is specified in a separate ODL group in the MCF.

In the granule metadata, the core attribute that is available to store product-specific metadata is called ParameterValue. The metadata describing this attribute is specified by the data provider through the AdditionalAttributes class at the collection-level. The units of measure, range, accuracy, and resolution for this are specified in the PhysicalParameterDetails class, also at the collection-level.

Product-specific metadata at the collection level is specified at the time the other collection level metadata attributes values are defined. At the granule-level, product-specific metadata is defined in the MCF. In both cases, a list of valid values and permitted ranges are specified in the ESDT data dictionary.

10.4 ESDT Maintenance

The ESDT Maintenance functionality is accomplished by using the ESDT Maintenance GUI. which provides the DAAC staff with functionality to view, update or remove installed ESDTs and to install new ESDTs.

The process of maintaining ESDTs will continue to rely on the ODL descriptors as the starting point. As part of an Update ESDT operation, changes to the descriptor will be propagated to the XML representation of the descriptor as well as an ESDT descriptor specific schema.

Figure 10.4-1 illustrates the ESDT descriptor files utilized in ECS and the components that generate them:

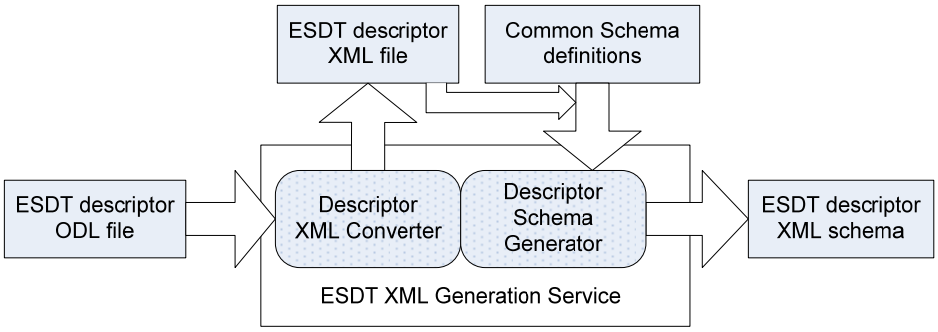


Figure 10.4-1. ESDT Descriptor File Transformations in ECS

The original ESDT descriptor ODL file is converted to its XML representation by the Descriptor XML Converter. This conversion occurs when an **Add ESDT** or **Update ESDT** or **View ESDT** process is selected from the ESDT Maintenance GUI.

The ESDT descriptor XML file is used to generate the descriptor XML schema. It is also copied into the Data Pool file system that corresponds to the ESDT, in the <ShortName.VersionId> directory and is available for anonymous FTP downloads.

The ESDT Descriptor XML file, together with a set of Common Schema Definitions file are used as input to the Descriptor Schema Generator which produces the ESDT descriptor XML schema. The schema is used for validating the granule XML metadata file. The Common Schema Definitions file contains definitions for all elements that are used by the supported ESDTs as well as the hierarchical relationships in which they can appear.

The Descriptor XML Converter and the Descriptor Schema Generator are part of the ESDT Descriptor XML Generation Service since they both produce descriptor related XML (the ESDT descriptor XML file and the ESDT specific schema respectively).

Figure 10.4-2 illustrates the high-level functionality flow that is provided by the ESDT Maintenance GUI for adding a new ESDT or updating an existing ESDT:

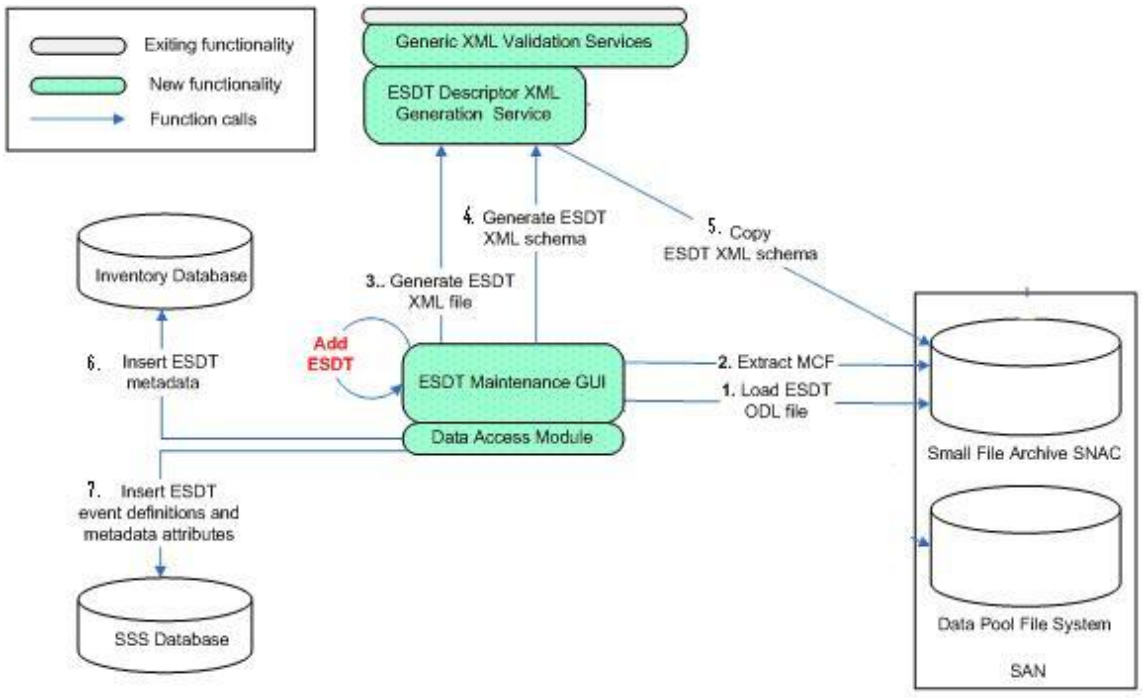


Figure 10.4-2. Adding/Updating an ESDT using the ESDT Maintenance GUI

The following functionality flow is used when an ESDT is added or updated using the ESDT Maintenance GUI.:

1. An ESDT ODL file from the location the user specified on the GUI is loaded. The directory will contain all the descriptor related files.
2. The Metadata Configuration File (MCF) is extracted from the descriptor ODL file and placed in the ESDT specific directory in the Small File Archive.
3. The ESDT descriptor XML file is generated. The ESDT Descriptor XML generation service contains the Descriptor XML Converter and the Descriptor Schema Generator modules.
4. The ESDT specific schema is generated, using the ESDT descriptor XML file produced in the previous step.
5. The generated ESDT descriptor XML schema is copied to the Small Archive File ESDT specific directory.
6. The ESDT collection metadata is inserted in the Inventory database.
7. The ESDT collection event definitions and metadata attributes that can be used to qualify subscriptions in the Spatial Subscription Server database are inserted.

Figure 10.4-3 illustrates the high-level functionality flow that is provided by the ESDT Maintenance GUI for removing an existing ESDT from the system:

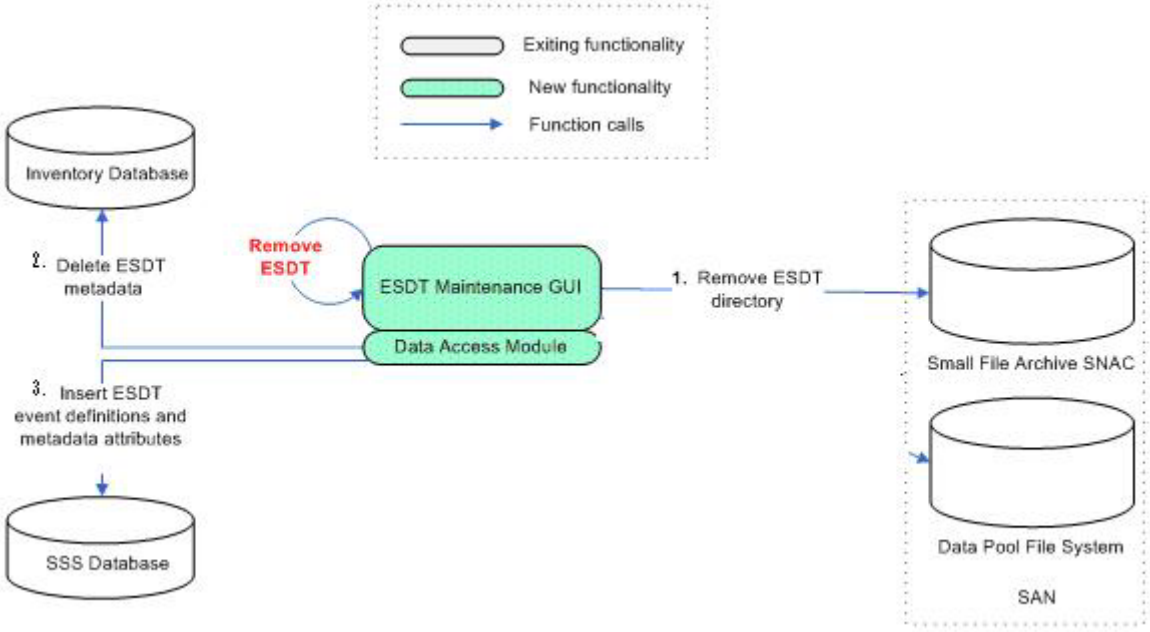


Figure 10.4-3. Removing an ESDT using the ESDT Maintenance GUI

The following functionality flow is used when an ESDT is removed:

1. ESDT specific files (ODL, MCF, XML schema, ESDT metadata directory <*shortname.VersionID*>) are removed.
2. The ESDT collection metadata from the Inventory database is deleted.
3. ESDT collection event definitions and metadata attributes from the Spatial Subscription Server database are deleted.

Note: Removal of an ESDT is not allowed if granules are present in the Inventory or DataPool. In addition, there can not be a Subscription on the ESDT within the Spatial Subscription Server. The appropriate Granule Deletion scripts must be run, if necessary, and all subscriptions removed before removing an ESDT.

Table 10.4-1 provides an activity Checklist for ESDT Maintenance.

Table 10.4-1. ESDT Maintenance - Activity Checklist

Order	Role	Task	Section
1	Database Admin	Launching the ESDT Maintenance GUI	(P) 10.4.1.1
2	Database Admin	Filter the ESDT List Page	(P) 10.4.1.2
3	Database Admin	View XML or ODL ESDT Descriptor Information	(P) 10.4.1.3
4	Database Admin	Re-generate an MCF or Schema	(P) 10.4.1.4
5	Database Admin	Remove an ESDT	(P) 10.4.1.5
6	Database Admin	Install/Update an ESDT	(P) 10.4.1.6
7	Database Admin	Update BMGT Configuration Files	(P) 10.4.1.7
8	Database Admin	Cleanup Failed ESDTs	(P) 10.4.1.8

10.4.1 Launching the ESDT Maintenance GUI

ESDT maintenance is accomplished by accessing the ESDT Maintenance GUI and is restricted to a single Database Username. This Username is configured in the ESDT Maintenance GUI Configuration file.

The ESDT Maintenance GUI will only allow for one authenticated session at a time. This is to prevent situations where multiple operators may perform conflicting actions. The time-out for authenticated sessions is configured in the Web application settings and is configurable via ECS Assist.

Note: The ESDT Maintenance GUI is configured to time out after 2 Minutes. You will need to log back into the GUI after each time out occurs.

10.4.1.1 Launching the ESDT Maintenance GUI

1 Access a terminal window logged in to a host (e.g., the Operations Workstation or Sun external server) that has access to the **Firefox** web browser.

- Examples of Linux external server host names include e4spl01 or n4spl01.

2 Type **firefox &** then press **Return/Enter**.

- It may be necessary to respond to dialogue boxes, especially if the browser is already being used by someone else who has logged in with the same user ID.
- The Mozilla Firefox web browser is displayed.

3 If a bookmark has been created for the **ESDT Maintenance GUI**, select the appropriate bookmark from those listed on the browser's Bookmarks pull-down window.

- The **Login:** prompt is displayed.

4 If no bookmark has been created for the **ESDT Maintenance GUI**, type **http://*host*:*port*** in the browser's **Location (Go To)** field then press **Return/Enter**.

- For example: http://f4dpl01.hitc.com:28000/ESDTMaint/

- The Login: prompt is displayed with the username configured for the GUI (see Figure 10.4-4)

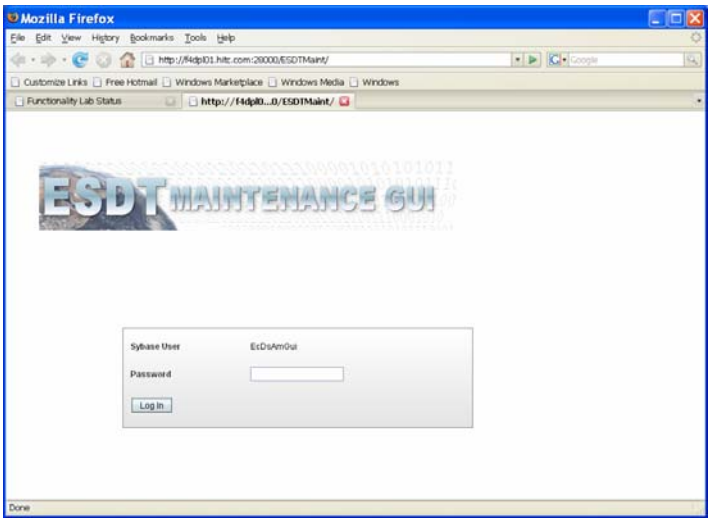


Figure 10.4-4. ESDT Maintenance GUI Log-in Screen

6 Type the appropriate password in the **Password** box of the security Login prompt.

7 Click on the **Log In** button:

- The **Installed ESDT** page is displayed (see Figure 10.4-5).

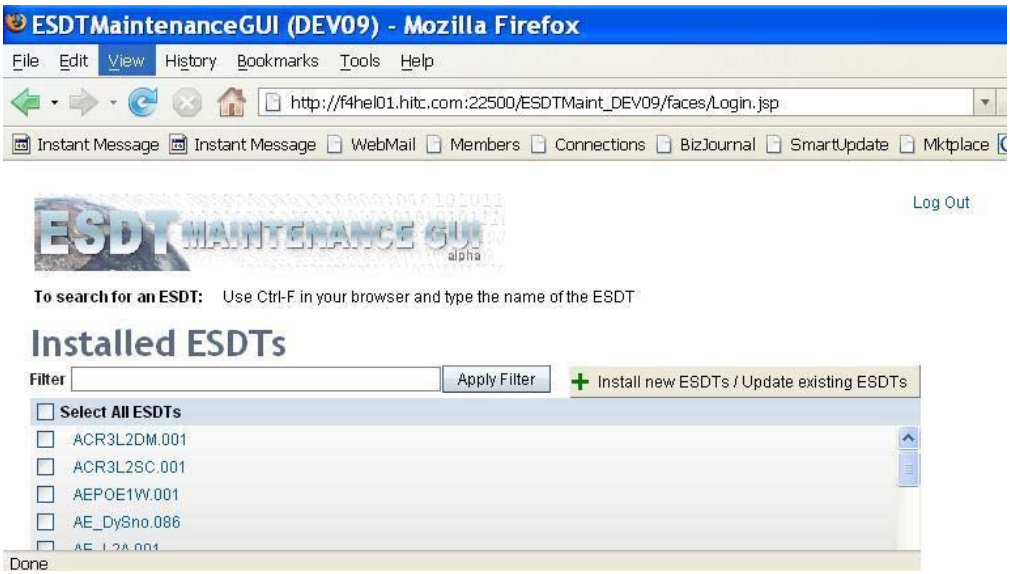


Figure 10.4-5. Installed ESDT Page

The ESDT List page lists all of the currently installed ESDTs. From this page, the operator can perform the following actions:

- Search for an ESDT by using the browser's built-in search function.
- View the ODL and XML descriptor information for a specific ESDT.
- Generate MCFs for one or more ESDTs.
- Generate Schemas for one or more ESDTs.
- Delete one or more ESDTs.
- Navigate to the ESDT installation/update page.

The ESDT List page includes a filter that can be applied to the list of ESDTs. This is useful for selecting particular types of ESDTs for bulk action (i.e., deletion, and MCF or ESDT Schema generation). This is a simple text search and will search ESDT Short Names. As shown in the example below, *MODIS* would return any ESDT with the MODIS anywhere in the name. The search is also case-insensitive.

10.4.1.2 Filter the ESDT List Page

1 Log in to the **ESDT Maintenance GUI**.
- The **Installed ESDT** page is displayed.

2 In the **Filter** box, enter the desired filter to be applied to the ESDT List**.**

3. Select **Apply Filter** button.
- The ESDT List will display the filtered ESDTs list as requested.

10.4.1.3 View XML or ODL ESDT Descriptor Information

1 Log in to the **ESDT Maintenance GUI**.
- The **Installed ESDT** page is displayed.

2 In the list of **Installed ESDT**, click on the desired ESDT name**.**
- The XML Descriptor Information page (see Figure 10.4-6) is displayed.

Figure 10.4-6. XML Descriptor Information Page

3. To display the Descriptor Information in ODL, click on the ODL button.
- The Descriptor Information page is displayed in ODL format.

During the initial install of an ESDT, the MCF and Schema files are generated from the descriptor file and copied to the Small Archive File ESDT specific directory. Certain situations (such as a corrupted MCF or Schema file) may require that these files be re-generated.

10.4.1.4 Re-generate an MCF or Schema

1 Log in to the **ESDT Maintenance GUI**.
- The **Installed ESDT** page is displayed (see Figure 10.4-5).

2 Select the ESDT(s) that require a re-generation of the MCF or Schema.
- A check mark is displayed in the box next to the selected ESDT.

3 Scroll to the bottom of the **Installed ESDT** list, click on the **Generate MCFs** or **Generate ESDT Schema** button.
- The MCF or Schema files stored in the Small File Archive will be used to re-generate the MCF or ESDT Schema.

Note: This action requires that the Data Pool Ingest Processing Service be restarted.

When an ESDT is removed, the following pre-conditions must be satisfied:

- All granules for this ESDT must not be present in the Inventory or DataPool. The Granule Deletion script must be run.
- The Data Pool collection for that ESDT must be removed using the Data Pool Maintenance GUI
- All subscriptions on the ESDT must be removed.

10.4.1.5 Remove an ESDT

1. Verify that Granules for the selected ESDT(s) have been removed from the Inventory and Data Pool.

2 Verify Subscriptions for the selected ESDT(s) have been removed.

3 Log in to the **ESDT Maintenance GUI**.

- The **Installed ESDT** page is displayed (see Figure 10.4-5).

4 Select the ESDT(s) that are to be deleted

- A check mark is displayed in the box next to the selected ESDT.

3 Scroll to the bottom of the **Installed ESDT** list, click on the **Delete selected ESDTs** button.

- The ESDT specific files (ODL, MCF, and XML schema) are removed.
- The ESDT Descriptor XML (ShortName, VersionID) directory on the file system where granules are stored is removed.
- The ESDT collection metadata from the Inventory database is deleted.
- The ESDT collection event definitions and metadata attributes from the SSS database is deleted.

Note: This action requires that the Data Pool Ingest Processing Service be restarted.

The operator can install a new or update an existing ESDTs from the ESDT Maintenance GUI.. On the List ESDT page of the ESDT Maintenance GUI, the operator selects the **Install new ESDTs/Update existing** button which displays a list **ESDTs to be installed.** The operator can review the file list and select the ESDTs to be installed or updated by checking the boxes for each ESDT. There are buttons to select the following descriptor files in the list: **all**, **none**, **installed**, **uninstalled** and **failed** ESDTs. Selection of these buttons will select all ESDTs in the category selected. Desired descriptors can be individually selected by clicking on the box next to the descriptor.

An operator performs installation or update on ESDTs by first selecting one, some, or all of the Descriptor files. Then the **Proceed with installation/update** button is used to perform installation or an update on the selected Descriptor file name. The column on the right contains the current status of an ESDT.

If the installation or update completes successfully for all ESDTs, the installation files will be removed from this list, and a message will be displayed at the top of the screen indicating the success.

If the installation or update did not succeed for one or more ESDTs, a general error message will be displayed at the top of the screen. A table at the top displays detailed error information next to each ESDT that failed.

If an error is encountered during the installation or update (e.g., a validation error), the installation for that particular ESDT will fail. Installation of the other ESDTs will continue processing until the selected list is completed. As ESDTs are successfully installed or updated, the descriptor files are removed from the installation source directory. Any remaining files in the list would be those that could not be installed due to an error or those that were not selected for processing.

Note: In order for products associated with this ESDT to be exported to ECHO, the following BMGT config files need to be updated:

- EcBmBMGTGroup.xml
- EcBmBMGTSpatialEsdts.xml

10.4.1.6 Install/Update an ESDT

1 Log on to the host where the **ESDT Maintenance GUI** is installed (e.g**., x4dpl01**).

- The **ESDT List** page is displayed.

2 Select the **Install new ESDTs/Update existing ESDTs** button.

- The **ESDTs to be Installed, Updated, or that have Failed** page is displayed (see Figure 10.4-7.

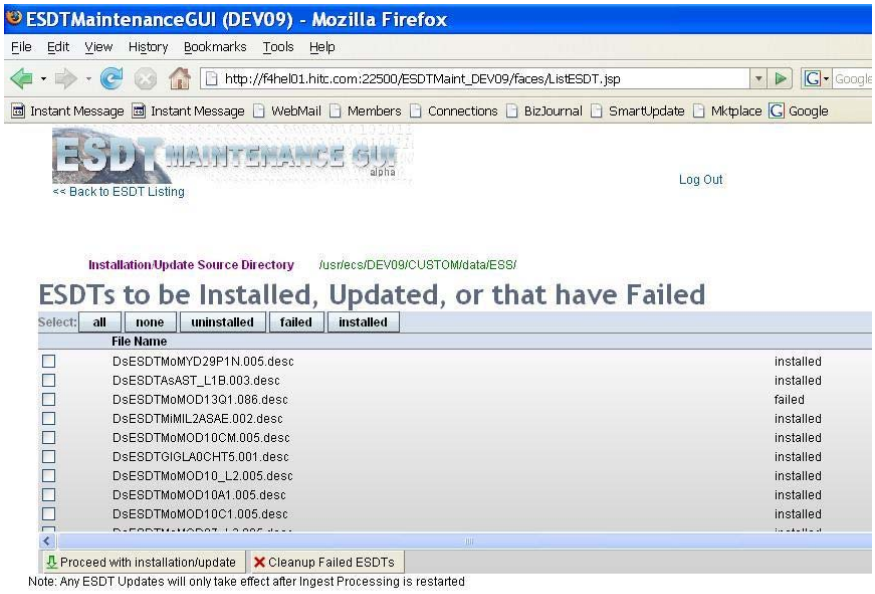


Figure 10.4-7. ESDTs to be Installed, Updated, or that Have Failed Page

3 Click on the box(es) next to the desired descriptor file(s).

- A check is displayed in the box.

Note: The five categories displayed above the list of descriptor files can be used if applicable (i.e. **all** - if you want all descriptor files selected; **uninstalled** - if you want all

uninstalled descriptor files selected; **failed** - if you want all failed descriptor files selected; **Installed** - if you want all installed descriptor files selected).

4 Select the **Proceed with installation/update** button.

- A message is displayed (see Figure 10.4-8) indicating the number of descriptors successfully installed and the installation files will be removed from the install list.
- If the installation is not successful, a message is displayed at the top of the page, indicating the number of descriptors that failed to be installed along with the associated error.

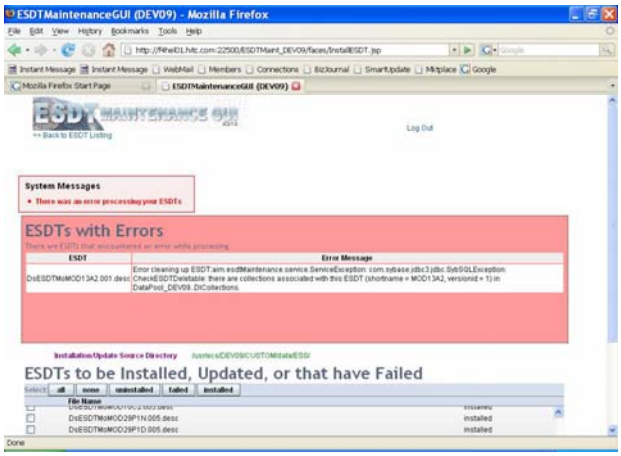


Figure 10.4-8. ESDTs Failure Screen

Note: This action requires that the Data Pool Ingest Processing Service be restarted.

10.4.1.7 Update BMGT Config Files

1 Log on to the host where BMGT is installed (e.g.**, x4oml01**).

2 Type the following:

- cd /usr/ecs/OPS/CUSTOM/cfg

3 Edit the EcBmBMGTGroup.xml by entering the following edit commands:

- **vi EcBmBMGTGroup.xml**

4 Following **<name>groupName</name>**, enter the following information for the ESDT added in procedure 10.4.1.6:

<ESDT>
<*ShortName*>**short name of the ESDT**</*ShortName*>
<*VersionID*>**VersionId of the ESDT**</*VersionID*>
<*CollExport*>**Y**</*CollExport*>
<*GranExport*>**Y**</*GranExport*>
</*ESDT*>

5 Exit the editor by typing **Ctrl Z.**

6 If the EcBmBMGTSpatialEsdts.xml has not been already configured for the ShortName of the newly installed ESDT, the following 4 lines need to be added by entering the following edit commands:

- **vi EcBmBMGTSpatialEsdts.xml**

7 Following **</spatialesdts>**, enter the following information for the ESDT added procedure in procedure 10.4.1.6:

<SP_ESDT>
*<ShortName>***ShortName of the Esdt** *</ShortName>*
*<SpatialRep>***SpatialRep as decided by the science team***</SpatialRep>*
</SP_ESDT>

8 Exit the editor by typing **Ctrl Z.**

If an error is encountered during the installation or update (e.g., a validation error), the installation for that particular ESDT will fail. Installation of the other ESDTs will continue processing until the selected list is completed. As ESDTs are successfully installed or updated, the descriptor files are removed from the installation source directory. Any remaining files in the list would be those that could not be installed due to an error or those that were not selected for processing. In cases when fatal error has occurred, the ESDT will be marked as failed in the list of **ESDTs to be Installed**. After reviewing the error, the operator will be able to initiate recovery for the failed ESDT by using the **Cleanup Failed ESDT**s command.

10.4.1.8 Cleanup Failed ESDTs

1 Log on to the host where the **ESDT Maintenance GUI** is installed (e.g**., x4hel01**).

- The **ESDT List** page is displayed.

2 Select the **Install new ESDTs/Update existing ESDTs** button.

- The **ESDTs to be Installed** page is displayed.

3 Click on the box(es) next to the desired descriptor file(s) to be recovered.

- A check is displayed in the box.

4 Select the **Cleanup Failed ESDTs** button.

- For each ESDT selected, (i.e., incomplete installation), any Descriptors, MCFs, and Schema present in the Small File Archive is removed.
- The ESDT is removed from the Inventory Database.
- The temporary backup descriptors, MCFs, and schema files are restored and information from the restored descriptor file is place in the Inventory Database.

This page intentionally left blank.

11. Bulk Metadata Generation Tool

11.1 BMGT Overview

The Bulk Metadata Generation Tool (BMGT) is an ECS component that is used to generate an external representation of the ECS metadata holdings. This external representation consists of a number of distinct data products that describes both the current state of the metadata holdings, as well as changes to that state (such as the insert, update, and deletion of collections and granules).

While the data products produced by the BMGT are generally ingested back into the ECS system so that value added providers may search, and order (or subscribe to) these products, the BMGT is also responsible for directly exporting the products to the EOS Clearing House (ECHO).

In general use, the BMGT is designed to be fully automatic. Running periodically, (with a frequency of once per hour up to once per day), BMGT will automatically generate the required products and export them to ECHO, as well as make them available for archiving via Data Pool Ingest. Additional BMGT tasks may be initiated as a result of other actions, such as Data Pool Cleanup. Alternatively, the operator may explicitly request BMGT to generate one or more products based on collection and/or granule selection criteria.

The metadata files sent to ECHO will be formatted in XML and will be compressed/consolidated into a single file for delivery using the UNIX compression utility zip. Currently, the exported files include the following:

- Collection metadata (ECSMETC-) files following the BMGTCollectionMetadata.dtd
- Granule metadata (ECSMETG-) files following the BMGTGranuleMetadata.dtd
- Granule QA update (ECSMETU-)files following the BMGTUpdateMetadata.dtd
- Bulk Browse (ECSBBR-) data files following the BMGTBrowseMetadata.dtd
- Bulk URL (ECSMETU-) files following the BMGTUpdateMetadata.dtd
- ECS browse files referenced in the ECSBBR files.

The following diagram (see Figure 11.1-1) shows the high level context in which the BMGT operates.

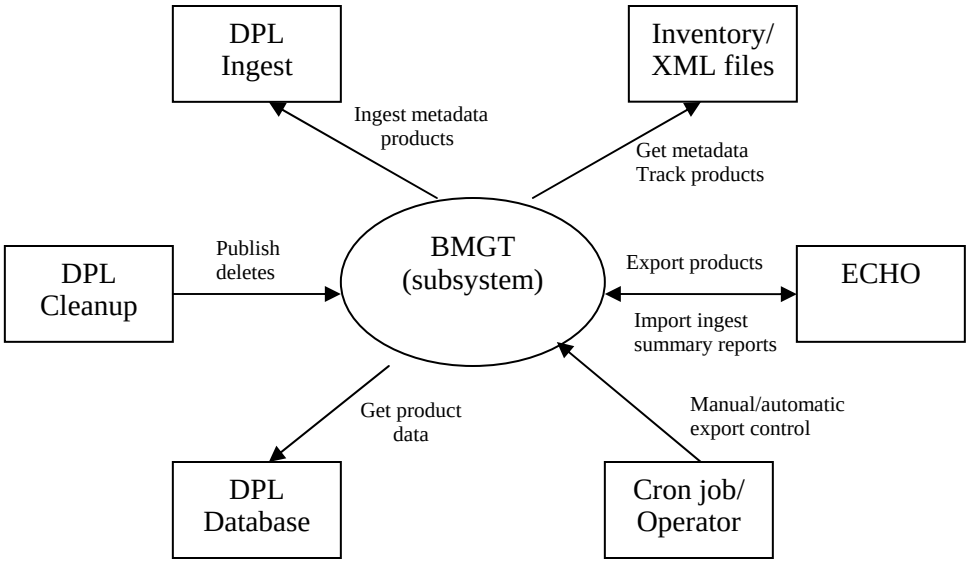


Figure 11.1-1. BMGT Context diagram

A BMGT export cycle can be initiated in one of three ways.

- Automatically: Based on its configuration and cron job setup, BMGT itself decides that it is time to initiate an export cycle. The Automatic Export process is responsible for selecting which Automatic export to run and populating export cycles to the BMGT database.
- Manually: This happens when the operator explicitly tells the BMGT to initiate an export cycle. This is handled by the Manual Export Process, which provides a large number of options for generating the export package. The Manual Export process is responsible for verifying that a manual cycle can be run and initiating the manual export generation.
- Cleanup: When Data Pool Cleanup is run, it will trigger the BMGT to produce an export package to ECHO. The idea here is to notify ECHO as quickly as possible of the removal of granules from the Data Pool. The Data Pool Cleanup Script has been modified to create a Cleanup Export Cycle at the end of its execution. BMGT Monitor will automatically poll the BMGT database looking for the Cleanup Export Cycle. BMGT will automatically export the package to ECHO. The frequency of the

polling is defaulted to 30 seconds but can be changed via configuration parameters found using the BMGT GUI. Data Pool Cleanup will not be permitted to actually remove any granules until the BMGT Monitor has picked up and initiated the requested Cleanup Export Cycle.

NOTE: There are collection level metadata values that cannot be automatically updated in ECHO. They include but are not limited to:

- Spatial search type - granule spatial representation
- Short Name/Version
- Long Name

Modifying the above collection level metadata values will require ECHO to drop the collection which means all granules in the ECHO inventory for that collection will be need to be deleted. In this circumstance, all historical granules for that collection will have to be re-exported to ECHO.

Table 11.1-1 provides an activity Checklist for BMGT.

Table 11.1-1. BMGT - Activity Checklist

Order	Role	Task	Section
1	Archive Technician	Launching the BMGT GUI	(P) 11.2.1.1
2	Archive Technician	Monitoring Recent Packages	(P) 11.2.2.1
3	Archive Technician	Cancelling Recent Packages	(P) 11.2.3.1
4	Archive Technician	Reviewing Failed Packages	(P) 11.2.4.1
5	Archive Technician	Changing Configuration Parameters	(P) 11.2.5.1
6	Archive Technician	BMGT Manual Mode	(P) 11.3.1
7	Archive Technician	BMGT Automatic Mode	(P) 11.4.1

11.2 BMGT GUI

The BMGT GUI allows the operator to monitor the export of BMGT packages (Automatic, Manual, and Cleanup). The primary purpose of the GUI is to provide the operator with a list of recent packages and their status. In addition, the operator will use it to configure various BMGT tuning parameters, such as the length of an Automatic cycle and the availability of the FTP service. Since it is possible for errors to occur during the FTP process, the third function of the GUI is to display any global FTP alerts.

11.2.1 BMGT GUI Functions

After a successful login, the user is presented with a navigation panel on the left-hand side of the screen, consisting of the following items:

- Home Page
- Monitoring
    - Recent Packages

    - Recent Failed Packages Only
- Configuration
    - Global Tuning

The GUI provides DAAC staff with the following functions:

- Display BMGT export processes that are currently in progress
- Monitor the status of the BMGT FTP service that exports products to ECHO
- Allows the operator to suspend/resume FTP of products to ECHO
- List the *N* most recent export packages and view detail information about them, where *N* is configurable by the DAAC staff
- Cancel an export package that is currently being transmitted to ECHO or waiting for transmission
- List the *N* most recently completed packages which resulted in errors and view detail information about them
- View and change BMGT configuration parameters, except for configuration items such as collection group/collection mapping that must be specified in XML configuration files. Changing the BMGT runtime configuration parameters will be restricted to DAAC staff that is logged in as BMGT administrator
- Display global alerts upon a configured number of BMGT FTP to ECHO failures

11.2.1.1 Launching the BMGT GUI

1 Access a terminal window logged in to a host (e.g., the Operations Workstation or Sun external server) that has access to the Mozilla Firefox web browser.

- Examples of Linux external server host names include e4dpl01, l4dpl01 or n4dpl01.

2 Type **firefox &** then press **Return/Enter**.

- It may be necessary to respond to dialogue boxes, especially if the browser is already being used by someone else who has logged in with the same user ID.
- The Mozilla Firefox web browser is displayed.

3 If a bookmark has been created for the **BMGT GUI**, select the appropriate bookmark from those listed on the browser's Bookmarks pull-down window.

- The **Login:** prompt is displayed.
- The Login page (see Figure 11.2-1) allows the operator to log in, either as an Administrator (with the ability to configure global tuning parameters) or a read-only Operator. The Administrator login requires a password, while the Operator login does not.

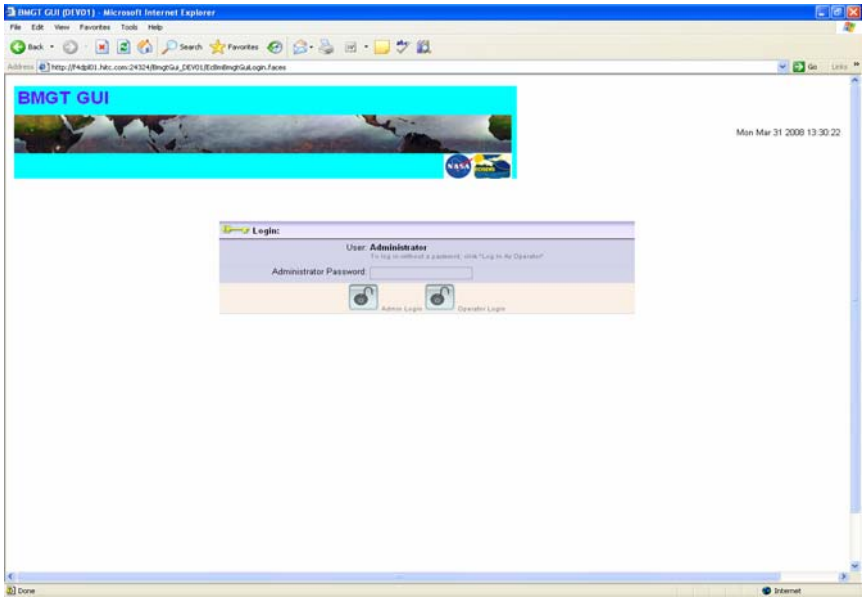


Figure 11.2-1. BMGT Login Page

4 If no bookmark has been created for the **BMGT GUI**, enter the URL in the Address window and click on the **Go** or press the **Return/Enter** button.

- For example: http://x4dpl01.hitc.com: 24320/BmgtGui/EcBmBmgtGuiLogin.faces.
- The Login: prompt is displayed.

5 If you are logging in as the **User: Administrator**, enter the appropriate password in the **Administrator Password** box.

6 Click on the **Admin Login** button.

- The **BMGT GUI Home** page (see Figure 11.2-2) is displayed.

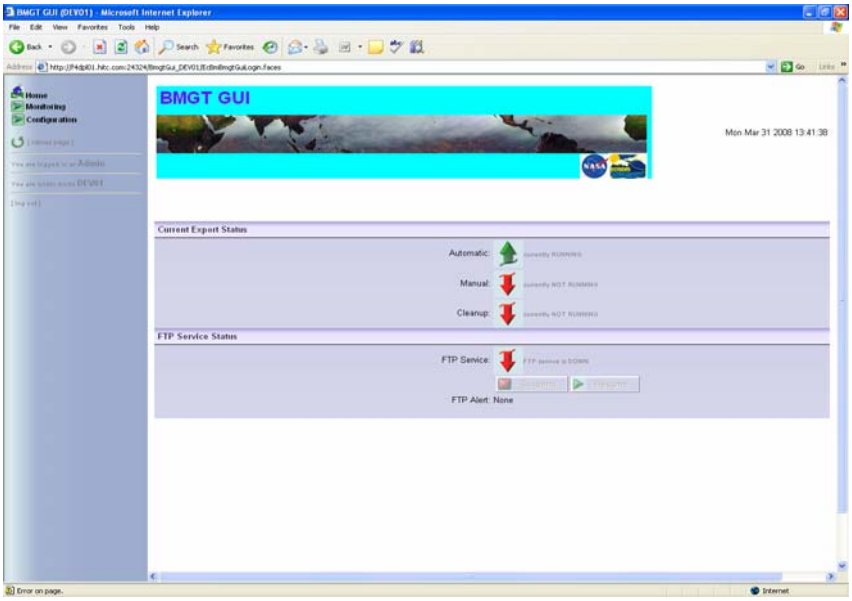


Figure 11.2-2. BMGT GUI Home Page

7 If you are logging in as an operator no password is required, just click on the **Operator Login** button**.**

- The **BMGT GUI Home** (read only) page is displayed.

The **BMGT GUI Home** Page provides an overview of the current system status, including any global alerts.

The first section shows whether or not the **Automatic**, **Manual**, or **Cleanup** mode is currently running.

The second section displays the current state of the BMGT FTP service. There may be three states for this service:

- The FTP service is down. This state corresponds to a red arrow pointing down. In this case, the Suspend and Resume buttons are both disabled.
- The FTP service is up and active. This state corresponds to a green arrow pointing up and a green "Active" light; the operator is allowed to manually suspend the FTP service.
- The FTP service is up and suspended by operator. This state corresponds to a green arrow pointing up and a red "Suspended" light; the operator can manually resume the service.

Additionally, this section displays the existence or absence of a global FTP alert. A single alert may be pending due to FTP errors; in this case, the FTP Alert line will show the alert description.

11.2.2 Monitoring Recent Packages

The **Recent Packages** page provides a listing of *N* most recent packages and their status (the number is configurable on the Global Tuning page).

The listing consists of the following columns:

- **Cycle ID**: A unique cycle ID. (Clicking on the underlined link will bring up the Package Details screen, discussed below.)
- **Package ID**: A unique package ID.
- **Export Type**: Automatic, Manual, or Cleanup, corresponding to the type of the cycle in which the package was generated.
- **Status**: The current status of the package, with the values defined in S_BGT_01250.
- **Last Status Update:** The date an time of the last change in the status of the package.
- **Coverage From:** The initial time point covered by the package.
- **Coverage To:** The last time point covered by the package.

The **Cancel Package** button allows the operator to select individual packages and cancel them if they are not yet in a terminal state. The cancellation process applies to FTP transmission only (rather than product generation). The **Cancel Package** button cancels all packages whose checkboxes are currently selected; the checkboxes appear only in those cases when the package can be cancelled.

11.2.2.1 Monitoring Recent Packages

1 Login to the BMGT GUI

- The **BMGT GUI Home** page is displayed.

2 Click on the **Monitoring** link from the navigation panel.

3 Select **Recent Packages** from the navigation panel.

- The **Recent Packages** page (see Figure 11.2-3) is displayed.

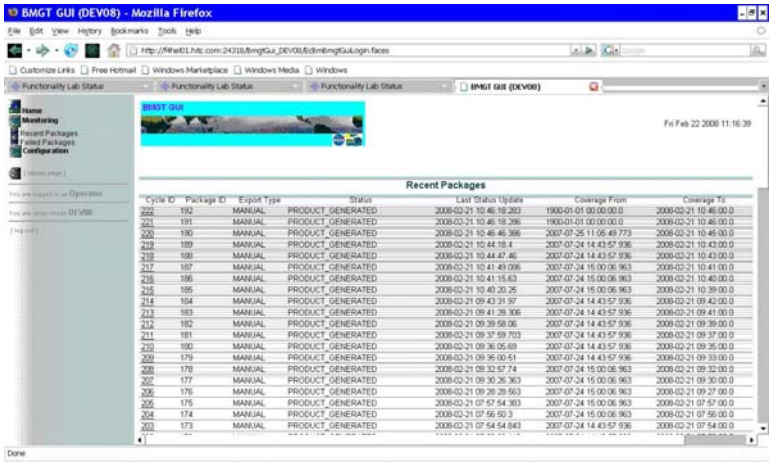


Figure 11.2-3. Recent Package Page

4. To see detailed information about a given package, click on the desired **Cycle ID *xxx.***

- The **Package Details: Package** *xxx* (see Figure 11.2-4) is displayed.

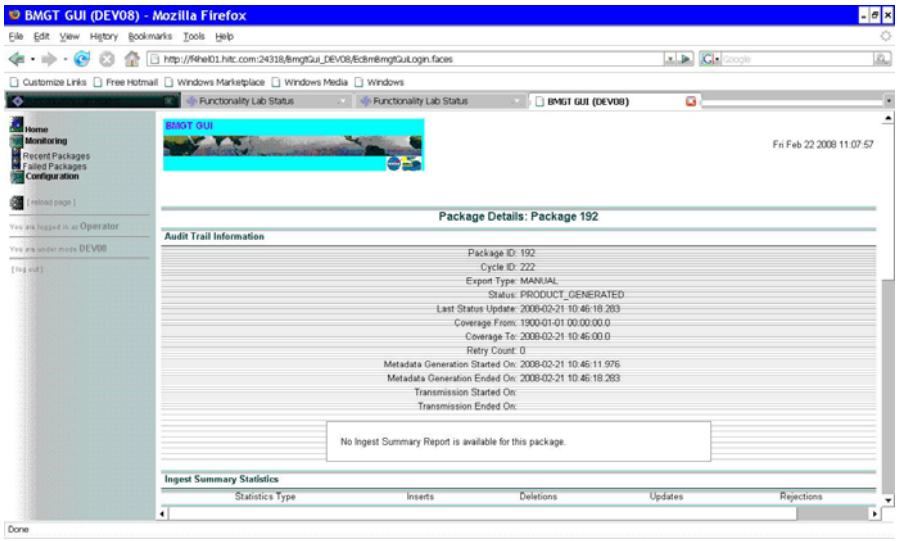


Figure 11.2-4. Package Details Page

- The Package Details page contains **Audit Trail Information**, **Ingest Summary Statistics** (if available) and **Product Information**.

11.2.3 Cancelling Recent Packages

From the **Recent Packages** page, the **Cancel Package** button allows the operator to select individual packages and cancel them if they are not yet in a terminal state. The cancellation process applies to FTP transmission only (rather than product generation). The **Cancel Package** button cancels all packages whose checkboxes are currently selected; the checkboxes appear only in those cases when the package can be cancelled.

11.2.3.1 Cancelling Recent Packages

1 Login to the BMGT GUI.

- The **BMGT GUI Home** page is displayed.

2 Click on the **Monitoring** link from the navigation panel.

3 Select **Recent Packages** from the navigation panel.

- The **Recent Packages** page is displayed.

4 Click on the box next to "non-terminal" packages to be cancelled.

- A check is placed in the box.

5 Click on the **Cancel Package** button.

- The status of the selected package(s) becomes CANCELING, and, upon successfully canceling, the status will be changed to CANCELLED.

11.2.4 Reviewing Failed Packages

The Failed Packages page shows a listing of *N* most recent packages which resulted in an error. The list columns are identical to those on the general Recent Packages page. The following additional detailed information is accessible from the Package Details page and can be viewed by clicking the underlined link of a failed package.

This detailed information includes:

- A summary of general package information, as presented on the Monitoring screens
- Information about the package's Ingest Summary Report, including the download link for the report
- The contents of a package, broken down by Product Type and Group
    - **Browse**: Multiple groups allowed per package; Inserts/Updates/Deletes applicable
    - **Granule**: Multiple groups allowed per package; Inserts/Updates/Deletes applicable

- **Collection**: Multiple groups allowed per package; Inserts/Updates/Deletes applicable
- **QA**: Multiple groups allowed per package; Inserts/Updates/Deletes applicable
- **Valids**: Only one allowed per package (not group-based); Inserts only
- **BulkURL**: Only one allowed per package (not group-based; Inserts/Deletes/Updates applicable)

11.2.4.1 Reviewing Failed Packages

1 Login to the BMGT GUI.

- The **BMGT Home** page is displayed.

2 Click on the **Monitoring** link from the navigation panel.

3 Select **Failed Packages** from the navigation panel.

- The **Failed Package** page (see Figure 11.2-5) is displayed.

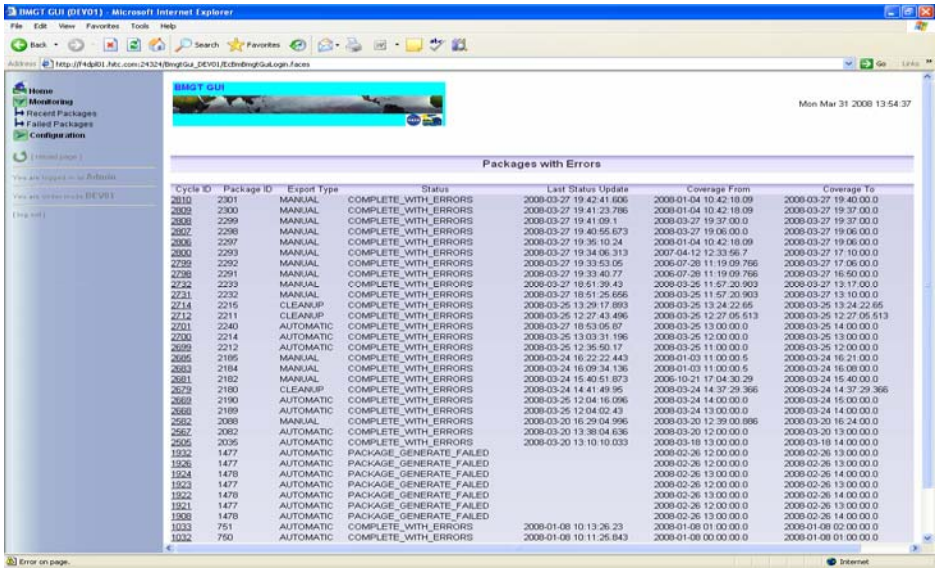


Figure 11.2-5. Package with Errors Page

4 Click on the underscored **Cycle ID** link.

- The **Package Details** (see Figure 11.2-6) is displayed.

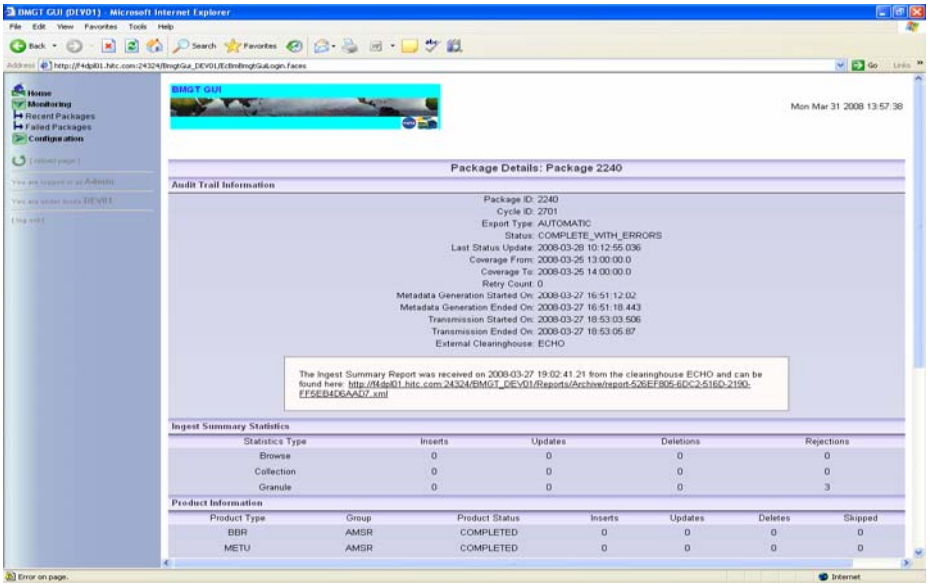


Figure 11.2-6. Package Details Page

11.2.5 Global Tuning Parameters

The Global Tuning page lists various BMGT configuration parameters. The list is a three-column table with the parameter name, description, and value. Clicking the **Apply Changes** button saves the current (possibly changed) values of all the parameters; the Reset button reverts to the default values. Checkboxes next to each value prevent accidental modifications.

The fields on the **Global Tuning** page are enabled only if the current user is the Administrator, and include the ability to change Admin Password. The fields are disabled if the current user is not the Administrator. Table 11.2-1 contains a description of the parameters that can be updated using the BMGT GUI.

Table 11.2-1. BMGT Configuration/Global Parameters (1 of 3)

Parameter Name	Description	Default Value
ADMIN_PASSWORD	The BMGT GUI administrator password. Note that this is stored in the database in encrypted form. When the password is changed on the BMGT GUI, the GUI will automatically encrypt the password before storing it.	xxxxxxx
FTP_PASSWORD	The encrypted password that will be used to authenticate the log in to the ECHO host. The BMGT does not need to be restarted for changes to this value to take effect	xxxxxxxx
AUTOMATIC_CYCLE_LENGTH_HRS	The length of the currently configured automatic export cycle, measured in hours. The BMGT does not need to be restarted if this value is changed, but note that the new value will not apply until the next day. Valid values are 1,2,3,4,6,8,12,24.	24
AUTOMATIC_CYCLE_RETRY_INTERVAL_MINS	The time interval, measured in minutes, between retries of a failed automatic export cycle. Recommend values in the range 30 to 60 minutes.	60
BMGT_PDR_POLLING_DIRECTORY	The DPL Ingest polling directory into which BMGT PDRs will be placed.	
BMGT_PDR_POLLING_HOST	The fully qualified host name where the DPL Ingest polling location is configured.	
CLEANUP_OLD_CYCLES_DAYS	Number of days before a package's audit trail information can be cleaned up.	10
DATABASE_RETRY_COUNT	The number of attempts that should be made to execute a database command.	5
DATABASE_RETRY_INTERVAL_SECS	The time, measured in seconds, between retries of a database command.	30
DATA_CENTER_ID	Value to use in generated METG,BBR xml for the DataCenterId value	
DESC_FILE_DIR	The directory where ESDT descriptor files are located.	
DISPLAY_MAX_PACKAGES	Determines how many recent packages will be displayed on the GUI Monitoring page.	100
DTD_LOC	The DTD host and port. This is the root URL where all of the DTDs can be found. The DTD file name will be appended after this value.	
EMAIL_HOST	The SMTP mail server full qualified host name that will be used to send emails.	

Table 11.2-1. BMGT Configuration/Global Parameters (2 of 3)

Parameter Name	Description	Default Value
FTP_HOST_NAME	The name of the ECHO host to which export packages will be pushed, and Ingest Summary Reports will be pulled. This may be either a hostname, or an IP address. The BMGT does not need to be restarted for changes to this value to take effect.	
FTP_PULL_DIRECTORY	The directory on the ECHO host from which Ingest Summary reports will be pulled. The BMGT does not need to be restarted for changes to this value to take effect.	
FTP_PUSH_DIRECTORY	The directory on the ECHO host into which the package files will be placed. The BMGT does not need to be restarted for changes to this value to take effect.	
FTP_RETRY_INTERVAL_MINS	The time interval, measured in minutes, between retries of a failed FTP export operation.	5
FTP_USERNAME	The user name that will be used to log in to the ECHO host. The BMGT does not need to be restarted for changes to this value to take effect.	
GENERAL_PKG_FILE	Absolute path for GENERAL Package file name, used to retrieve CollectionPackage information.	
GENERATOR_CHECK_INTERVAL_SECS	Determines how frequently the BMGT checks the database for new packages to generate. Recommend values in range 30 to 300 seconds.	30
GROUPS_CONFIG_FILE	The absolute path of the ESDT group configuration file.	
INGEST_SUMMARY_RPT_ARCHIVE	The directory on the local host in which the ECHO Ingest Summary Reports will be archived.	
INGEST_SUMMARY_RPT_DIR	The temporary directory on the local host into which the BMGT will place ECHO Ingest Summary Report files for processing.	
INGEST_SUMMARY_RPT_URL	The URL where the ECHO Ingest Summary Report files can be downloaded.	
MAX_DATA_SKIPPED	The maximum number of data-related errors that the BMGT may encounter when generating an export package before the package will fail.	10
MAX_FTP_PACKAGE_INTERVAL_HRS	The maximum number of hours that may pass before a warning email is sent if an export package has not started transferring to ECHO.	12
MAX_SIZE_ECSBBR	The maximum number of browse inserts/deletes allowed for ECSBBR files. Export products larger than this will have their output split into multiple files.	200
MAX_SIZE_ECSMETG_KB	The maximum size for ECSMETG files, measure in KB. Export products larger than this will have their output split into multiple files.	

Table 11.2-1. BMGT Configuration/Global Parameters (3 of 3)

Parameter Name	Description	Default Value
MAX_SIZE_ECSMETU	The maximum number of updated granules that may be allowed per ECSMETU file. Export products larger than this will have their output split into multiple files.	10000
MAX_WAIT_FOR_INGEST_REPORT_HRS	Maximum number of hours to wait for an Ingest Summary Report from ECHO before issuing a warning email to the DAAC operator that the expected report has not arrived.	72
MISR_PROCESSING	Indicates whether MISR processing is enabled. Reserved for use by ASDC. Do not change this configuration parameter while the system is running.	Y
MONITOR_CHECK_INTERVAL_SECS	This determines how frequently the BMGT checks for completed export packages. Recommend values in the range 60 to 300 seconds. This will also determine how often BMGT checks for Cleanup Export Packages requested by DataPool Cleanup.	120
NUM_RETRIES_FOR_ALERT	The number of ECHO FTP retries that will trigger an alert.	5
PACKAGER_RETRY_INTERVAL_MINS	The time interval, measured in minutes, between retries of a failed attempt to package up the export product files.	60
PRODUCT_ROOT_DIRECTORY	The root directory under which the temporary package directories will be created. These are used to store the product/package files for ingest or export.	
SPATIAL_ESDT_FILE	The absolute path for the BMGT Spatial ESDTs' configuration file	
SPECIAL_CASE_FILE	The absolute path for the special case file name, used to retrieve cost estimate information for collections.	
STYLESHEET_DIR	The absolute path for the location of Collection and Granule style sheets.	
TEMP_DESC_DIR	Temporary directory for writing XML descriptor files to retrieve collection metadata.	

11.2.5.1 Changing Configuration Parameters

1 Login to the BMGT GUI as the system administrator.

- The BMGT **GUI Home** page is displayed.

2 Click on the **Configuration** link from the navigation panel.

3 Select **Global Tuning** from the navigation panel.

- The Global Tuning page (see Figure 11.2-7) is displayed.

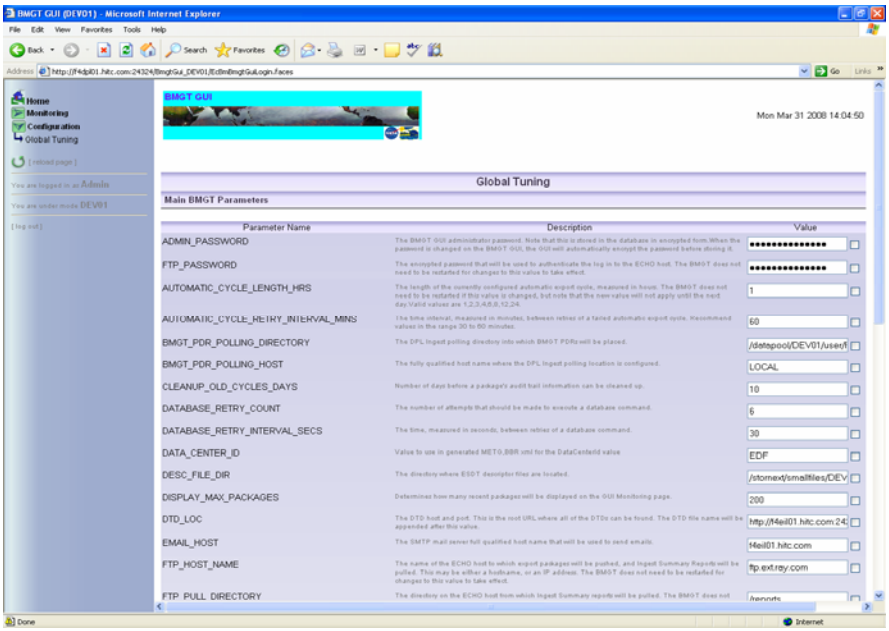


Figure 11.2-7. Global Tuning Page

4 Click on the **Value** box of the parameter to be changed.

- A flashing input curser is displayed.

5 Enter the desired parameter update.

- Change is displayed in the **Value** box.

6 Click on the checkbox next to the **Value** box.

- A check is placed in the checkbox.

7 Scroll to the bottom of the Global Parameter page and select the **Apply Changes** button.

- Changes will be saved for the parameters which have had their checkboxes checked. Unintentional changes to other parameters will not be saved.

Note: Most configuration changes made through the Global Tuning Page do not take effect until all BMGT servers are re-started. The exception is the **FTP_PUSH_USERNAME** and **FTP_PUSH_PASSWORD** input which are applied when the **Apply Change** button selected.

11.3 BMGT Manual Mode

The BMGT manual preprocessor provides an interface through which the operator can initiate an export of ECS metadata through BMGT. Unlike a normal 'AUTOMATIC' instantiation of BMGT, which exports metadata in response to changes, or 'events', a 'MANUAL' BMGT run

will export the current metadata for an operator provided set of granules and collections. An operator is able to specify which metadata products are desired, rather than retrieving all of them. An operator is also able to use the Manual Preprocessor to re-run a previous AUTOMATIC export that has failed. Once the Manual Preprocessor is run, the desired products will be created by the BMGT Generator server. These products can be exported to ECHO and/or ingested into ECS, or neither of these depending on what that operator specifies. The Manual Preprocessor is meant primarily for reconciling ECS and ECHO metadata, or for other situations where the normal, automatic export of BMGT metadata is not sufficient. The Manual Preprocessor does not prevent the Operator from exporting duplicate metadata to ECHO. The Operator is responsible for specifying options carefully to minimize the risk and for consulting with ECHO when exporting data outside of the normal sequence.

Note: For the purposes of this section, "Inventory database" refers to either the Science Data server database, or the Inventory database. With this in mind, this section is applicable to ECS releases 7.20+ and 7.21.

Table 11.3-1 contains a listing/description of the arguments used by the Manual Export Script.

Table 11.3-1. Manual Export Command Line Arguments (1 of 3)

Option	Notes	Description
GENERAL OPTIONS		
--help | -h	Overrides all other options	Display a detailed help page.
--mode <MODE>	Required	Run in ECS mode <MODE>.
PRODUCT OPTIONS		
--metg	Requires one or more SELECTION CRITERIA options	Generate an ECSMETG(granule metadata) product. URL and visibility products will also be generated as well where appropriate (ie. **--url** is implied and does not need to be specified explicitly).
--metv	Requires **--noexport**	Generate an ECSMETV(valids) product.
--metc	requires **--collections** or **--collectionfile**	Generate an ECSMETC(collection metadata) product.
--bbr	requires one or more SELECTION CRITERIA options	Generate an ECSBBR(browse) product. The BBR product generated will contain any browse granules explicitly specified by the SELECTION CRITERIA options, as well as browse files associated with any granules specified by those options. Browse linkages to science files will also be generated. If a METG is being generated for an associated science granule, it will include the linkage, otherwise the linkage will be recorded in a METU file.
--url	requires one or more SELECTION CRITERIA options	Generate a BULKURL(DataPool public URL) product.

Table 11.3-1. Manual Export Command Line Arguments (2 of 3)

Option	Notes	Description
RUN TYPE		
--delete	requires one or more SELECTION CRITERIA options	Generate deletion metadata. If this option is omitted, insertion metadata will be generated. Granules and collections being processed in a deletion cycle must be either physically or logically deleted. Similarly, granules and collections specified for a normal insert cycle must currently exist in ECS. If a granule is physically deleted from the archive, it must be explicitly specified (with the **--granules** or **--granulefile** option) by geoid rather than dbid.
SELECTION CRITERIA		
--collections | –c <shortname.versionID>[,<shortname.versionID>,...]		Generate metadata for collection <shortname.versionID>. Multiple collections can be specified, separated by a comma and no space.
--collectionfile | --cf <filename>		Same as **--collections**, but specifies a file which contains one or more collections. The collections can be on one or multiple lines and must be separated either by newlines or whitespace.
--granules | -g <ID>[,<ID>,...]		Where <ID> is either a dbid or a geoid in the form: **<SC/BR>:<SHORTNAME>.<VERSIONID>:<DBID>** Generate metadata for the listed granules. Multiple granules can be specified, separated by a comma and no space.
--granulefile | --gf <filename>		Same as **--granules**, but specifies a file which contains one or more dbids and/or geoids. The ids can be on one or more lines and must be separated either by newlines or whitespace.
--starttime | --st <datetime>	requires **--collectionfile** or **--collections**	Defines the starting time (inclusive) of a datetime range for which to generate granule metadata. This parameter is used only if **--collection**, or **--collectionfile** is specified. It will be used to select a subset of granules from the specified collection(s) for which metadata will be generated. <datetime> should be in the format "YYYY-MM-DD HH:MM:SS" [quotes are required].
--endtime | --et <datetime>	requires **--collectionfile** or **--collections**	Defines the end time (non-inclusive) of a datetime range for which to generate granule metadata. This parameter is used only if **--collection**, or **--collectionfile** is specified. It will be used to select a subset of granules from the specified collection(s) for which metadata will be generated. <datetime> should be in the format "YYYY-MM-DD HH:MM:SS" [quotes are required].

Table 11.3-1. Manual Export Command Line Arguments (3 of 3)

Option	Notes	Description
OUTPUT OPTIONS		
--noexport | --ne	implies **--nosequence**	Do not export the generated package to ECHO, and do not assign it a sequence number.
--noarchive | --na		Do not archive the generated package into ECS.
--nosequence | --ns		Generated package should not be assigned a sequence number. This is automatically implied when **--noexport** is specified.
--daacstring | -d		A string up to 40 characters long and consisting only of valid Unix file name characters (excluding period) to be included as part of the file names in the metadata export package created by a manual export operation. For example, using "**--daacstring AnnMiltEchoSmallMetgEchoTest**" will produce a package named: EDFManualExport.AnnMiltEchoSmallMetgEchoTest.200800710.200800710.2008007110752.000717.zip
CONCURRENCY OPTIONS		
--excludeAuto | -x		Prevent the execution of any Automatic export cycles concurrently with this manual cycle.
ERROR RECOVERY OPTIONS		
--regenerate | -r <package ID >	incompatible with **--excludeAuto** and **--delete**. Overrides all other options besides OUTPUT OPTIONS	Attempt to regenerate the AUTOMATIC package specified by the packageId <package ID>. Must specify **--noexport** if package to be regenerated is in COMPLETE state. NOTE: **packageId** must be given, NOT **cycleId**.

11.3.1 BMGT Manual Mode

1 Log in at the machine where the Bulk Metadata Generation Tool (BMGT) manual script is installed (e.g., e4oml01 and n4oml01).

2 Type **cd /usr/ecs/<*MODE*>/CUSTOM/utilities** then press **Return/Enter**.

3 To run the BMGT manually, at the UNIX prompt enter (as applicable):

EcBmBMGTManualStart.pl.

4 Select the desired command arguments using the table above.

- Example 1: Run the Manual script to generate Browse and Granule Information by collection and time and forward to ECHO.

  Enter the following:

  EcBmBMGTManualStart.pl

  -mode<MODE>

-metg

-bbr

-cf<*file*>

-starttime<*YYYY-MM-DD HH:MM:SS*>

-endtime<*YYYY-MM-DD HH:MM:SS*>

-gr<*file*>

-e

- Example 2: Run the Manual script to generate Browse inserts

  Enter the following:

  EcBmBMGTManualStart.pl

  -mode<MODE>

  -bbr

  -url

  -gf<*file*>

11.4 BMGT Automatic Mode

The BMGT Automatic Preprocessor is used by DAAC Operations Staff to export changes to the holdings of the ECS inventory at a regular interval. The DAAC will choose and configure a cycle length, which defines the time period for which metadata changes are aggregated into a single package for export to ECHO. The time period can be any whole number of hours between 1 and 24 which splits a day into a whole number of parts (e.g. 6 hours would be valid, as 4 intervals would add up to an entire day. 5 hours would not). The Preprocessor should be run at least once per export interval, and will cause the metadata changes for any preceding un-exported interval(s) to be generated and exported to ECHO. Extraneous runs of the preprocessor will have no effect. The first run of the preprocessor for a particular day will populate the export cycles for the entire day. Since the preprocessor can be run with basically no operator interaction, it can be added as a cron job such that it will run automatically at a set interval. For instance, setting a cron to run the automatic preprocessor every hour at 5 minutes past the hour would ensure that regardless of the export cycle length being used, an export package would begin generation 5 minutes after each cycle ends. On hours where a cycle is not ending, the preprocessor would simply return, with no effect.

11.4.1 BMGT Automatic Mode

1 Log in at the machine where the Bulk Metadata Generation Tool (BMGT) is installed (e.g., e4oml01 and n4oml01).

2 At the UNIX prompt, enter:
crontab -e
A vi editor window will appear. Use the arrow keys to scroll through the file and verify that there is not already an entry for the automatic preprocessor in the desired mode.

3 If there is not already an entry for the desired mode:

- Type **'o'** to open a new line.
- On this line type: **<min> <hr1>,<hr2>,…<hrn> * * * (/bin/csh -c "cd /usr/ecs/<MODE>/CUSTOM/utilities ; /usr/ecs/<MODE>/CUSTOM/utilities/EcBmBMGTAutoStart <MODE>")**
  - Where **<min>** is the number of minutes after the hour (0-59) to run at and **<hr1…n>** are the hours (0-23) during which the cron should run.
- Hit **escape** and then type **':wq'** to save the file.

4 If there is already an entry for the desired mode, but the frequency of the cron needs to be changed:

- Determine the correct values for the new frequency in the format:
- **"minute hour day month dayofweek command"**
- Use the arrow keys to navigate to the value that you with to change.
- With the cursor over the beginning of the value to change, type **'cw'** followed by the new value to change the value.
- Hit escape.
- Repeat the same for all values to be changed.
- Type **':wq'** to save the file.

5 If there is already an entry for the desired mode, but you would like to disable it:

- Use the arrow keys to navigate to the line where the entry is located.
- Type **'I'** to insert at the beginning of the line.
- Type **'#'** to comment out the line.
- Hit **escape** and then type **':wq'** to save the file.

6 If there is already an entry for the desired mode, but it is disabled by a '#' at the beginning of the line:

- Use the arrow keys to navigate to the line where the entry is.
- Type **'^x'** to remove the '#' from the beginning of the line.
- Hit **escape** and then type **':wq'** to save the file.

This page intentionally left blank.

12. Quality Assurance

In release 7.21 the Data Pool subsystem contains a new update utility for managing QA metadata. The new **QA Update Utility** consolidates the QAMUT utility from the Release 7.20 SDSRV subsystem with the QA Update Utility from the Data Pool subsystem into a single utility. The new tool receives an input file which contains header information indicating the format of the data in the file and the flags that are being updated.

12.1 Using the QA Update Tool

The QA Update Utility is an operational support tool used for updating the values of the Quality Assurance (QA) flags in the inventory metadata. The QA Update Utility sets QA values for data granules containing one or more measured parameters after they have been assessed by Science Computing Facility (SCF) or DAAC staff to determine their quality.

Data granules have Operational and Science QA flags. Operational QA flags can have the following values:

- Passed.
- Failed.
- Being Investigated.
- Not Investigated.
- Inferred Passed.
- Inferred Failed.
- Suspect.

In addition to these Operational QA flag values, Science QA flags can also have the following value:

- Hold

Table 12.1-1 provides an Activity Checklist for Using the QA Update Tool.

Table 12.1-1. Using the QA Update Tool - Activity Checklist

Order	Role	Task	Section	Complete?
1	System Administrator/ Database Administrator	Configure the QA Update Tool	(P) 12.1.1.1	
2	System Administrator/ Database Administrator	Configure the QA Update Email script	(P) 12.1.2.1	
3	Production Monitor	Prepare QA Update Request File	(P) 12.1.4.1	
4	Production Monitor	Update QA Flags Using QA Update Utility	(P) 12.1.5.1	

During one run, the QA Update Utility can update the metadata QA flags for multiple granules. In fact, the strength of the tool derives from its ability to update batches of granules at a time. There is a configurable limit on the number of granules that may be specified for a run (MAX_NUM_GRANULES in Table 12.1-2). However, this limit may be overriden by the operator. In fact, depending on how frequently the originators of requests for QA flag updates submit their requests, the DAAC may receive requests for updates of thousands of granules at a time. However, this creates the potential for extreme database loading (e.g., requirements for temporary storage of granule information). Specific practical limits may depend on individual site capacities and requirements, and the DAAC may need to work with the originators of requests to formulate requests of appropriate size to minimize QA Update Utility processing times and associated database impacts. If a request is for significantly more than that, consideration should be given to breaking it up into multiple requests.

The granules with QA flags to be updated using the QA Update Utility may each contain several different measured parameters. The tool can update the QA flag associated with each parameter for each granule listed in a metadata update request. Updates for different measured parameters related to a particular granule may be grouped contiguously on separate lines in the request so that all the updates for the granule are accomplished at the same time.

The input needed to run the QA Update Utility is a uniformly formatted update request Each update request contains an e-mail header (including the requester's return address) and an attachment request file containing a list of the granules to be updated, along with the new QA flag values for the specified parameters.

Requests sent via e-mail are automatically placed in the input request directory by the QA Update email script. Requests not sent by e-mail must be placed in the input request directory.

After the data has been copied to this input request directory, the metadata can be updated by using QA Update Utility.

The QA Update Utility retrieves a batch of granules to update from the processing table and is updated within the XML Archive. When all files listed in the batch are updated, the processing table is updated to record the work as completed. Additionally, the updates are recorded in a history file with the original and new values. When all batches are completed, the updates are recorded for BMGT to export. Finally, the files in the Data Pool file system are replaced by the newly updated files in the XML Archive. The QA Update Utility is accessible on the x4dpl01 server.

12.1.1 Configure QA Update Utility

The **EcDsAmQaUpdateUtility.properties** contains the QA Update configuration parameters and is used by the System Administrator/Database Administrator to manage the configuration of the QA Update Utility. The site installer or Database Administrator is responsible for maintaining this file. Table 12.1-2 contains a list and description of the QA Update Utility Configuration Parameters.

Table 12.1-2. Configuration File Parameters for QA Update Utility (1 of 2)

Parameter Name	Description
SYBASE_SQL_HOST	The host for the Inventory and Data Pool databases
SYBASE_SQL_SERVER	The name of the Sybase server for the Inventory and Data Pool databases
SYBASE_JDBC_DRIVER_CLASS	The java class used for connecting the QAUU java application to Sybase
SYB_DBNAME	The name of the Inventory database
SYB_DPL_DBNAME	The name of the Data Pool database
SYB_PORT	The port number used to connect to the Inventory and Data Pool databases
SYB_USER	The username used to connect to and perform queries for the Inventory and Data Pool databases
PGM_ID	The ECS Program ID for the QAUU user (SYB_USER)
DB_NUM_RETRIES	The number of times to retry failed DB operations
DB_SLEEP_SEC	The number of seconds between DB operation retries
EMAIL_SERVER_HOST	Host name where email server runs
EMAIL_SMTP_USER	Email SMTP user name
EMAIL_QAUU_FROM_ADDRESS	Email notification sender address
FILE_NUM_RETRIES	The number of times to retry failed file operations
FILE_SLEEP_SEC	The number of seconds between file operation retries
QA_REQUEST_DIR	Path of directory containing QA update requests
QA_ERROR_REQUEST_DIR	Path of directory containing QA update requests that have failed.
QA_COMPLETED_REQUEST_DIR	Path of directory containing successfully completed QA update requests
QA_TEMP_DIR	Path of directory containing temporary files
QA_HISTORY_DIR	Path of directory containing QA update history files
DAAC_EMAIL_ADDRESSES	List of valid DAAC email notification addresses
<SCFSite>_EMAIL_FROM_ADDRESSES	List of valid email notification from addresses for a <SCFSite>
<SCFSite>_EMAIL_REPLY_ADDRESSES	List of valid email notification reply addresses for a <SCFSite>
<SCFSite>_NOTIFICATION_ON_SUCCESS	Flag indicating (if = “Y”) that email notification should be sent upon successfully processing QA update requests for a <SCFSite> or for requests that fail. If = ‘N’, email should only be sent for requests that fail.
VALID_SCIENCE_QA_FLAGS	List of valid science QA flag values
VALID_OPERATIONAL_QA_FLAGS	List of valid operational QA flag values

Table 12.1-2. Configuration File Parameters for QA Update Utility (2 of 2)

Parameter Name	Description
NUM_XML_THREADS	The number of threads to be used. One thread will operate upon an UPDATE_BATCH_SIZE of QA updates.
MAX_NUM_GRANULES	The maximum number of granules that can be updated per run
UPDATE_BATCH_SIZE	The number of QA updates to be performed at a time.
XML_ARCHIVE_DIRECTORY	Pathname of XML Archive file system
SOCKS_PROXY_HOST	SOCKS proxy hostname
SOCKS_PROXY_PORT	SOCKS proxy port
BCP_EXEC_PATH	Path to unix bcp executable
SHELL_PATH	Path to unix sh shell needed to perform unix commands
application.name	Name of this application
log.operations.level	Level of logging desired in operational log: NONE, INFORMATION, VERBOSE or XVERBOSE
log.debug.level	Level of logging desired in debug log: NONE, INFORMATION, VERBOSE or XVERBOSE
log.performance.level	Level of logging desired in performance log: NONE, INFORMATION, VERBOSE or XVERBOSE
log.overwrite	If true, log file will be overwritten for each run
log.threshold	Size of log files before new ones are created.
log.rotation.number	Number of log files that will be rotated through.

12.1.1.1 Configure the QA Update Utility

1 Log into the host for the QA Update Utility (e.g., x4dpl01).

2 Change to the directory for configuration files, and then press the **Return/Enter** key.

cd /usr/ecs/<*MODE*>/CUSTOM/cfg

- The working directory is changed to **/usr/ecs/<*MODE*>/CUSTOM/cfg**.

3 Type **ls** and then press the **Return/Enter** key.

- Configuration files are displayed.

4 Find and highlight the **EcDsAmQaUpdateUtility.properties** file.

5 To start the vi editor and specify **EcDsAmQaUpdateUtility.properties** as the name of the file to be updated, type the following:

vi EcDsAmQaUpdateUtility.properties

- A new file is opened for editing and the cursor is displayed on the first character at the upper left corner of the file.
- ***Note*:** This procedure assumes use of the vi editor. Other editors may be used.

6 Type **i** to put the **vi** editor into the insert mode.

- The **vi** editor is in the insert mode, but no feedback is provided.

7 Enter/Update data to specify how to connect to the Sybase database and provide necessary DAAC-specific configuration information (see Table 12.1-2).

8 To leave the insert mode and return to the command mode, press the **Esc** key.

- The cursor moves one character to the left and the **vi** editor is in the command mode.

9 Type **ZZ** to save the file and exit the **vi** editor.

12.1.2 Configure QA Email Script

A perl script allows remote sites to submit update request input files via email as attachments. The script (EcDsQAUUEmailScript.pl) parses the request, gets the attached request file and moves it to the QAUU request directory. It will reside on the central mail servers while the rest of the QAUU will reside on other boxes. The directories containing the email script output (/usr/ecs/<mode>/CUSTOM/data/DSS/QAUU/ and subdirectories) will be created on the boxes holding the QAUU and remote mounted on the central mail servers. Email aliases need to be set up in the /etc/aliases file on the central mail servers to direct email QAUU update request to the email script. One email alias is required for each mode supporting QAUU.

12.1.2.1 Configure QA Email Aliases

1 Log into the host for the QA Update Utility (e.g., x4dpl01).

2 Change to the directory for utilities, and then press the **Return/Enter** key

cd /usr/ecs/<*MODE*>/CUSTOM/utilities

3 Type **ls** and then press the **Return/Enter** key.

4 Set up email aliases on the central mail servers (x4eil01) by entering the following:

QAUU_<OPS>:|/usr/ecs/<MODE>/CUSTOM/utilities/EcDsQAUUEmailScript.pl

QAUU_<TS1>:|/usr/ecs/<MODE>/CUSTOM/utilities/EcDsQAUUEmailScript.pl

QAUU_<TS2>:|/usr/ecs/<MODE>/CUSTOM/utilities/EcDsQAUUEmailScript.pl

12.1.3 Input File Name Format

The input file name must adhere to the following:

<MODE>_<Site>_QAUPDATE<description>.<YYYY><MM><DD><HH><MM><SS>

The following example shows the filename from site LDOPE for OPS mode at 12:20:30 on Feb. 29, 2008:

OPS_LDOPE_QAUPDATE.20080229122030

Note: All the files in the request directory will be processed alphabetically by file name and stored in the Inventory Database. The timestamp in the filenames guarantee that all the requests coming from the same site will be processed in the right order.

12.1.4 Request Format

The body of the request starts with the statement "begin QAMetadataUpdate <Science or Operational> <LGID, GranuleUR or ESDT>" and ends with an "end QAMetadataUpdate" statement. Each request can be based on 3 possible origins:

- **LGID**
- **GranuleUR**
- **ESDT with temporal range.**

In between the "begin QAMetadataUpdate" and "end QAMetadataUpdate" statements is at least one parameter/QA value statement with the following components (which must be separated by tabs):

- Short Name
- Version ID
- LGID, GranuleUR, or Range Beginning Date <tab> Range Ending date, depending on whether "LGID", "GranuleUR", or "ESDT" is specified, respectively, on the "begin" statement
- Measured Parameter Name or "ALL"
- QA Flag Value
- QA Flag Explanation Value

This information must be properly arranged and placed in the Inventory database (a designated directory or file). Figures 12.1-1 through Figure 12.1-3 contain examples the different requests.

```
From LaRC
begin QAMetadataUpdate Operational ESDT
MOD13A1    1Jul 18 2000  Jul 27 2000    ALL   Being Investigated  ESDT Update for Perf
MOD13A1    1   Jun 9 2000 Jul 11 2000    ALL   Being Investigated  ESDT Update for Perf
MOD13A1    1   Oct 2 2000 Oct 15 2000    ALL   Being Investigated  ESDT Update for Perf
end QAMetadataUpdate
```

Figure 12.1-1. Sample Metadata QA Update Request ESDT with Temporal Range

```
From LaRC
begin QAMetadataUpdate Science LGID
AIRHASCI    77  AIRHASCI.A2001181.2359.077.2003129185118.hdf    RadianceCounts
Passed          LGID  EDC Syn IV
AIRHASCI    77  AIRHASCI.A2001181.2359.077.2003133150736.hdf    RadianceCounts
                Passed LGID  EDC Syn IV
AIRHASCI    77  AIRHASCI.A2001181.2359.077.2003134164830.hdf    RadianceCounts
                Passed LGID  EDC Syn IV
AIRHASCI    77  AIRHASCI.A2001181.2359.077.2003141142634.hdf    RadianceCounts
                Passed LGID  EDC Syn IV
AIRHASCI    77  AIRHASCI.A2001181.2359.077.2003147145008.hdf    RadianceCounts
                Passed LGID  EDC Syn IV
AIRHASCI    77  AIRHASCI.A2001181.2359.077.2003148174646.hdf    RadianceCounts
                Passed LGID  EDC Syn IV
AIRHASCI    77  AIRHASCI.A2001181.2359.077.2003149211207.hdf    RadianceCounts
                Passed LGID  EDC Syn IV
AIRHASCI    77  AIRHASCI.A2001181.2359.077.2003150132315.hdf    RadianceCounts
                Passed LGID  EDC Syn IV
end QAMetadataUpdate
```

Figure 12.1-2. Sample Metadata QA Update Request with LGID

```
From LaRC
begin QAMetadataUpdate Science GranuleUR
AST_L1BUR:10:DsShESDTUR:UR:15:DsShSciServerUR:13:[PVC:DSSDSRV]:24:SC:AST_L1B.001:2007640312    ALL
Failed   SynergyIV  QA  2        Update
end QAMetadataUpdate
```

Figure 12.1-3. Sample Metadata QA Update Request with GranuleUR

12.1.4.1 Prepare QA Update Request File

1 Log into the host for the **QA Update Request File** (e.g., x4dpl01).

2 To change to the directory for **QA Update Request File**, type the following and then press the **Return/Enter** key:

cd /usr/ecs/<*MODE*>/CUSTOM/data/DSS/QAUU/QAUURequest

- The working directory is changed to **cd/usr/ecs/<MODE>/CUSTOM/data/DSS/QAUU/QAUURequest**

3 To start the **vi** editor and specify **OPS_<SITE> QAUPDATE.<yyyymmddhrminsec>** as the name of the Request file to be used by QA Update Utility, type the command:

vi OPS_<SITE> QAUPDATE.< yyyymmddhrminsec >

- The Request file is opened for editing and the cursor is displayed on the first character at the upper left corner of the file.
- ***Note*:** This procedure assumes use of the **vi** editor. Other editors may be used.

4 Type **i** to put the **vi** editor into the insert mode.

- The **vi** editor is in the insert mode, but no feedback is provided.

5 Enter request data following the proper format.

6 To leave the insert mode and return to the command mode, press the Esc key.

- The cursor moves one character to the left and the **vi** editor is in the command mode.

7 Type **ZZ** to save the file and exit the **vi** editor.

12.1.5 Update QA Metadata Flags Using QA Update Utility

Access to the QA Update Utility must be gained through the use of UNIX commands. The QAUU is started by executing the following:

EcAmQAUUStart modename [-file <filename>] [-noprompt] [-noExitonError][-recoverOnly] [-abortRecovery] [-skipRecovery] [-recoverInvestigated]

All parameters, except for **modename**, are optional.

- **modename:** The mode to run in
- **-file <filename>**: The name of the request file containing the QA updates to be applied. If omitted, all request files in the configured request directory are processed.
- **-noprompt**: if specified, the utility will not prompt the user for confirmations
- **-noExitonError**: if specified, the utility will not exit on the first error. This allows the operator to determine all errors that may occur during processing.
- **-recovery options:** These are all mutually exclusive; only one may be specified. Note that if none of these options are specified, the utility will recover, if necessary, and process new requests:
  - **-recoverOnly**: . recover and do NOT process new requests (assume we do NOT recover failures flagged as investigating)
  - **-abortRecovery**: delete all failures in working table and process new requests

- **-skipRecovery**: flag (don't process) failures for investigatation (InvestigateFlag = 'Y') and process new requests
- **-recoverInvestigated**: set InvestigateFlag = null, recover (including formally investigated failures) and process new requests

The process of updating QA metadata flags using the QA Update Utility start-up script starts with the following assumptions:

- The applicable servers are running.
- The DAAC operator has logged in to the system.
- A request for metadata update has been received in an acceptable format.
- The update request has been saved with the appropriate file name **<MODE>_<Site>_QAUPDATE<description>.<year><month><day><hour><minute><second>** (i.e., **OPS_<SITE>_QAUPDATE.< yyyymmddhrminsec)** and placed in the **/usr/ecs/<*MODE*>/CUSTOM/data/DSS/QAUU/QAUURequest directory** found on.x4dpl01 machine.

12.1.5.1 Update QA Flags Using the QA Update Utility

1 Log into the host for the QA Update Utility (e.g., x4dpl01).

2 Enter:

Run EcAmQAUUStart <**MODE**> -file <***QAUpdate Request File*****> -noexitonerr**

- The QA Update Utility retrieves the batch of granules to update from the processing table. The files listed in the batch are updated within the XML Archive.
- The history file is updated
- The DIMeasuredParameter table within the Data Pool database is updated
- The Data Pool file system are replaced by the newly updated files from the XML Archive.
- Tthe updates are recorded for BMGT to export.

This page intentionally left blank.

13. Ingest

13.1 Ingest Process

The DPL Ingest Subsystem is the part of the Science Data Processing component that the Ingest Technician uses when getting data from external data providers into the system. The Ingest Technician has access to Data Pool Ingest through the DPL Ingest GUI.

The Data Pool Ingest service is used for SIPS, S4P, Secure Copy and cross-DAAC ingest. This service supports the ingest protocol known as 'Polling with Delivery Record, and inserts the ingested data into the Data Pool SAN and archive.

The Data Pool Ingest service is also used for ingesting EMOS Detailed Activity Schedules and data type ActSched which is supported by the ingest protocol known as 'Polling without Delivery Record, and inserts the ingested data into the Data Pool SAN and archive. Data Pool Ingest does not send PANS or PDRDs to the EMOS data provider. EMOS data files do not contain checksum values and it is assumed that processed data files in the EMOS polling directory are periodically cleaned out by a DUE.

Figure 13.1-1 provides an illustration of the Data Pool ingest with Polling Delivery Record and archiving processes and are described in the following steps:

1. SIPS providers place their data and Product Delivery Record (PDR) files into a polling directory. The directory can be local, e.g., accessible via a mount point; or remote, i.e., accessible via FTP or SCP.
2. The DPL Ingest Service will poll these directories as configured by the DAAC and retrieve all new PDR files in those directories.
3. The DPL Ingest Service queues ingest requests for all PDRs that it finds. To decide which validated PDR will be processed next, it checks available resources and DAAC configured priorities.
4. The granule files are copied into the Data Pool SAN, using hidden directories for that purpose unless the DAAC requested that the data be published on insert.
5. Preprocessing events include checksum verification and translation of ODL files to XML if needed.
6. For non-SIPS ESDTs, Ingest will retrieve the MCF from a configured location in the Small File Archive.
7. Ingest validates the incoming granule metadata using the XML Validation Utility. The validated science xml metadata will be copied to a location in the StorNext Archive.
8. The Data Pool Ingest Utility (DPIU) registers the granule in the DPL database

9. The DPIU then copies the granules to the StorNext Archive. This may involve a copy to both a primary and backup archive depending on how the ESDT is configured for archiving.

10. Once all granules within the PDR are completed, the provider is notified of the outcome, which could be immediately via Product Delivery Discrepancy Report (PDRD) if PDR validation failed, or later via a short or long Product Acceptance Notification (PAN).

11. If the ESDT is configured for public Data Pool insert, granule made public in the Data Pool and populates the warehouse tables using the XML version of the metadata.

12. If the ESDT is not configured for public DPL Ingest, the granule will be cleaned up when it expires

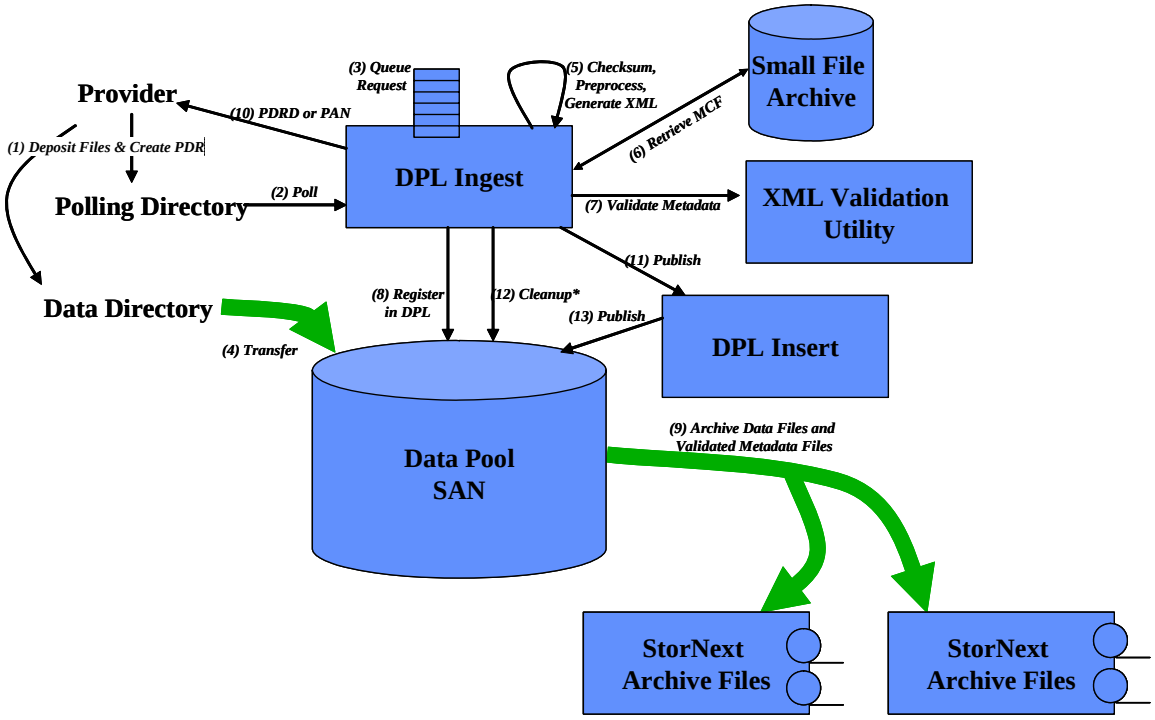


Figure 13.1-1. Data Pool Ingest High Level Architecture

The DPL Ingest Service is split into the three distinct, contiguous components. Polling (step 2) will be responsible for the provision of work to the service via transferring Product Delivery Records (PDRs) into the system and registering them.

Processing (steps 3-9) will pick up registered PDRs and attempt to ingest the inventory they describe into the Data Pool, perform any additional processing required for specific inventory (for example inventory may relate to a pending order causing processing to inform the OMS), and archive the inventory. The processing component will update the status of a particular PDR

on completion of various steps during processing, and queue a notification to be sent when all processing has completed (either successfully or unsuccessfully).

Notification (step 10) will detect the queued notification and notify the provider associated with that PDR with details of its completion state. Terminal states are Successful, Partially Failed and Failed. Terminal states are conveyed to the provider by means of a Product Acceptance Notification (PAN) or Product Delivery Discrepancy Report (PDRD).

Subsequent sections related to Ingest address the following topics:

- **Section 13.2** Contains procedures for logging in to Data Pool Ingest System Hosts.
- **Section 13.3** Contains procedures for Monitoring Data Pool Ingest System.
- **Section 13.4** Contains procedures for resolving ingest requests with open interventions and Data Pool System alerts.
- **Section 13.5** Contains procedures for modifying DPL Ingest configuration parameters.
- **Section 13.6** Contains procedures for reviewing and generating reports.
- **Section 13.7** Contains procedures accessing Help Pages.
- **Section 13.8** Contains procedures for monitoring Data Pool Collections from the Data Pool Maintenance GUI.

13.2 Logging in to System Hosts

The following procedure presents the steps required to log in to system hosts.

Table 13.2-1 contains the activity checklist for Login to the Systems Hosts.

Table 13.2-1. Login to System Hosts - Activity Checklist

Order	Role	Task	Section	Complete?
1	Ingest Technician	Log in to System Hosts	(P) 13.2.1	

13.2.1 Log in to System Hosts

1 At the UNIX command line prompt enter**: setenv DISPLAY <client name>:0.0**

- Use either the X terminal/workstation IP address or the machine-name for the client name.
- When using secure shell, the DISPLAY variable is set just once, before logging in to remote hosts. If it were to be reset after logging in to a remote host, the security features would be compromised.

2 In the terminal window (at the command line prompt) log-in to the appropriate host by entering:

ssh <host name>

- Examples of Data Pool Ingest Server host names include **e4dpl01, e4eil01**, **e4lil01**, **e4spl01** at the LP DAAC; **n4dpl01, n4eil01**, **n4lil01**, **n4spl01** at NSIDC**; l4dpl01, l4eil01**, **l4lil01**, **l4spl01 at ASDC**
- If you receive the message, "Host key not found from the list of known hosts. Are you sure you want to continue connecting (yes/no)?" enter **yes** ("y" alone will not work).
- If you have previously set up a secure shell passphrase and executed sshremote, a prompt to Enter passphrase for RSA key '<user@localhost>' appears; continue with Step 3.
- If you have not previously set up a secure shell passphrase, go to Step 4.

3 If a prompt to **Enter passphrase for RSA key '<user@localhost>'** appears, enter:

<passphrase>

- If a command line prompt is displayed, log-in is complete.
- If the passphrase is unknown, press **Return/Enter**, which should cause a **<user@remotehost>'s password:** prompt to appear (after the second or third try if not after the first one), then go to Step 4.
- If the passphrase is entered improperly, a **<user@remotehost>'s password:** prompt should appear (after the second or third try if not after the first one); go to Step 4.

4 If a prompt for **<user@remotehost>'s password:** appears, enter:

<password>

- A command line prompt is displayed.
- Log-in is complete.

13.3 Monitoring the Ingest System

The central feature for monitoring the Ingest System is the DPL Ingest GUI. The DPL Ingest GUI is a web-based interface that allows operators to access and manipulate the DPL Ingest system. Using this GUI, an operator can monitor and fix Ingest requests, view system alerts, and see at a glance the status of the DPL Ingest system. The DPL Ingest GUI also allows in-depth configuration of the entire DPL Ingest system without having to touch a database. It provides a fast and secure way to easily manage the entire DPL Ingest system, complete with full operator permission configuration and management so that only authorized persons may perform actions or change configuration settings.

Data Pool Ingest servers are started by the following three scripts:

- EcDlInProcessingService
- EcDlInPollingService
- EcDlInNotificationService

Since the DPL Ingest GUI is a web-based interface, it can be accessed from virtually anywhere there is access to the internal network. No custom software installation is required – all that is needed is a web browser (Firefox is recommended and supported) running on a Windows 2000/XP PC or a compatible Linux OS (e.g., Red Hat) that can run Firefox.

13.3.1 DPL Ingest GUI

The DPL Ingest GUI is a new feature available with delivery of Release 7.20. The ECS Data Pool Ingest GUI, illustrated in Figure 13.3-1, has five major functional areas accessible through the Navigation Panel located in the left panel of the Home Page:

- Home - Displays General System Statistics, DPL Ingest Status, Email Service Status, Notification Service Status, Polling Service Status and Processing Service Status.
- Monitoring - Allows operators to monitor currently active Ingest Requests, History of Ingest Request, Provider Status, File System Status, Transfer Host Status, ECS Service Status, and PDR List.
- Interventions & Alerts - Allows operator to review, resume, cancel and process ingests requests that have Open Interventions.
- Configuration - Allows operator to alter configuration parameters for Data Providers, Data Types, Transfer Hosts, File Systems, ECS Service Hosts, Global Tuning, Volume Groups and Operators.
- Reports - Displays information across several data providers or data types.
- Help - Provides General help topics and Context Help Information.

The navigation panel also contains a section below the menus that displays a synopsis of the current logged-in operator and provides some tools to the perform the following actions

- Log out
- Change your password
- Show all of your permissions

Operator GUI security standards require the following mulitiple levels of permissions to be assigned to each operator that has access to the DPL Ingest GUI:

- View Only
- Ingest Admin
- Ingest Control
- Security Admin
- Tuning Control

Full-capability operators have the ability to configure parameters and perform all other actions that can be accomplished with the **DPL Ingest GUI**. Limited-capability operators are able to view a lot of information; however, on the limited-capability GUI some buttons and links have been disabled so it is not possible to perform certain actions or access certain pages.

The DPL Ingest GUI is certified for use with any browser supporting the Mozilla Firefox 2.0 standard. Launching the DPL Ingest GUI starts with the assumption that the Ingest Technician has logged in to the system. Table 13.3-1 provides an activity Checklist for Monitoring DPL Ingest.

Table 13.3-1. Monitoring DPL Ingest (1 of 2)

Order	Role	Task	Section
1	Ingest Technician	Launching the DPL Ingest GUI	(P) 13.3.1.1
2	Ingest Technician	Changing Requests Status Filters	(P) 13.3.2.1
3	Ingest Technician	Monitoring Request Status	(P) 13.3.2.2
4	Ingest Technician	Cancel, Suspend, Resume or Change Requests Priority	(P) 13.3.2.3
5	Ingest Technician	Changing Suspended Granules Status	(P) 13.3.2.4
6	Ingest Technician	Viewing Historical Requests	(P) 13.3.3.1
7	Ingest Technician	Viewing Provider Status	(P) 13.3.4.1
8	Ingest Technician	Suspend or Resume Data Providers	(P) 13.3.4.2
9	Ingest Technician	Suspend or Resume Individual Polling Locations	(P) 13.3.4.3
10	Ingest Technician	Viewing File System Status	(P) 13.3.5.1
11	Ingest Technician	Suspend or Resume File System	(P) 13.3.5.2
12	Ingest Technician	Viewing Transfer Host Status	(P) 13.3.6.1
13	Ingest Technician	Suspend or Resume Transfer Host	(P) 13.3.6.2

Table 13.3-1. Monitoring DPL Ingest (2 of 2)

Order	Role	Task	Section
14	Ingest Technician	Viewing ECS Service Status	(P) 13.3.7.1
15	Ingest Technician	Suspend or Resume ECS Service(s)	(P) 13.3.7.2
16	Ingest Technician	Re-Ingesting a PDR	(P) 13.3.8.1

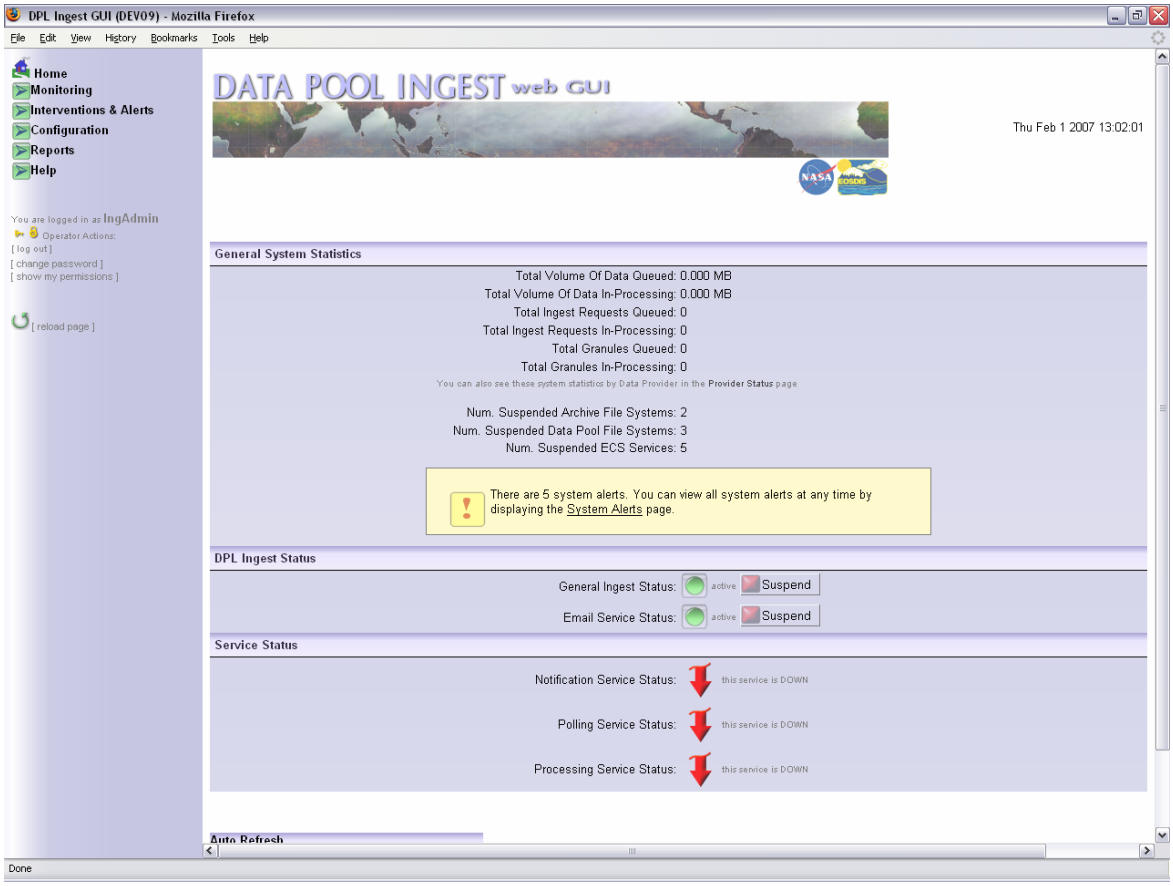


Figure 13.3-1. Data Pool Ingest GUI Home Page

The DPL Ingest GUI Home Page provides a general overview of the Data Pool Ingest system status. This page is divided into the following three sections.

- General system statistics.
- DPL Ingest Status
- Service Status

The **General System Statistics** section provides general information about current requests and granules in the system, as well as the various services and file systems used in processing.

Detail descriptions of the data found in this section is available in Table 13.3-2.

Table 13.3-2. Home Page Field Descriptions

Field Name	Description
Total Volume of Date Queued	Sum of the size of all files of all granules that have not yet been activated
Total Volume of Data In-Processing	Sum of the size of all files of all granules that are currently active, and not suspended or in a terminal state
Total Ingest Requests Queued	Total number of requests that have not yet been activated
Total Ingest Requests In-Processing	Total number of requests that are currently active, and not suspended or in a terminal state
Total Granules Queued	Sum of all granules in active or queued requests that have not yet been activated
Total Granules In-Processing	Sum of all granules in active or queued requests that are currently active, and not suspended or in a terminal state
Num Suspended Archive File Systems	Total archive file systems that have been suspended, either automatically by the server or manually by operator
Num Suspended Data Pool File Systems	Total data pool file systems that have been suspended, either automatically by the server or manually by operator
Num Suspended ECS Services	Total ECS service hosts that have been suspended, either automatically by the server or manually by operator

The **DPL Ingest Status** section consists of two buttons that enable the user to halt various actions throughout the data pool ingest system.

By pressing the **General Ingest Status** button, the operator is able to stop polling from all polling locations and prevent any new granules from being activated. Any granules that are already active will complete ingest. These actions can easily be resumed by pressing the **Resume** button (see Figure 13.3-2).

Figure 13.3-2. General Ingest Status/Resume Button

Pressing the **Email Service Status** button will stop any further email notifications (i.e. completed, cancelled, failed, or terminated requests). Once the button is pressed again, email notifications will resume (see Figure 13.3-3).

Figure 13.3-3. General Ingest Status/Resume Buttons

Service Status section provides status for three primary services that make up the Data Pool Ingest system. Ingest services cannot be started and stopped via the Data Pool Ingest GUI. Instead, they are managed using start and stop scripts found in the utilities directory of the given mode. For the status of these services to be accurate, the IngestServiceMonitor script must also be running. This script is installed in the utilities directory

(ie, /usr/ecs/<MODE>/CUSTOM/utilities) of each mode and can be started with the command: EcDlIngestServiceMonitorStart [MODE].

The services are as follows:

- **Notification Service Status** - Indicates whether the notification service is active or suspended. If suspended, no notifications will be sent, but a queue of notifications will be collected and distributed once the service is restarted (not done via the DPL Ingest GUI).
- **Polling Service Status** - Indicates whether the polling service is active or suspended. If suspended, PDRs will not arrive from any configured polling location, but any PDRs that remain in the directories will be added once the service is restarted (not done via the DPL Ingest GUI).
- **Processing Service Status** - Indicates whether the processing service is active or suspended. If suspended, no actions on any requests or granules will start, continue, or complete. Granules will "hang" in whatever state they are in (not done via the DPL Ingest GUI).

13.3.1.1 Launching the DPL Ingest GUI

1 Access a terminal window logged in to a host (e.g., the Operations Workstation or Sun external server) that has access to the **Firefox** web browser.

- Examples of Linux external server host names include e4spl01 or n4spl01.

2 Type **firefox &** then press **Return/Enter**.

- It may be necessary to respond to dialogue boxes, especially if the browser is already being used by someone else who has logged in with the same user ID.
- The Mozilla Firefox web browser is displayed.

3 If a bookmark has been created for the **DPL Ingest GUI**, select the appropriate bookmark from those listed on the browser's Bookmarks pull-down window.

- The **Login:** prompt is displayed.

4 If no bookmark has been created for the **DPL Ingest GUI**, type **http://*host*:*port*** in the browser's **Location (Go To)** field then press **Return/Enter**.

- For example: http://n4dpl01.nsidc.ecs.nasa .gov:25000/Ingest
- For example: http://p4dpl01.pvc.nasa.gov:25000/Ingest.
- For Example: http://f4dpl01.hitc.com:25010/Ingest_DEV01
- The **Login:** prompt (see Figure 13.3-4) is displayed.

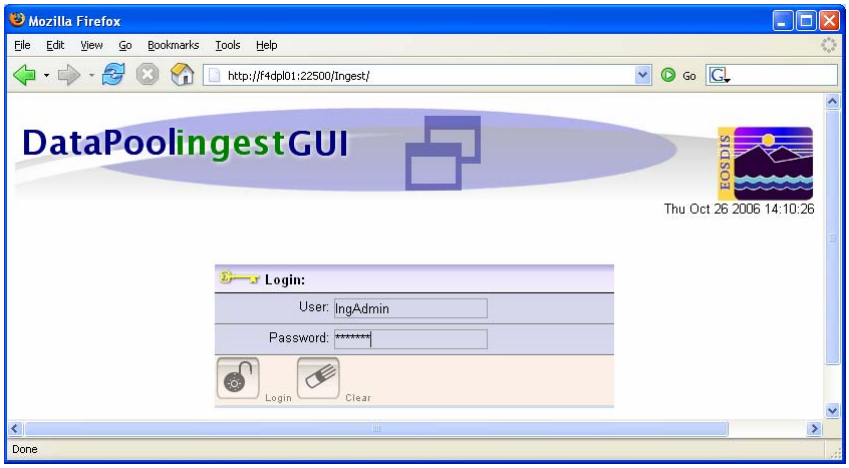


Figure 13.3-4. Ingest GUI Login Screen

5 Type the appropriate user name in the **User** box of the security **Login** prompt.
6 Type the appropriate password in the **Password** box of the security **Login** prompt.
7 Click on the **Login** button:

- The dialogue box is dismissed.
- The **DPL Ingest GUI** ["Home" Page] is displayed.

13.3.2 Monitoring Requests Status

The DPL Ingest Request Status screen is used to check the status of current active ingest requests. Table 13.3-3 provides descriptions of the information available for each request. Table 13.3-4 lists actions allowed for different status types.

This page displays the current active ingest requests. The limited-capability operator can use the Request Status page to filter and view Ingest request information.

Table 13.3-3. Request Status Page Column Descriptions

Field Name	Description
Request ID	Unique ID for an ingest request
Status	Status of the request (see Table 13.3.4 for list of possible statuses)
Priority	The precedence which a request will have for activation and various processing actions (XPRESS, VHIGH, HIGH, LOW or NORMAL).
Provider Name	Name of the provider from which the request was obtained
Size [MB]	Sum of the size of all granules in the request
Granules	Total granules included in the request
Granules Completed Processing	Total granules that have reached a successful state
When Queued	Time the request was encountered by the polling service
Last Update	Time of the last change made by the ingest services to the status of the request or its granules

Table 13.3-4. Ingest Request Status Allowed Actions

Request Status	Request Actions			
	Suspend	Cancel / Change Priority	Resume	No Actions Allowed
New		X		
Validated		X		
Active	X	X		
Partially_Suspended		X		
Suspending / Suspended		X	X	
Resuming	X			
Failed				X
Partial_Failure				X
Canceling				X
Partially_Cancelled				X
Successful				X

Sometimes it may be desirable to change the filters for the **Ingest Requests** screen. This change will alter the contents of the Ingest Requests screen. Filter settings stay the same until they are changed again.

13.3.2.1 Changing Requests Status Filters

1 Click on the **Monitoring** link in the navigation frame of the **DPL Ingest GUI.**

- The **Monitoring** menu is expanded.

2 Click on the **Requests Status** link in the navigation frame of the **DPL Ingest GUI.**

- The **Ingest Request** page is displayed.

3 Click on the **Show/Hide Filters** button.

- The **Active Ingest Request List Filter Panel** (see Figure 13.3-5) is displayed.

Figure 13.3-5. Active Ingest Request List Filter Panel

4 Click on the **Criteria Based Filtering** button.

- This allows filtering the attributes of various requests.
  - If you want to filter a single granule ID select **Filter By Request ID**.

5 Select the desired **Data Provider** from the drop-down menu.

6 Select the **Request Detail Criteria**.

- **Error Type.**
  - Selection of this state forces the filter for the **Request Detail Criteria** to be error types.
- **Request States.**

– Selection of this state forces the filter for the **Request Detail Criteria** to be requests states.

7 Select the **Target Archives** criteria**.**

8 Select the **Date Range Criteria.**

- To view entries for a particular **Date/Time Criteria**, click and hold on the option button, move the mouse cursor to the desired selection (**SHOW ALL, Last Update, Queued, Queued Within Last Hour**), then release the mouse button.
- If you selected **Last Updated or Queued** select the appropriate **FROM** Date/Time range (Month, Day, Year, Hour, Minute) and **TO** Date/Time range (Month, Day, Year, Hour, Minute),
  - Use the 24-hour format to designate the hour (e.g., type **14** to designate 2 p.m.) in the **hour** fields.
- Use the **Tab** key to advance from one field to the next.

9 If the selected filters are to be the desired default filters, click in the box next to **Save As Default Settings.**

- A checkmark is placed in the box.

10 Select the **Apply Filter** button.

- The **Ingest Requests** screen is displayed with the new filters.

13.3.2.2 Monitoring Requests Status

1 Click on the **Monitoring** link in the navigation frame of the **DPL Ingest GUI.**

- The **Monitoring** menu is expanded.

2 Click on the **Requests Status** link in the navigation frame of the **DPL Ingest GUI.**

- The **Ingest Requests** (see Figure 13.3-6) page is displayed.

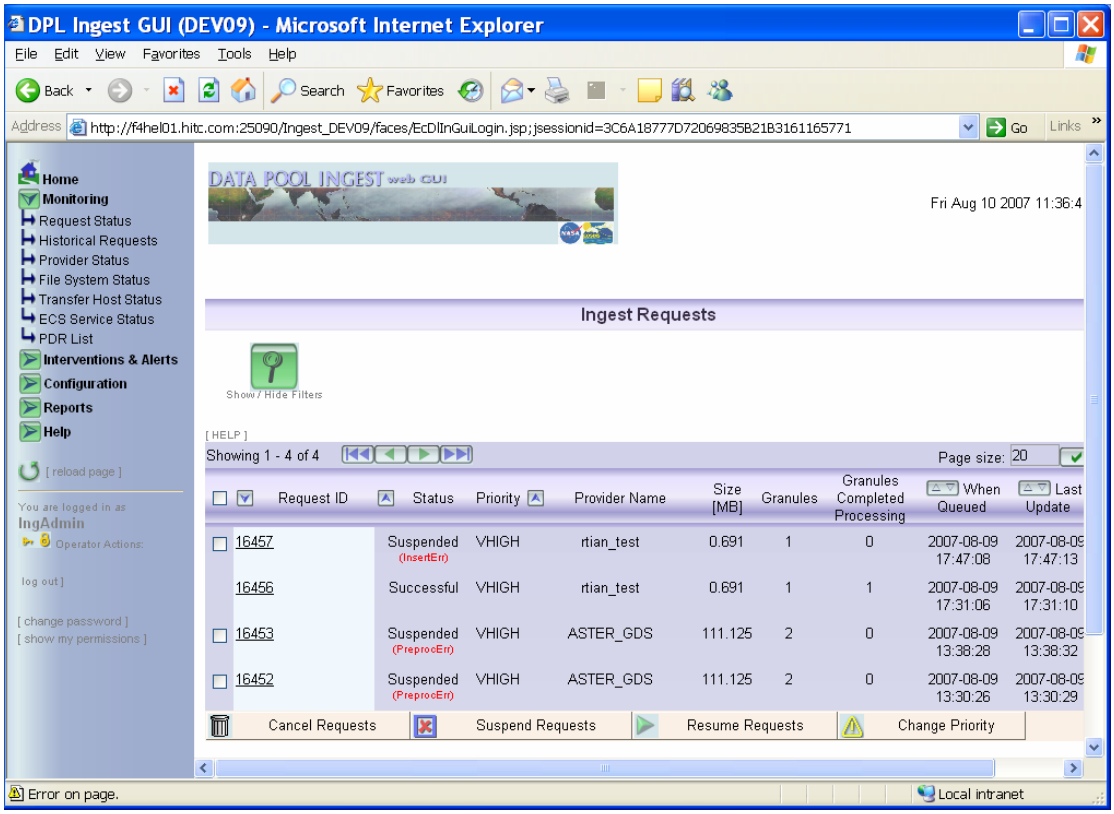


Figure 13.3-6. Ingest Request Page

- The **Ingest Requests** status page has the following columns:
  - **Request ID.**
    - Displays a unique ID for each ingest request.
  - **Status.**
    - Provides status of a request (i.e. **New, Validated, Active, Partially_Suspemded. Suspended, Canceling, Resuming, Successful, Cancelled, Partially_Cancelled, Failed, Partial_Failure** or **Terminated).**

– **Priority.**

  · The precedence which a request will have for activation and various processing actions (**XPRESS, VHIGH, HIGH, LOW** or **NORMAL).**

– **Provider Name.**

  · Name of provider from which the request was originated.

– **Size [MB].**

  · Sum of the size of all granules in the request.

– **Granules.**

  · Total Granules included in the request.

– **Granules Completed Processing.**

  · Total Granules that have reached a terminal state.

– **When Queued.**

  · Time the Request was encountered by the polling service.

– **Last Update.**

  · Time of the last change made by the ingest services to the status of the request or its granules.

3 To view the details of an ingest request (see Figure 13.3-7), click on the desired **Request ID.**

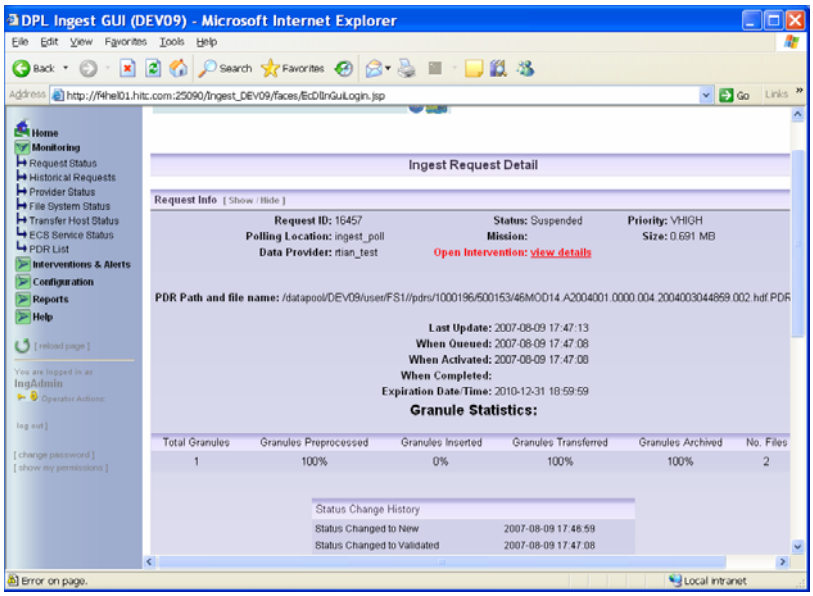


Figure 13.3-7. Ingest Request Detail Page

- The **Ingest Request Detail** page is divided into three parts.
  - **Request Info.**
  - **Granule Statistics.**
  - **List of Granules.**
- The **Request Info** contains summarized data from the **Ingest Request** status page and is located at the top of the page. The following information highlights the fields that are found only on this page
  - **Request ID.**
  - **Status.**
  - **Priority.**
  - **Polling Location.**
    - Polling path.
  - **Mission.**
  - **Size [MB].**
  - **Data Provider.**
    - Name of provider from which the request was originated.
  - **Open Interventions** link.
    - Link to the Intervention Detail page.
  - **PDR Path and file name.**
    - Location of Request.
  - **Last Update.**
  - **When Queued.**
  - **When Activated.**
  - **When Completed.**
  - **Expiration Date/Time.**
- The **Granule Statistics** contains the following information for all the granules associated with this request:
  - **Total Granules.**
    - Total number of granules included in the request**.**

- **Granules PreProcessed.**
  - Percentage of granules that have moved from the preprocessing state to the archiving state.
- **Granules Inserted.**
  - Percentage of granules that have been inserted into the Science Data Server.
- **Granules Transferred.**
  - Percentage of granules transferred from the provider to the temp directories.
- **Granules Archived.**
  - Percentage of granules that have been inserted into the archive.
- .**No. Files.**
  - Total number of files included in the request**.**

- **Status Change History** is displayed for the selected Request ID.
- **Request Notes.**
  - Notes added by the operator.
- The **File Detail** at the bottom of the screen contains a list of granules with the following associated status information**;**
  - **Seq Number.**
    - The order in which a granule was found in the PDR.
  - **Ingest Gran ID.**
    - Unique Identifier assigned to the granule.
  - **Data Type.**
    - Data Type found in the PDR describing the granule.
  - **Version.**
    - Version found in the PDR describing the granule.
  - **Status.**
    - Current granule status and detailed error information.
  - **Granule Size.**
    - Sum of the size of all files associated with the granule.

- **No. of Files.**
  - Number of files found associated with the granule in the PDR.
- **Last Status Change.**
  - Time the granule's status was last updated.
- **Path.**
  - Location of the **Granule ID.**
- **Name.**
  - Name of the **Granule ID.**
- **Type.**
  - Type of the **Granule ID.**
- **Status.**
  - Status of the **Granule ID.**

- Any granule(s) encountering problems during any point in their processing are initially flagged as "suspended". The following actions can be performed depending on the granule state:
  - **Retry Selected Granules.**
  - **Retry Selected Granules From Start.**
  - **Fail Selected Granules.**
  - **Cancel Selected Granules.**

Sometimes it may be necessary to cancel, suspend or resume the processing of one or more ingest request. The procedure for canceling, suspending or resuming granule processing starts with the assumption that all applicable servers and the DPL Ingest GUI are currently running.

13.3.2.3 Cancel, Suspend, Resume or Change Requests Priority

1 Click on the **Monitoring** link in the navigation frame of the **DPL Ingest GUI.**
- The **Monitoring** menu is expanded.

2 Click on the **Requests Status** link in the navigation frame of the **DPL Ingest GUI.**
- The **Ingest Requests** page is displayed.

3 To change one or more Request Statuses (cancel, suspend or resume) (see Figure 13.3-8), select the desired request(s) by checking the boxes on the left side of the request list.

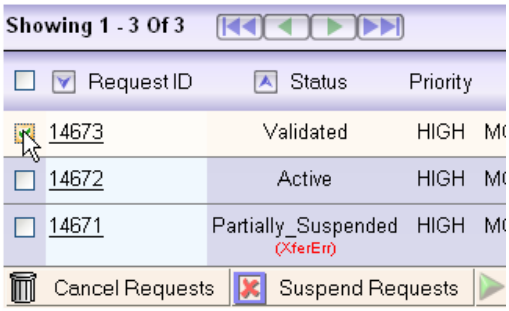


Figure 13.3-8. Cancel Request/Suspend Requests Buttons

- A checkmark is displayed in the box.
- Click on the desired action at the bottom of the list:
  - **Cancel Requests**
    - This is an irreversible action, there is no way to 'un-cancel' a request.
  - **Suspend Requests**
    - This action may be performed only if the selected request(s) are not already suspended or cancelled and is used to stop new granules from being activated. Active granules in suspended requests will continue through processing.
  - **Resume Requests**
    - This action may be performed only if the selected requests are suspended.
  - **Change Priority**
    - To change the priority of an ingest request, select the desired requests and click on the Change Priority button at the bottom of the list. A dropdown lists appears to select the new priority.
  - The Change Priority box (see Figure 13.3-9) will appear to enter a reason for the status change.

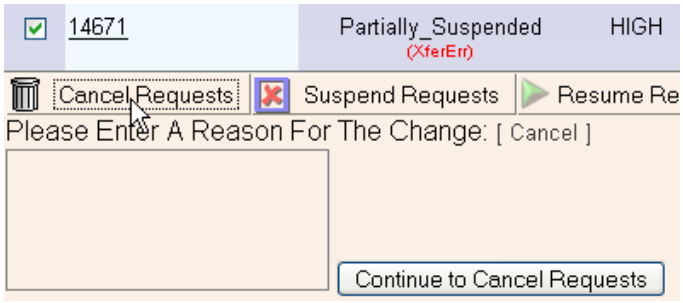


Figure 13.3-9. Change Priority Box

- Enter the reason for the change in the **Reason For Change** box.
- To cancel this action click on the **Cancel** button.
- Select the **Continue to [Cancel or Resume]** button.
- Or, Select the **OK** button.

Any granule(s) encountering problems during any point in their processing are initially flagged as "suspended". They are not failed until the operator explicitly takes an action to fail such granules. The following actions may be performed on granules that have been initially suspended:

- Retry selected granules: This applies only to granules that are currently suspended and retries them from the last known good state of processing.
- Retry from START selected granules: This applies only to granules that are currently suspended and retries them from the beginning of processing.
- Fail selected granules: This applies only to granules that are currently suspended and transitions the granule into the failed state, with the status indicating the type of error that originally caused the suspensions
- Cancel selected granules: This applies to granules that are in the New state, Active state, or Suspended state and can be cancelled by selecting this icon. If the state is Successful, Failed or any Terminal state, the granule may not be cancelled. This action manually fails the granules, marking them 'canceled.'

13.3.2.4 Changing Suspended Granules Status

1 Click on the **Monitoring** link in the navigation frame of the **DPL Ingest GUI.**

- The **Monitoring** menu is expanded.

2 Click on the **Requests Status** link in the navigation frame of the **DPL Ingest GUI.**

- The **Ingest Requests** page is displayed.

3 Click on the desired **Request ID.**

- **Request Detail Page** displayed for the selected Request ID.
- At the bottom of the **Request Detail Page** (see Figure 13.3-10), granules for the selected request(s) are listed.

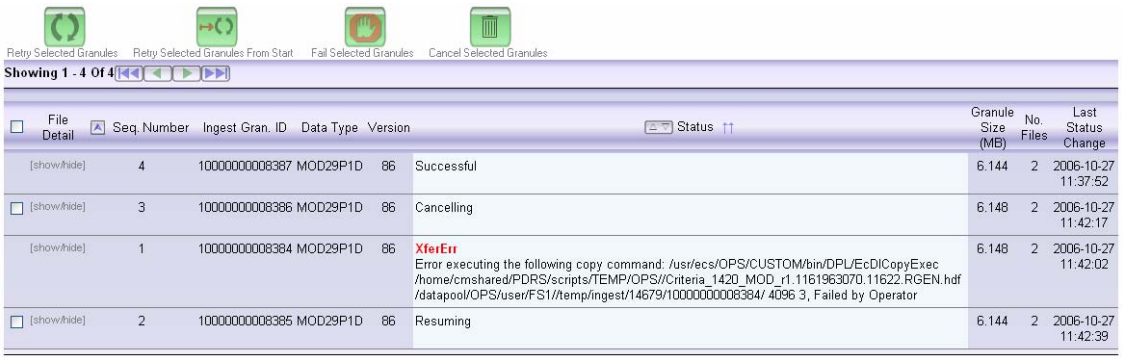


Figure 13.3-10. Request Detail Page

- Any granule(s) that encountered problems during any point in their processing are initially flagged as "suspended".

4 Click on the box next to the granule you want to change the status.

- A checkmark is displayed.

5 Select one of the following appropriate actions:

- **Retry Selected Granules.**
  - This applies only to granules that are currently suspended and will retry them from the last known good state of processing.
- **Retry Selected Granules From Start.**
  - This applies only to granules that are currently suspended and will retry them from the beginning of processing.
- **Fail Selected Granules.**
  - This applies only to granules that are currently suspended and transitions the granule into the failed state, with the status indicating the type of error that originally caused the suspensions.
- **Cancel Selected Granules.**
  - This applies only to granules that are not yet in a terminal state. It manually fails the granules, marking them 'canceled.'
- A selected action is executed and status is updated.

13.3.3 Viewing Historical Requests

When an ingest transaction has been completed, several things happen:

- A notice is automatically sent to the data provider indicating the status of the ingested data.
- The data provider sends an acknowledgment of that notice.
- Receipt of the acknowledgment is logged by Ingest.
- The **Request ID** of that ingest request is removed from the list of active requests.
- The DPL Ingest History receives statistics on the completed transaction.

The DPL Ingest Historical Requests provides the following information:

- A summary of ingest requests that have been processed.
- Historical Requests Detail gives detailed information about each completed ingest request.
- Request Timings provides ingest request processing statistics to include time required to perform Transfer, Checksum. Preprocess, DPL Insert and Archive.
- Granule List provides detailed information about each granule.

Since the Historical Requests are completed requests, no action can be processed from these pages.

13.3.3.1 Viewing Historical Requests

1 Click on the **Monitoring** link in the navigation frame of the **DPL Ingest GUI.**

- The **Monitoring** menu is expanded.

2 Click on the **Historical Request** link in the navigation frame of the **DPL Ingest GUI.**

- The **Historical Ingest Requests** (see Figure13.3-11) page is displayed with the following fields.
  - **Request ID.**
    - Unique ID for an ingest request.
  - **Status.**
    - Terminal state reached by the request.
  - **Priority.**
    - The final priority assigned to the request during processing.
  - **Provider Name.**
    - Name of the provider from which the request was obtained.

– **Size [MB].**

  · Sum of the size of all granules in the request.

– **No. Granules.**

  · Total number of successful granules included in the request.

– **Ingest Method.**

  · Whether the request was processed by Classic Ingest, or the new DataPool Ingest system.

– **When Queued.**

  · Time the request was encountered by the polling service.

– **When Proc. Started.**

  · Time the request was activated by processing.

– **When Processing Completed.**

  · Time the request reached a terminal state.

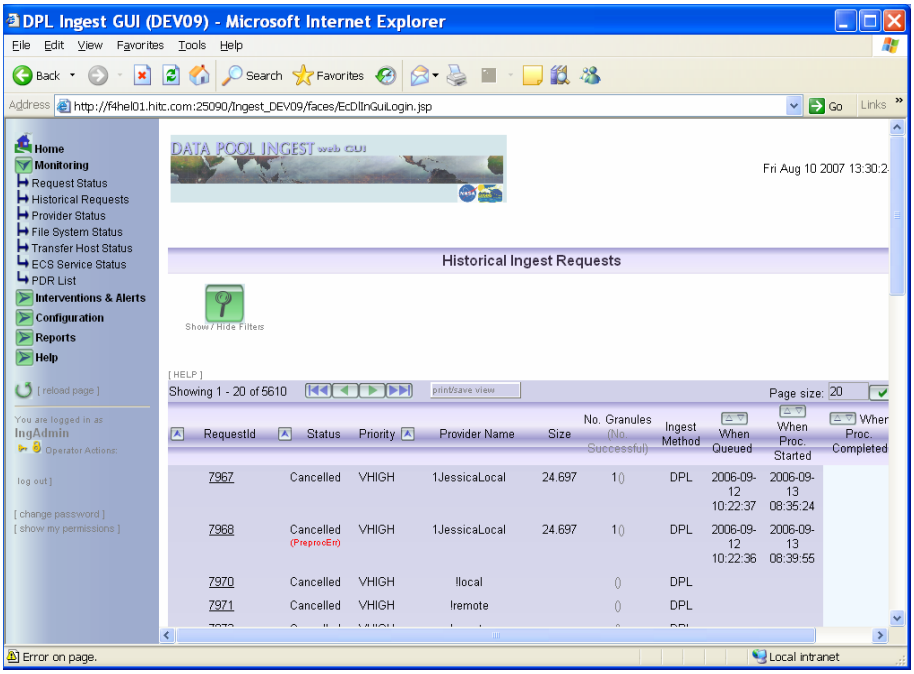


Figure 13.3-11. Historical Ingest Requests Page

3 Click on the **Show/Hide Filters** button.

- The following filters are displayed.
  - **Provider** (e.g., SHOW ALL, EDOS, ASTER, MODAPS, MISR).
  - **Request States** (SHOW ALL, Resuming, Successful, Cancelled, Partially_Cancelled, Failed, Partial_Failure, Terminated).
  - **Date/Time Range Filter** (SHOW ALL, When Completed, When Queued, Queued Within 24 Hours, Start Date and Stop Date).
  - **Data Type** (e.g., SHOW ALL, AST_L1B, MOD021KM.003, MISL0CA).

4 Select the desired **Provider** (e.g., **EDOS**) by highlighting the desired provider from the pull-down window.

5 Select the desired **Request State** (ie. **SHOW ALL, Resuming, Successful, Cancelled, Partially_Cancelled, Failed, Partial_Failure, Terminated**) by highlighting the desired request state from the window.

6 Select **Date/Time Range Filter** (ie, **When Completed, When Queued, Queued Within 24 Hours, Start Date and Stop Date**), by highlighting the desired Date/Time Range Filter from the pull-down window.

- If you selected **When Completed** or **When Queued** select the appropriate **From Date/Time Range** (**Month, Day, Year, Hour, Minute**) and **To Date/Time range** (**Month, Day, Year, Hour, Minute**).
- Use the 24-hour format to designate the hour (e.g., type **14** to designate 2 p.m.) in the **hour** fields.

7 Select a particular **Data Type** (e.g., AST_L1B) by highlighting the desired data type from the pull-down window.

8 Select the **Apply Filter** button.

- **The Historical Ingest Request** page is displayed with the new filters.
- This page shows all of the ingest requests that have been processed. The DPL database keeps a persistent record of all requests that have undergone ingest processing and can thus be viewed on this page.

9 To view the **Historical Ingest Request Detail** for a given **Request ID** , click on the desired **Request ID.**

- The **Historical Ingest Request Detail** page (see Figure 13.3-12) is displayed.
  - The layout of the request detail page for historical requests consist of **Request Info** (top section), **Request Timings** and **Status Change History** (middle section) and **Granule List** (bottom section) and are very similar to the data contained on the **Active Ingest Request** page.
  - The details on this page pertain to historical data only and can not be changed.

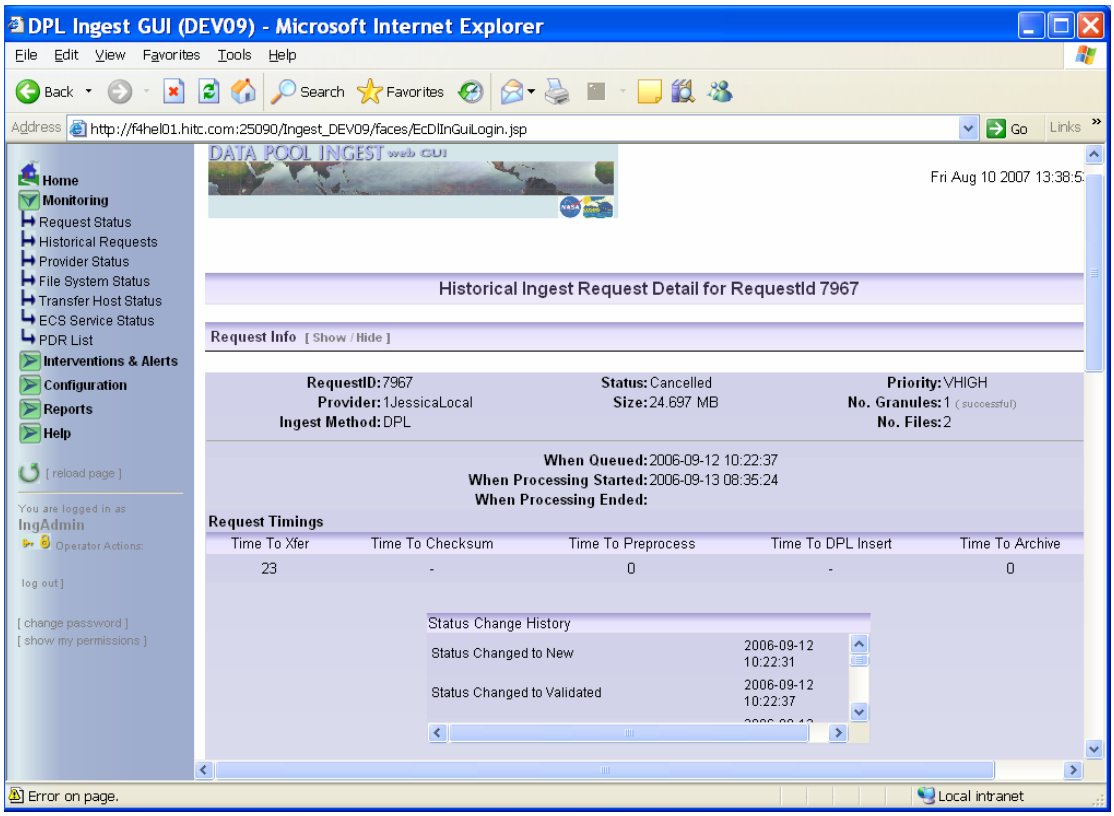


Figure 13.3-12. Historical Ingest Requests Detailed Page

– **Seq Number**.

  · The order in which a granule was found in the PDR.

– **Ingest Gran ID.**

  · Unique Identifier assigned to the granule by the DPL Ingest System.

– **ECS Gran ID.**

  · Unique Identifier assigned to the granule for insert in the SDSRV.

– **DPL Gran ID.**

  · Unique Identifier assigned to the granule for registration in the Data Pool.

– **Data Type.**

  · Data Type found in the PDR describing the granule.

– **Version.**

  · Version found in the PDR describing the granule.

- **Status.**
  - Terminal state reached by the granule.
- **Granule Size (MB).**
  - Sum of the size of all files associated with the granule.
- **No. Files**.
  - Number of files found associated with the granule in the PDR.
- **Proc. Start.**
  - Time of granule activation.
- **Proc. End.**
  - Time granule reached a terminal state.
- **Total Proc. Time.**
  - Total seconds that lapsed in between granule activation and completion.
- **Time to Checksum.**
  - Seconds that passed during granule checksum.
- **Retry Count.**
  - Number of times the granule was resumed or retried from start.

13.3.4 Provider Status

The Provider Status link provides access to the status and information about each configured data provider in the ingest system. This page provides the following:

- List of all configured providers along with general statistics for each provider.
- Provides the status of the provider (i.e. Active, Suspended by Server, or Suspended by Operator). This is the only changeable field on this page. From this page a provider can be Resumed or Suspended.
- Provides individual status for polling locations (ie total number of active or suspended polling location).
- Provides access to detailed provider status that shows individual status of each polling location associated with a provider. From this page, an individual polling location can be suspended or resumed accordingly.

13.3.4.1 Viewing Provider Status

1 Click on the **Monitoring** link in the navigation frame of the **DPL Ingest GUI.**

- The **Monitoring** menu is expanded.

2 Click on the **Provider Status** link in the navigation frame of the **DPL Ingest GUI.**

- The **Provider Status** page (see Figure 13.3-13) is displayed with the following fields.
  - **Provider.**
    - Provider name configured to identify an External Data Provider.
  - **Status.**
    - Whether the provider is **Active**, **Suspended by Server**, or **Suspended by Operator.**
  - **Polling Locations.**
    - Total number of active polling locations on the provider, or the number of polling locations that are suspended out of the total number configured.
  - **Request Queued.**
    - Total Number and Volume of requests waiting for activation.
  - **Request In Process.**
    - Total Number and Volume of requests that are active and not suspended.
  - **Granules Queued.**
    - Total number of granules waiting for activation in requests from the provider.
  - **Granules In Process.**
    - Total number of granules waiting for activation in requests from the provider.

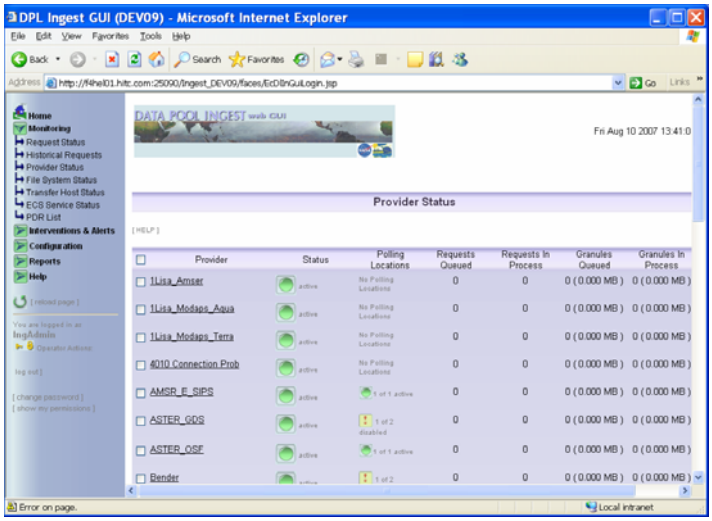


Figure 13.3-13. Provider Status Page

3 To view the individual status of each polling location associated with a given provider, Click on the desired **Provider** (ie ASTER.OSF, JPL etc.).

- The **Provider Status Detail** page (see Figure 13.3-14) is displayed.

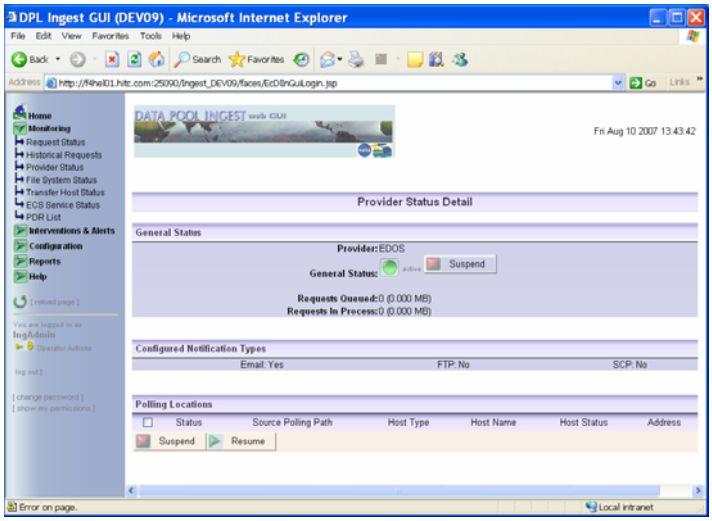


Figure 13.3-14. Provider Status Detail Page

- Displays the following **General Status** for a given Provider:
  - **Provider:**
    - Identifies the selected provider.

- **General Status:**
  - Identifies stutus of the provider (**Active** or **Suspend**).
- **Requests Queued.**
  - Total Number and Volume of requests waiting for activation.
- **Request In Process.**
  - Total Number and Volume of requests that are active and not suspended.

- **Configured Notification Types.**
  - Status of each notification method (i.e**. Email: No**, or **FTP: Yes**).
- **Polling Locations.**
  - **Status**.
    - Whether the polling location is **Active, Suspended by Server, or Suspended by Operator.**
  - **Source Polling Path.**
    - Directory being polled.
  - **Host Type.**
    - Method being used for polling – **Local, FTP, or SCP**.
  - **Host Name.**
    - Label assigned to the host on which the polling location is found.
  - **Host Status.**
    - Whether the host where the polling location is found is active or suspended.
  - **Address.**
    - Address where the polling directory can be found.

A provider may be suspended or resumed from the Provider Status page. Suspending a Data Provider will stop the activation of Ingest Requests from that Provider, but Ingest Requests that are already active will be completed. Ingest will also stop polling any of the Polling Locations associated with that Data Provider. This means that no new Requests from that suspended Data Provider will be queued except if a polling cycle is in progress, in which case the polling cycle will be completed.

13.3.4.2 Suspend or Resume Data Providers

1 Click on the **Monitoring** link in the navigation frame of the **DPL Ingest GUI.**

- The **Monitoring** menu is expanded.

2 Click on the **Provider Status** link in the navigation frame of the **DPL Ingest GUI.**

- The **Provider Status** page is displayed.

3 Select the desired provider(s) (ie ASTER.OSF, JPL etc.) by clicking in the box next to the name of the provider.

- A checkmark is displayed.

4 Select either the **Suspend** or **Resume** button located at the bottom of the page.

- You will be prompted for confirmation before the action is carried out.
- If you selected **Suspend**, the activation of Ingest Requests from that Provider will be stopped, but Ingest Requests that are already active will be completed. Ingest will also stop polling any of the Polling Locations associated with that Data Provider.
- If you selected **Resume**, the activation of Ingest Requests from that Provider will be resumed.
- The **Status** field will be updated accordingly when the requested action is completed.

Polling Locations for a provider may be suspended or resumed from the Provider Status Detail page. Each Data Provider has a list of associated Polling Locations, which are directories on FTP or local Hosts. Polling locations can be suspended or resumed in order to halt or resume data to be sent through these providers, without impacting the status of the Host on which that polling location resides.

13.3.4.3 Suspend or Resume Individual Polling Locations

1 Click on the **Monitoring** link in the navigation frame of the **DPL Ingest GUI.**

- The **Monitoring** menu is expanded.

2 Click on the **Provider Status** link in the navigation frame of the **DPL Ingest GUI.**

- The **Provider Status** page is displayed.

3 Click on the desired **Provider** (ie ASTER.OSF, JPL etc.).

- The **Provider Status Detail** page is displayed.
- At the bottom of the **Provider Status Detail**, the **Polling Locations** of the selected provider are displayed.

4 Select the desired Polling Locations to be suspended or resumed by clicking in the box associated with the **Source Polling Path**.

- A checkmark is displayed in the box.

5 Select either the **Suspend** or **Resume** button located at the bottom of the page.

- You will be prompted for confirmation before the action is carried out.
- If you selected **Suspend**, the Polling Path(s) will be stopped but Ingest Requests that are already active will be completed.
- If you selected **Resume**, the activation of Ingest Requests from that polling location will be resumed.
- The **Status** field will be updated accordingly when the requested action is completed.

13.3.5 File System Status

The **File System Status** page displays the following information on the Data Pool Archive File Systems and Data Pool File Systems:

- Name(s) and directory paths for **Archive** and **Data Pool File Systems**.
- Provides the statuses of the Archive and Data Pool File Systems (i.e. **Active, Suspended by Operator** or **Suspended by Server**). This is the only changeable field on this page. From this page Archive and Data Pool File Systems can be **Resumed** or **Suspended**.
- Provides File System space threshold metrics.

13.3.5.1 Viewing File System Status

1 Click on the **Monitoring** link in the navigation frame of the **DPL Ingest GUI.**

- The **Monitoring** menu is expanded.

2 Click on the **File System Status** link in the navigation frame of the **DPL Ingest GUI.**

- The **File System Status** page (see Figure 13.3-15) is displayed with the following fields.for **Data Pool** and **Archive File Systems.**

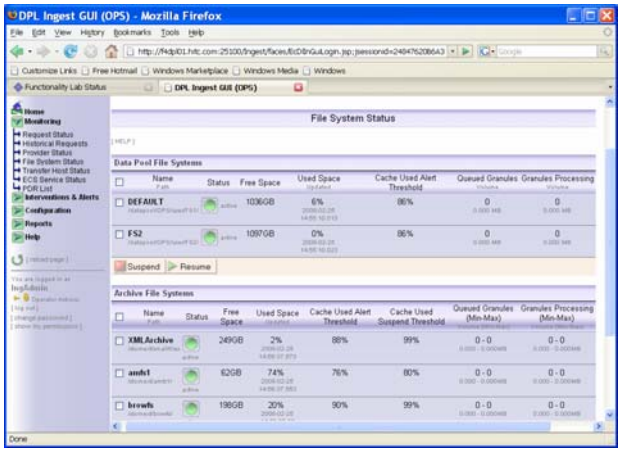


Figure 13.3-15. File System Status Page

- **Name.**
  - Unique name assigned to the file system.
- **Status.**
  - Whether the file system is active, suspended by operator, or suspended by server.
- **Free Space.**
  - Space (GB) remaining on the file system.
- **Used Space.**
  - Percentage of used space and the time this information was gathered.
- **Cache Used Alert Threshold.**
  - The percentage of used space in the cache at which point an alert would be raised. For example, if the threshold was set to 80%, an alert would be raised as soon as more than 80% of the cache was available.
- **Cache Used Suspend Threshold** (Archive File System only)**.**
  - The percentage of used space in the cache at which point the Archive File System would be suspended. For example, if the threshold was set to 90%, the Archive File System would be suspended as soon as more than 90% of the cache was available.
- **Queued Granules.**
  - Total granules waiting for activation set to ingest on the file system and the sum of the size (MB) of those granules.
- **Granules Processing.**
  - Total granules active set to ingest on the file system and the sum of the size (MB) of those granules.

Each of these archives (Data Pool File System and Archive File System) can also be suspended or resumed from the File System Status page. Suspending a File System will prevent the occurrence of any activity on the selected File System. Conversly, resuming a File System will allow activity on a File System to resume.

13.3.5.2 Suspend or Resume File Systems

1 Click on the **Monitoring** link in the navigation frame of the **DPL Ingest GUI.**

- The **Monitoring** menu is expanded.

2 Click on the **File System Status** link in the navigation frame of the **DPL Ingest GUI.**

- The **File System Status** page is displayed.

3 Click on the box next to the desired Data Pool File System or Archive File System.

- A checkmark is displayed in the box.
- Multiple selections may be made.

4 Select either the **Suspend** or **Resume** button located at the bottom of the page.

- You will be prompted for confirmation.
- The page will reload with the status of the selected archives changed.
- The application will not allow the same action to be taken twice on an Archive File System. For example, an already active status can not be resumed. However, an Archive File system that was suspended by the server may be manually suspended by the operator.

13.3.6 Transfer Host Status

The Transfer Host Status page shows the status of each configured FTP, SCP and Local transfer host. These hosts can be suspended or resumed manually, or by the Data Pool Ingest service

When an operator suspends a host, the Data Pool Ingest Service will complete any ongoing transfers, polling cycles, or notifications with that host, but not start any new ones. When an operator resumes a host, this will resume all traffic with that host. This includes polling for any previously suspended polling locations, that is, resuming a host will resume all polling locations on that host that may have been suspended.

13.3.6.1 Viewing Transfer Host Status

1 Click on the **Monitoring** link in the navigation frame of the **DPL Ingest GUI.**

- The **Monitoring** menu is expanded.

2 Click on the **Transfer Host Status** link in the navigation frame of the **DPL Ingest GUI.**

- The **Transfer Host Status** page (see Figure 13.3-16) is displayed with the following fields for all **Existing FTP Hosts and Existing SCP Hosts**.

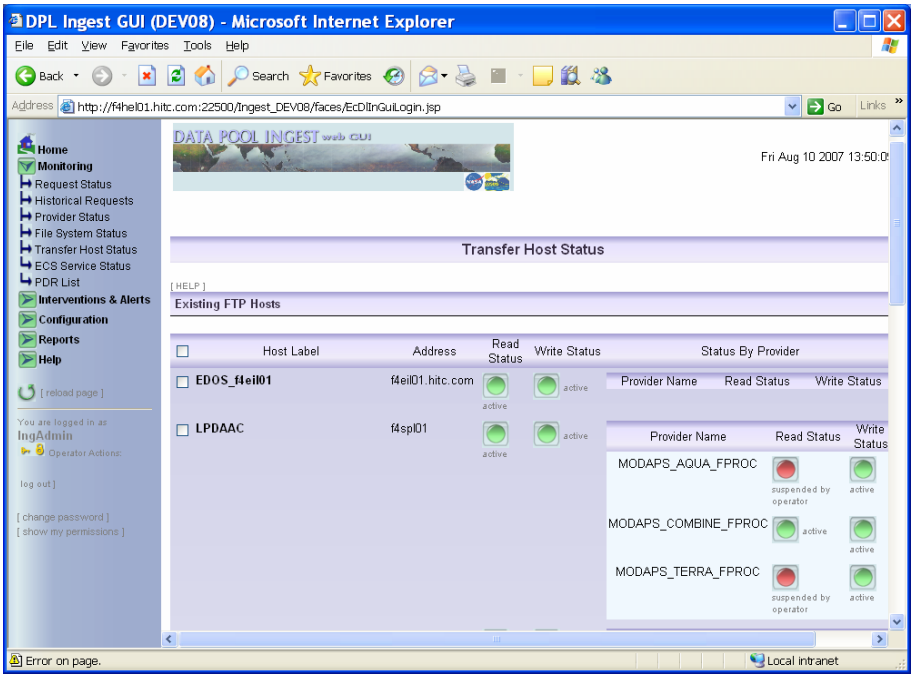


Figure 13.3-16. Transfer Host Status Page

– **Host Label.**

  · Label assigned to the host on which the polling location is found.

– **Address.**

  · The IP address or the name and port number of the host.

– **Read Status.**

  · Whether or not read actions (such as polling location listings) are active, suspended by operator, or suspended by server on the host.

– **Write Status.**

  · Whether or not write actions (such as transferring notifications) are active or suspended on the host.

– **Status By Provider.**

  · Name of the provider.

  · Whether or not read actions are active or suspended for a specific provider.

  · Whether or not write actions are active or suspended for a specific provider.

- The **Transfer Host Status** page is displayed with the following fields for all **Local Host Operations**.
  - **Read Status.**
    - Whether or not read actions (such as polling location listings) are active, suspended by operator, or suspended by server on the host.
  - **Write Status.**
    - Whether or not write actions (such as transferring notifications) are active or suspended on the host.

Each of the FTP hosts, as well as Local Host Operations, can be suspended or resumed. The status columns show a green (active) or red (suspended) icon and indicate which operation (read, write, or both) are suspended.

13.3.6.2 Suspend or Resume Transfer Host

1 Click on the **Monitoring** link in the navigation frame of the **DPL Ingest GUI.**

- The **Monitoring** menu is expanded.

2 Click on the **Transfer Host Status** link in the navigation frame of the **DPL Ingest GUI.**

- The **Transfer Host Status** page is displayed.

3 Click on the box next to the desired FTP or Local Host.

- A checkmark is displayed in the box.
- Multiple selections may be made.

4 Select either the **Suspend** or **Resume** button located at the bottom of the page.

- You will be prompted for a confirmation.
- The page will reload with the status of the selected archives changed.
- If the status is suspended, the GUI indicates whether it was suspended by an operator or automatically by the server.
- If a PDR is sent through processing with a host configured in the PDR that does not show up on the GUI, a new host will automatically be added to the list of FTP Hosts with the name UNDEFHOST_[Provider]_[RequestID]. No provider status will be displayed until the operator manually configures a provider to use that host.

13.3.7 Viewing ECS Service Status

The ECS Service Status page shows the status of each of the various ECS Services. There are two types of ECS Services.

- Services that can run on any number of hosts that have been configured for that purpose. Examples are checksumming, archiving, and transfers.
  - The service on each host is independent of the same type of service on the other hosts, in that its configuration and status is host specific. For example, checksumming on one host may be suspended but may be operating just fine on the other. As a result, the GUI shows the status information for that service separately for each host. These services are called ***Host Specific Services.***
- The status columns show a green (active) or red (suspended) icon/
- Some services exist only once and run on the host on which they were installed. An example is the Science Data Server (SDSRV) Service.
  - The DPL Ingest GUI shows only one configuration and status entry for each of those services. These services are called ***Non-Host Services***.

13.3.7.1 Viewing ECS Services Status

1 Click on the **Monitoring** link in the navigation frame of the **DPL Ingest GUI.**

- The **Monitoring** menu is expanded.

2 Click on the **ECS Service Status** link in the navigation frame of the **DPL Ingest GUI** (see Figure 13.3-17).

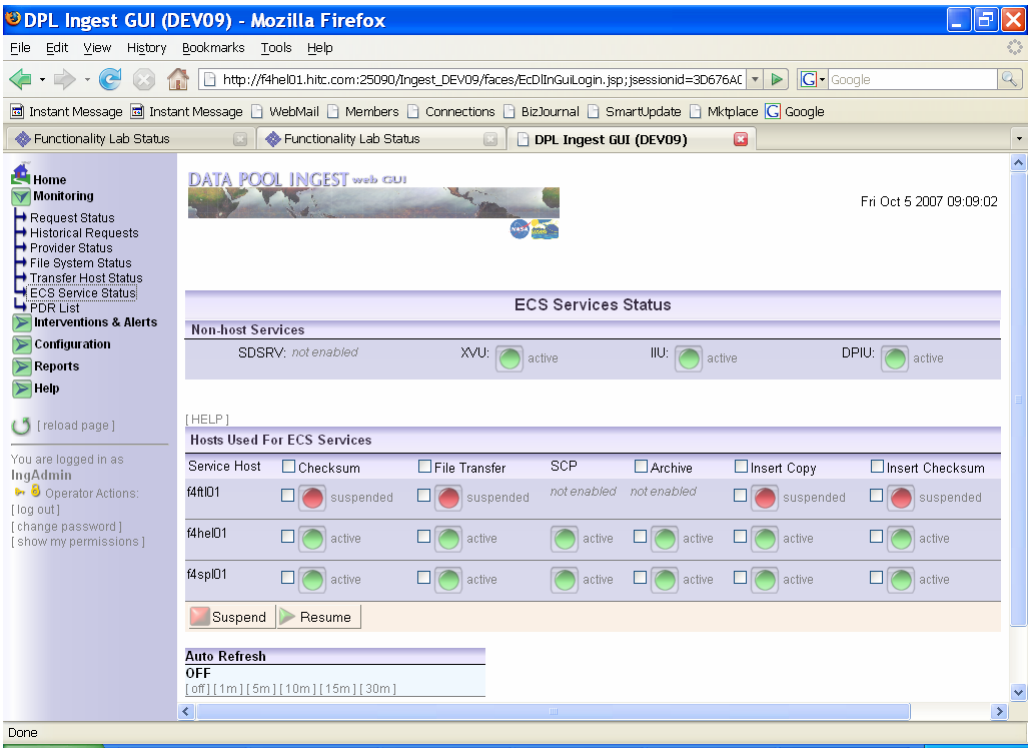


Figure 13.3-17. ECS Services Status Page

- The **ECS Service Status** page (see Figure 13.3-17) displays the status for all **Non-Host Services.**
  - **SDSRV** (in Release 7.21, the SDSVR is a configurable option and is being phased out).
    - Status: **Active**, **Suspended by Server** or **Suspended by Operater.**
  - **XVU.** The XML Validation Utility (XVU) is used to validate the incoming granule metadata.
    - Status: **Active**, **Suspended by Server** or **Suspended by Operater.**
  - **IIU**. The Inventory Insert Utility (IIU) is used to insert granule metadata into the Inventory Db
    - Status: **Active**, **Suspended by Server** or **Suspended by Operater.**
  - **DPIU**. The Data Pool Insert Utility is used to register data and populate Db tables in the Data Pool.
    - Status: **Active**, **Suspended by Server** or **Suspended by Operater.**
- The **Host Used For ECS Services** displays the following services that are tied to a specific host.
  - **Service Host.**
    - Label assigned to the host.
  - **Checksum.**
    - Status: **Active**, **Suspended by Server** or **Suspended by Operater.**
  - **File Transfer.**
    - Status: **Active**, **Suspended by Server** or **Suspended by Operater**.
  - **SCP.**
    - Status: **Active**, **Suspended by Server** or **Suspended by Operater**.
  - **Archive.**
    - Status: **Active**, **Suspended by Server** or **Suspended by Operater.**
  - **Insert Copy.**
    - Status: **Active**, **Suspended by Server** or **Suspended by Operater.**
  - **Insert Checksum.**
    - Status: **Active, Suspended by Server** or **Suspended by Operater**.

Non-Host services are not tied to a particular host. These services can be suspended or resumed by simply clicking on the button next to the indicated service status.

Suspending a service on a host specific location, will let all service operations of that type that are currently executing on that host complete, but no new requests for that service will be dispatched to that host. For example, if the Checksum service is suspended for HOST_A, ongoing checksumming operations will complete, but then no more checksumming operations will be dispatched on that host (regardless of the type of checksum involved).

13.3.7.2 Suspend or Resume ECS Service(s)

1 Click on the **Monitoring** link in the navigation frame of the **DPL Ingest GUI.**

- The **Monitoring** menu is expanded.

2 Click on the **ECS Service Status** link in the navigation frame of the **DPL Ingest GUI.**

- The **ECS Services** page is displayed.

3 In the **Non_Host Service,** click on the **Suspend** or **Resume** box next to the SDSVR **Non_Host Service**.

- The service will respond accordingly.

4 In the **Hosts Used For ECS Services.** click on the box next to the desired Service Host.

- A checkmark is displayed in the box.
- Multiple selections may be made.

5 Select either the **Suspend** or **Resume** button located at the bottom of the page.

- The page will reload with the status of the selected Service Host changed.

13.3.8 Monitoring PDR List

The PDR List page displays the PDR information retrieved from the Ingest database. The PDRs are listed by Polling Location Name and PDR File Name. Selecting a PDR from the PDR List allows the operator to re-ingest the data from the selected polling location.

13.3.8.1 Re-Ingesting a PDR

1 Click on the **Monitoring** link in the navigation frame of the **DPL Ingest GUI.**

- The **Monitoring** menu is expanded.

2 Click on the **PDR List** link in the navigation frame of the **DPL Ingest GUI.**

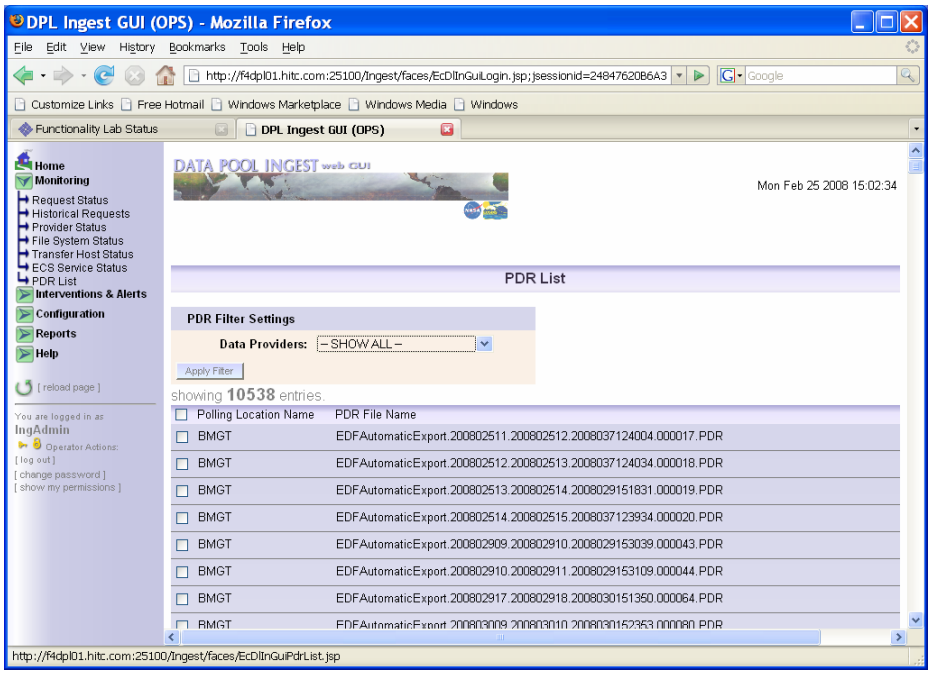


Figure 13.3-18. PDR List Page

- The **PDR List** page (see Figure 13.3-18) displays a list of Polling Locations and File Name for existing PDRs**.**

3 Click in box next to the desired PDR/Polling location name**.**

- A checkmark is placed in the box.

4 Select **Ingest Selected PDRS Again**

- The data is re-ingested.

13.4 Interventions & Alerts

The Interventions & Alerts link provides the operator access to **Ingest Requests** with open interventions. The operator may select any eligible request and either cancel the request(s) or resume the request(s). Additionally, the **Interventions & Alerts** link displays Data Pool System Alerts as they are raised in the DPL database. These warn the operator when the Ingest Service runs into a problem that it believes is associated with a resource or service it is using. After

raising an alert, the Ingest Service will check in regular intervals whether the problem has been resolved and clear the alert if that is the case. An alert may also be cleared manually once the operator determined that the problem has been resolved. An operator might do that to avoid waiting until the next auto-retry of the resource. Table 13.4-1 provides an activity Checklist for Interventions & Alerts.

Table 13.4-1. Interventions & Alerts

Order	Role	Task	Section
1	Ingest Technician	Viewing Open Intervention Ingest Requests	(P) 13.4.1.1
2	Ingest Technician	Changing Request E-Mail Configuration	(P) 13.4.1.2
3	Ingest Technician	Changing Open Interventions Ingest Requests	(P) 13.4.1.3
4	Ingest Technician	Viewing Open Intervention Detail Page	(P) 13.4.1.4
5	Ingest Technician	Changing Suspended Granules Status	(P) 13.4.1.5
6	Ingest Technician	Viewing System Alerts	(P) 13.4.2.1
7	Ingest Technician	Changing E-Mail Recipient Configuration	(P) 13.4.2.2
8	Ingest Technician	Viewing Detailed System Alert Information	(P) 13.4.2.3
9	Ingest Technician	Clearing An Alert	(P) 13.4.2.4

13.4.1 Open Intervention

The **Interventions & Alerts** link provides the operator access to **Ingest Requests** with open interventions. The operator may select any eligible request and perform one of two actions:

- Cancel the request(s) – This is an irreversible action, there is no way to 'un-cancel' a request. Processing for this ingest request will be terminated and any granules that did not yet complete processing will be considered failed. A PAN will be sent to the provider that will report the failed granules and the failure reasons.
- Resume the request(s) – Used only if the selected requests are suspended or not cancelled. Resuming a request will resume processing for all granules that are currently suspended, restarting each from the last known good state. To disposition individual granules differently, the operator needs to access the intervention detail page.

13.4.1.1 Viewing Open Intervention Ingest Requests

1 Click on the **Interventions & Alerts** link in the navigation frame of the **DPL Ingest GUI.**

- The **Interventions & Alerts** menu is expanded.

2 Click on the **Interventions** link in the navigation frame of the **DPL Ingest GUI.**

- The **Open Interventions** page (see Figure 13.4-1) is displayed and contains the following **Open Information Management** information for all interventions**:**

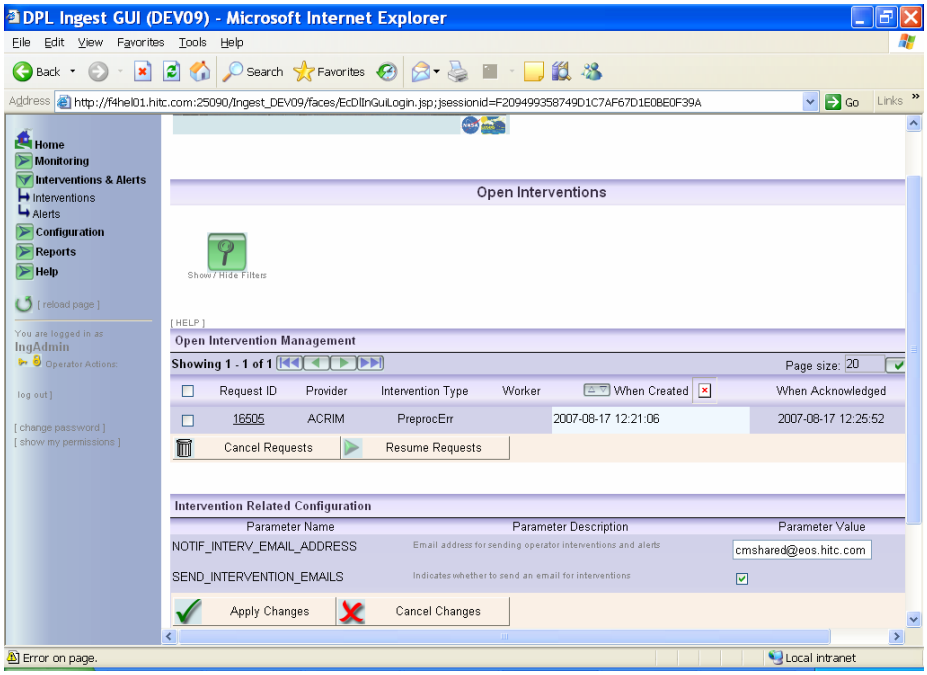


Figure 13.4-1. Open Interventions Page

– **Request ID.**

  · Unique Data Pool Ingest identifier assigned to the request in intervention.

– **Provider.**

  · Name of the provider from which the request was obtained.

– **Intervention Type.**

  · Type of error encountered during processing of at least one of the request granules: (i.e. XferErr, ChecksumErr, PreProcErr, ArchErr, InsertErr, PubErr, InitErr, Multiple).

– **Worker.**

  · Name of a worker assigned to address the intervention.

- **When Created.**
  - Time the intervention was generated (which may have been after several retries after the error was first encountered).
- **When Acknowledged.**
  - Time the intervention was first viewed by an operator.

- The **Intervention Related Configuation** fields are as follows and can be changed:
  - **Parameter Name.**
    - **SEND_INTERVENTION _EMAIL.** Indicates whether to send an Email for Intervention.
    - **NOTIFY_INTERV _EMAIL_ADDRESS.** Email address for sending operator interventions and alerts.
  - **Parameter Description.**
    - Indicates whether to send an Email for Intervention.
    - Email address for sending operator interventions and alerts.
  - **Parameter Value.**
    - **SEND_INTERVENTION _EMAIL.** Contains a box to select this parameter.
    - **NOTIFY_INTERV _EMAIL_ADDRESS.** Contains a place to enter an Email address for sending operator interventions and alerts.

The **Interventions & Alerts** link provides the operator access to **Ingest Requests** with open interventions. The operator may change the e-mail recipient configuration from this page.

13.4.1.2 Changing E-Mail Recipient Configuration

---

1 Click on the **Interventions & Alerts** link in the navigation frame of the **DPL Ingest GUI.**

- The **Interventions & Alerts** menu is expanded.

2 Click on the **Interventions** link in the navigation frame of the **DPL Ingest GUI.**

- Displays the **Open Interventions** page.

3 In the **Intervention Related Configuration** section, enter an address next to the **NOTIFY_INTERV _EMAIL_ADDRESS in the Parameter Value field.**.

4 Click on the box next to the **SEND_INTERVENTION _EMAIL** parameter field**.**

- A checkmark is displayed in the box.

5 Click on **Apply Changes** button displayed at the bottom of the **Intervention Related Configuration** section.

- A confirmation window is displayed. To confirm, click on **OK**.
- The page will reload with the new e-mail address.

When a request completes its processing, a review is made to determine the status of each granule. If at least one granule from a request is suspended because it ran into some error, the entire request is suspended and goes into Operator Intervention Status as Suspended.

From the **Open Intervention Ingest Requests** an operator can resume suspended requests regardless of the reason for a failure.

13.4.1.3 Changing Open Intervention Ingest Requests

---

1 Click on the **Interventions & Alerts** link in the navigation frame of the **DPL Ingest GUI.**

- The **Interventions & Alerts** menu is expanded.

2 Click on the **Interventions** link in the navigation frame of the **DPL Ingest GUI.**

- Displays the **Open Interventions** page.

3 In the **Open Intervention Management** section of the Open Intervention page, click on the box next to the desired **Request ID**.

- A checkmark is displayed in the box.
- Multiple selections may be made.

4 Select either the **Cancel Requests** or **Resume Requests** button located at the bottom of the page as appropriate.

- Selecting **Cancel Requests** is is an irreversible action. There is no way to 'un-cancel' a request. Processing for this ingest request will be terminated and any granules that did not yet complete processing will be considered failed. A PAN will be sent to the provider that will report the failed granules and the failure reasons.
- Selecting Resume Requests will resume a request if the selected requests are suspended or not cancelled. Resuming a request will resume processing for all granules that are currently suspended, restarting each from the last known good state. To disposition individual granules differently, the operator needs to access the intervention detail page.
- The page will reload with the status of the selected Request ID changed.

The **Interventions Detail** page is the operator's link to taking action on specific granules that have been intervened. The Request information contained in the Open Intervention page is listed at the top of the page. A list of granule(s) along with detailed information is displayed at the bottom of the page.

An Operator Intervention remains open as long as there are suspended granules. The operator can take one of several actions to 'close' the intervention (i.e., take the request out of suspension and allow the Ingest Request to be processed normally).

Once all granules issues have been resolved, the Operator Intervention status will automatically be removed. No explicit action on the part of the operator is required to do this.

If an Operator Intervention is not resolved after being viewed, it will remain in the intervention list and can be worked on at any time after navigating to a different page or even logging out of the session.

13.4.1.4 Viewing Open Intervention Detail Page

1 Click on the **Interventions & Alerts** link in the navigation frame of the **DPL Ingest GUI.**

- The **Interventions & Alerts** menu is expanded.

2 Click on the **Interventions** link in the navigation frame of the **DPL Ingest GUI.**

- The **Open Interventions** page is displayed.

3 Click on the specific **Request ID.**

- **The Open Intervention Detail** page (see Figure 13.4-2) is displayed.

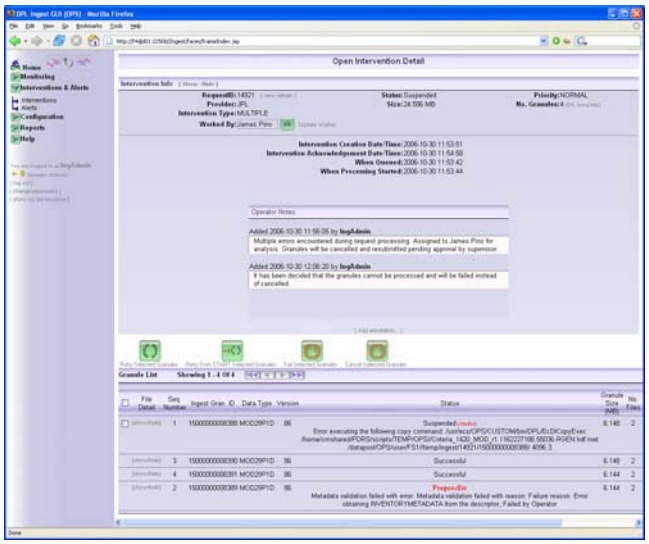


Figure 13.4-2. Open Interventions Detail Page

- The following detailed **Intervention Info** is displayed:
  - **Request ID.**
    - Unique Data Pool Ingest identifier assigned to the request in intervention.
  - **Status.**
    - Provides status of a request (i.e. New, Validated, Active, Partially_Suspemded. Suspended, Cancelling, Resuming, Successful, Cancelled, Partially_Cancelled, Failed, Partial_Failure or Terminated).
  - **Priority.**
    - The precedence which a request will have for activation and various processing actions (**XPRESS, VHIGH, HIGH, LOW** or **NORMAL).**
  - **Provider.**
    - Name of provider from which the request was originated.
  - **Size [MB].**
    - Sum of the size of all granules in tbe request.
  - **No. Granules.**
    - Total Granules included in the request.
  - **Intervention Type.**
    - Type of error encountered during processing of at least one of the request granules: (i.e. XferErr, ChecksumErr, PreProcErr, ArchErr, InsertErr, PubErr, InitErr, Multiple).
  - **Worker.**
    - Name of a worker assigned to address the intervention.
  - **Intervention Creation Date/Time.**
    - Time the intervention was generated (which may have been after several retries after the error was first encountered).
  - **Intervention Acknowledgement Date/Time.**
    - Time the intervention was first viewed by an operator.
  - **When Queued.**
    - Time request was queued.
  - **When Processing Started.**
    - Time when request processing started.

- **Operator Notes.**
    - Displays all notes along with the name of the operator who added the annotation.

- The **Granule List** contains the following information for all the granules associated with this request:

- **File Detail** contains the following detailed file information**.**
  - **Show/Hide.** A toggle button that displays the following Detail for a Granule:
    - **Path:** Directory identified in the PDR where the file can be found.
    - **Name.** Name of file.
    - **Type:** Type of file, as identified by the file extension (such as SCIENCE or METADATA).
    - **Status:** Last action performed on the file or the most recent, unresolved, error encountered while processing the file.
  - **Seq Number.**
    - The order in which a granule was found in the PDR.
  - **Ingest Granule ID.**
    - Unique Identifier assigned to the granule.
  - **Data Type.**
    - Data Type found in the PDR describing the granule.
  - **Version.**
    - Version found in the PDR describing the granule.
  - **Status.**
    - Current granule status (whether the granule is queued, its stage in processing, an error status, or its terminal state) and detailed error information.
  - **Granule Size.**
    - Sum of the size of all files associated with the granule.
  - **No. Files.**
    - Number of files found associated with the granule in the PDR.
  - **Processing Start.**
    - Time the granule's processing started.

- **Processing End.**
    - Time the granule's processing ended.

- Any granule(s) encountering problems during any point in their processing are initially flagged as "suspended". The following actions can be performed depending on the granule state:
    - **Fail Selected Granules.**
        - Suspended granules can be failed. This is a permanent action and cannot be reversed. The granule transitions into one of the ingest granule error states that indicates the type of error the granule encountered.
    - **Retry Selected Granules.**
        - The granule is retried from the last point of processing (For example: Checksum), at which point it was suspended. This is effective for most cases and requires the least amount of time to reprocess.
    - **Retry Selected Granules From Start.**
        - The granule is retried from the start of processing, no matter where in the processing chain it failed.
    - **Cancel Selected Granules.**
        - The granule is cancelled.

The **Interventions Detail** page is the operator's link to taking action on specific granules that have been intervened. The **Request** information contained in the Open Intervention page is listed at the top of the page. A list of granule(s) along with detailed information is displayed at the bottom of the page.

An Operator Intervention remains open as long as there are suspended granules. The operator can take one of several actions to 'close' the intervention (i.e., take the request out of suspension and allow the Ingest Request to be processed normally):

- The suspended granules can be failed. This is a permanent action and cannot be reversed. The granule transitions into one of the ingest granule error states that indicates the type of error the granule encountered.

- The suspended granules can be retried in one of two ways:
    - **Retry Selected Granules.**
        - The granule is retried from the last point of processing (For example: Checksum), at which point it was suspended. This is effective for most cases and requires the least amount of time to reprocess.

– **Retry From Start.**

· The granule is retried from the start of processing, no matter where in the processing chain it failed.

13.4.1.5 Changing Suspended Granules Status

1 Click on the **Interventions & Alerts** link in the navigation frame of the **DPL Ingest GUI.**

- The **Interventions & Alerts** menu is expanded.

2 Click on the **Interventions** link in the navigation frame of the **DPL Ingest GUI.**

- The **Open Interventions** page is displayed.

3 Click on the specific **Request ID.**

- **The Open Interventions Detail** page is displayed.

4 In **Granule List** section of the Open Intervention Detail page, click on the box next to the desired **Granule ID**.

- A checkmark is displayed in the box.
- Multiple selections may be made.

5 Any granule(s) encountering problems during any point in their processing are initially flagged as "suspended". Click on one of the the appropriate button.

- **Fail Selected Granules.**
  - Suspended granules can be failed. This is a permanent action and cannot be reversed. The granule transitions into one of the ingest granule error states that indicates the type of error the granule encountered.
- **Retry Selected Granules.**
  - This applies only to granules that are currently suspended. The granule is retried from the last point of. This is effective for most cases and requires the least amount of time to reprocess.
- **Retry Selected Granules From Start.**
  - This applies only to granules that are currently suspended. The granule is retried from the start of processing, no matter where in the processing chain it failed.
- **Cancel Selected Granules.**
  - The granule is cancelled.
- Once the suspended granule retry is successful, the system will automatically close the Operator Intervention.

13.4.2 Viewing System Alerts

The Ingest Operator is able to monitor Data Pool System Alerts as they are raised in the DPL database. These alerts warn the operator when the Ingest Service runs into a problem that it believes is associated with a resource or service it is using. Ingest Services checks in regular intervals whether the problem has been resolved and clear the alert if that is the case. An alert may also be cleared manually once the operator determines that the problem has been resolved.

13.4.2.1 Viewing System Alerts

1 If the **DPL Ingest GUI Home** page is displayed, and an alert has been generated, the General System Statistics section will contain a link to **System Alerts.** Clicking on this link will take you to the **Alerts** page.

- The Alerts page is displayed.

OR

2 Click on the **Interventions & Alerts** link in the navigation frame of the **DPL Ingest GUI.**

- The **Interventions & Alerts** menu is expanded.

3 Click on the **Alerts** link in the navigation frame of the **DPL Ingest GUI.**

- The **Alerts** page (see Figure 13.4-3) containing the following **Alert Management** information for the Data Pool is displayed**:**

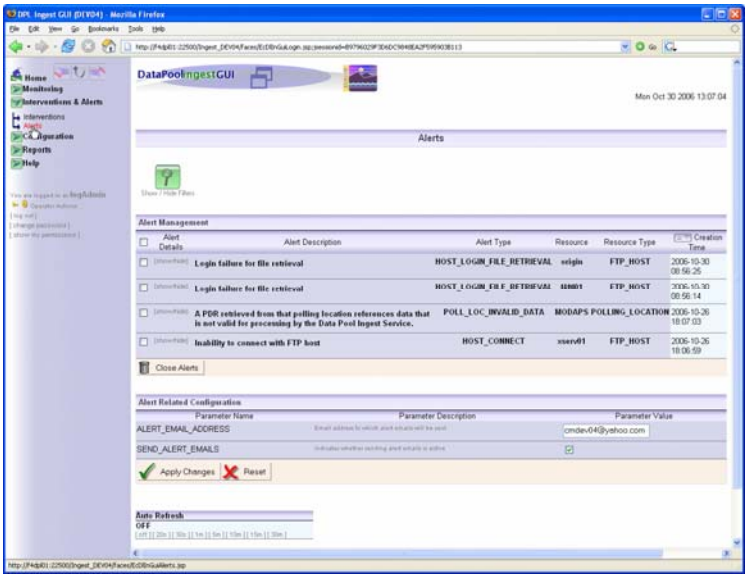


Figure 13.4-3. Open Interventions Detail Page

- **Alert Details.**
  - Buttons for displaying detailed alert information.
  - Selection of the **Show/Hide** button will display expanded detailed alert information to include **Symptom** and **Solutions**of the alert.
- **Alert Description.**
  - Basic description of the error that generated the alert.
- **Alert Type.**
  - Unique name for the type of error that was encountered.
- **Resource.**
  - The name of the resource affected by the alert.
- **Resource Type.**
  - The type of resource affected by the alert, such as SCO/FTP Host, Polling Location, or Archive.
- **Server Name.**
  - .The name of the server affected by the alert.
- **Creation Time.**
  - The time by the alert was generated.

- The **Alert Related Configuration** fields displayed as follows:
  - **Parameter Name.**
    - **ALERT _EMAIL_ADDRESS.**
    - **SEND_ALERT _EMAIL**
  - **Parameter Description.**
    - Email address to which alert emails will be sent.
    - Indicates whether sending alert emails is active.
  - **Parameter Value.**
    - Contains a place to enter an Email address for sending alerts notification.
    - Contains a box to select this parameter.
    - Contains a place to enter an Email address for sending operator interventions and alerts.

In addition to being displayed on this page, alerts can also be sent as email to a specified address. Use the following procedure to set the email address and permit email notification.

13.4.4.2 Changing E-Mail Recipient Configuration

1 Click on the **Interventions & Alerts** link in the navigation frame of the **DPL Ingest GUI.**

- The **Interventions & Alerts** menu is expanded.

2 Click on the **Alerts** link in the navigation frame of the **DPL Ingest GUI.**

- Displays the **Alerts** page.

3 In the **Alert Related Configuration** section, enter an address in the **Parameter Value** field for the **ALERT _EMAIL_ADDRESS** parameter.

4 Click on the box in the **SEND_ALERTS _EMAIL** Parameter Value field parameter**.**

- A checkmark is displayed in the box.

5 Click on **Apply Changes** button displayed at the bottom of the **Alert Related Configuration** section.

- The page will reload with the new e-mail address.

The Ingest Operator is able to monitor Data Pool System Alerts as they are raised in the DPL database. Additional details can be viewed by using the Show/Hide button and will include **Symptoms** of the alert. If the **Resource Type** is an **archive** or **file system** the Alert Details will contain **Symptoms**, **Data Provider, Request Status** information. These alerts warn the operator when the Ingest Service runs into a problem that it believes is associated with a resource or service it is using. The Ingest Services check in regular intervals whether the problem has been resolved and clear the alert if that is the case. An alert may also be cleared manually once the operator determined that the problem has been resolved.

13.4.2.3 Viewing Detailed System Alert Information

1 Click on the **Interventions & Alerts** link in the navigation frame of the **DPL Ingest GUI.**

- The **Interventions & Alerts** menu is expanded.

2 Click on the **Alerts** link in the navigation frame of the **DPL Ingest GUI.**

- The **Alerts** page is displayed**.**

3 In the **Alert Management** section of the screen**,** click on the **Show/Hide** button to display the following Detailed Alert information:

- **Symptom**
  - Information about the specific action or item that caused the alert.
- If the Resource Type is an Archive or File System, the following alert details are displayed:

- **Data Providers affected:**
  - List of providers that will be suspended as a result of the alert.
- **Number of PDRs.**
  - Total number of PDRs "stuck" in a particular state as a result of the alert.
- **Number of granules.**
  - Total number of granules "stuck" as a result of the alert.
- **Total amount of data queued.**
  - Sum of the size of the files in the granules that require the file system and will not be activated while it is suspended.
- **Total amount of data processing.**
  - Sum of the size of the files in the granules that require the file system, but will get "stuck" in an active state as a result of the alert.

An alert may be cleared manually once the operator determines that the problem has been resolved. In response, the Ingest Service will resume using that resource and all the associated resources. The Ingest Service may find that it is still unable to use the resource (e.g., still cannot connect), in which case the alert will be raised again.

It is not necessary for an operator to clear an alert manually. Normally, the Ingest Service will test in regular intervals whether the error situation has been resolved and if so, clear the alert automatically. However, it may be appropriate to clear an alert manually, for example, if the operator took some manual steps to resolve the reported problem (such as restarting an ECS Host) and then wants the Ingest Service to try using that resource immediately.

13.4.2.4 Clearing an Alert

---

1 Click on the **Interventions & Alerts** link in the navigation frame of the **DPL Ingest GUI.**

- The **Interventions & Alerts** menu is expanded.

2 Click on the **Alerts** link in the navigation frame of the **DPL Ingest GUI.**

- The **Alerts** page is displayed.

3 In **Alert Management** section of the **Alerts** page, click on the box next to the Alert(s) to be cleared.

- A checkmark is displayed in the box.
- Multiple selections may be made.

4 After selecting all alerts to be cleared, click on the **Close Alert** button**.**

- A confirmation prompt is displayed. Select **OK** or **Cancel.**
- If you selected **OK,** the page will be reloaded with the selected alerts no longer appearing on the list.

13.5 DPL Ingest Configuration

This section contains descriptions of how to modify DPL Ingest configuration values.

The DPL Ingest Configuration pages provide the full-capability operator with a means of modifying (if necessary) the values assigned to the following types of DPL Ingest configuration parameters:

- Providers.
- Data Types.
- Transfer Hosts.
- File Systems
- ECS Services
- Global Tuning
- Volume Groups.
- Operators

Table 13.5-1 provides an activity Checklist for Modifying DPL Ingest Configuration.

Table 13.5-1. Modifying DPL Ingest Configuration (1 of 2)

Order	Role	Task	Section
1	Ingest Technician	Edit a Data Provider	(P) 13.5.1.1
2	Ingest Technician	Edit a Polling Location	(P) 13.5.1.2
3	Ingest Technician	Remove a Data Provider	(P) 13.5.1.3
4	Ingest Technician	Remove a Polling Location	(P) 13.5.1.4
5	Ingest Technician	Add a Data Provider	(P) 13.5.1.5
6	Ingest Technician	Changing Default Retention Times for Data Types	(P) 13.5.2.1
7	Ingest Technician	Changing Data Types Attributes	(P) 13.5.2.2
8	Ingest Technician	Remove FTP or SCPTransfer Host	(P) 13.5.3.1
9	Ingest Technician	Add FTP or SCP Transfer Host	(P) 13.5.3.2
10	Ingest Technician	Edit FTP or SCP Transfer Host	(P) 13.5.3.3
11	Ingest Technician	Edit Local and Default Host Configuration	(P) 13.5.3.4
12	Ingest Technician	Change File System Threshold	(P) 13.5.4.1

Table 13.5-1. Modifying DPL Ingest Configuration (2 of 2)

Order	Role	Task	Section
13	Ingest Technician	Remove Checksum Type	(P) 13.5.5.1
14	Ingest Technician	Add Checksum Type	(P) 13.5.5.2
15	Ingest Technician	Update Host Used For SDSVR Operations	(P) 13.5.5.3
16	Ingest Technician	Add an ECS Service Host Type	(P) 13.5.5.4
17	Ingest Technician	Edit an ECS Service Host Type	(P) 13.5.5.5
18	Ingest Technician	Change Global Tuning Parameters	(P) 13.5.6.1
19	Ingest Technician	Add a Volume Group For a New Versioned Data Type	(P) 13.5.7.1
20	Ingest Technician	Add a Volume Group For an Existing Versioned Data Type	(P) 13.5.7.2
21	Ingest Technician	Modify Volume Groups	(P) 13.5.7.3
22	Ingest Technician	Modify Operator Permission Settings	(P) 13.5.8.1
23	Ingest Technician	Add Operator Permissions	(P) 13.5.8.2
24	Ingest Technician	Remove Operator Permission Settings	(P) 13.5.8.3

13.5.1 Data Provider Configuration

The Provider Configuration page lists all of the Data Providers for the DPL Ingest System along with the following selected attributes for each Provider.

- **Checksum Mandatory.**
- **% Files To Checksun.**
- **Default Priority.**
- **Notification Method.**

By clicking on the Provider name, the operator can view/change configuration parameters for a data provider. Table 13.5-2 contains a list of changeable Data Provider parameters and their descriptions.

Table 13.5-2. Edit a Data Provider Configuration Parameter Descriptions (1 of 2)

Field Name	Entry	Description
Name	Required	Name for an external data provider
ProviderType	Required	Indicates the type of the provider (such as Polling with DR, Polling without DR, EDOS)
Checksum Mandatory	Optional	Indicates that the Data Provider *must* provide checksum information in the PDR.
% Files to Checksum	Required	Percent of requests to checksum for this provider
Default Priority	Required	Default priority for ingest requests for this provider

Table 13.5-2. Edit a Data Provider Configuration Parameter Descriptions (2 of 2)

Field Name	Entry	Description
Preprocessing Type	Required	Type of ingest processing to occur (such as SIPS or DDIST)
Max Active Data Volume	Required	Maximum total volume that will be active on a provider if requests for other providers are pending
Max Active Granules	Required	Maximum total granules that will be active on a provider if requests for other providers are pending
Transfer Type	Required	Method used for obtaining files from the external data provider (local, FTP, or SCP with various cipher types)
Notification Method	Required	Method for providing notifications to the provider (email, SCP, FTP, or combination of SCP/FTP and email)

13.5.1.1 Edit a Data Provider

1 Click on the **Configuration** link in the navigation frame of the **DPL Ingest GUI.**

- The **Configuration** menu is expanded.

2 Click on the **Providers** link in the navigation frame of the **DPL Ingest GUI.**

- The **Provider Configuation** page is displayed. (Please see Figure 13.5-1.)

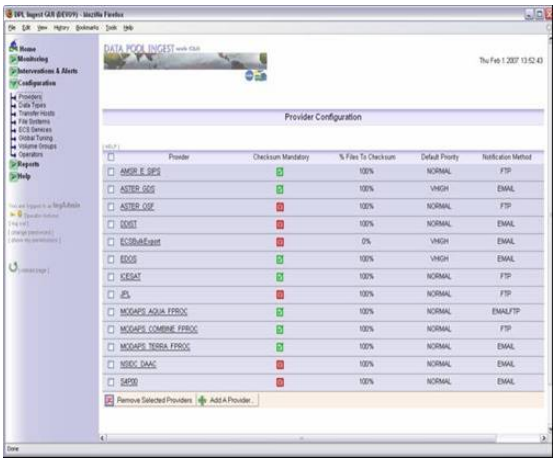


Figure 13.5-1. Provider Configuration Page

3 Click on the provider to be modidied.

- The **Edit a Provider** page (see Figure 13.5.2) is displayed.

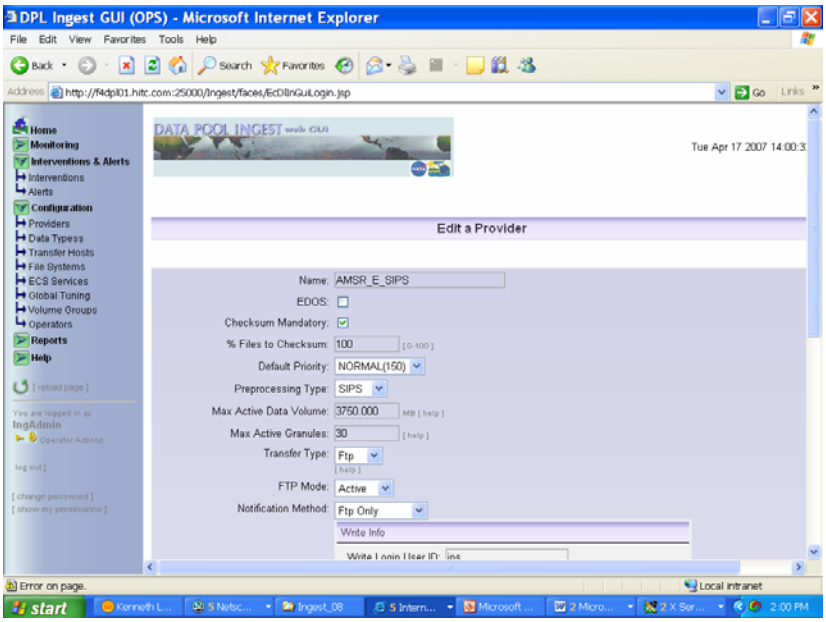


Figure 13.5-2. Edit a Provider Page

4 Select the fields to be modified.

- Certain fields are required and must have an input entered or selected**.**

5 When ready to accept changes, select the **Apply Changes** button.

- A Confirmation prompt is displayed. Select **OK**
- Changes are accepted.

Table 13.5-3 describes the changeable Polling Location fields. The following procedure contains the steps required to Edit a polling location.

Table 13.5-3. Polling Location Page Field Descriptions (1 of 2)

Field Name	Entry	Description
Parent Provider	Not Editable	Name of the provider with which this polling location is associated
Polling Location Name	Required	Name used to uniquely identify the polling location
Source Polling Path	Required	Source Directory where the PDRs are located.
Polling Frequency	Required	Number of seconds the ingest service will wait between scanning the polling path for new PDRs (must be greater than 120 seconds)
DPL Ingest Enabled	Optional	Indicates whether this polling location is enabled for ingest via DPL

Table 13.5-3. Polling Location Page Field Descriptions (2 of 2)

Field Name	Entry	Description
Polling Method	Required	Transfer method used for obtaining PDRs from the polling location (Local Disk. Or Ftp Host)
Host Name	Required if using a remote transfer method	Host where the polling directory is found

13.5.1.2 Edit a Polling Location

1 Click on the **Configuration** link in the navigation frame of the **DPL Ingest GUI.**

- The **Configuration** menu is expanded.

2 Click on the **Providers** link in the navigation frame of the **DPL Ingest GUI.**

- The Provider Configuation page is displayed.

3 Click on the provider to be modidied.

- The **Edit a Provider** page is displayed (see Figure 13.5-2).

4 Scroll down to the bottom of the information panel.

- The **Existing Polling Locations** are displayed.

5 Select the polling location name.

- The **Edit a Polling Location** page (see Figure 13.5-3) is displayed.

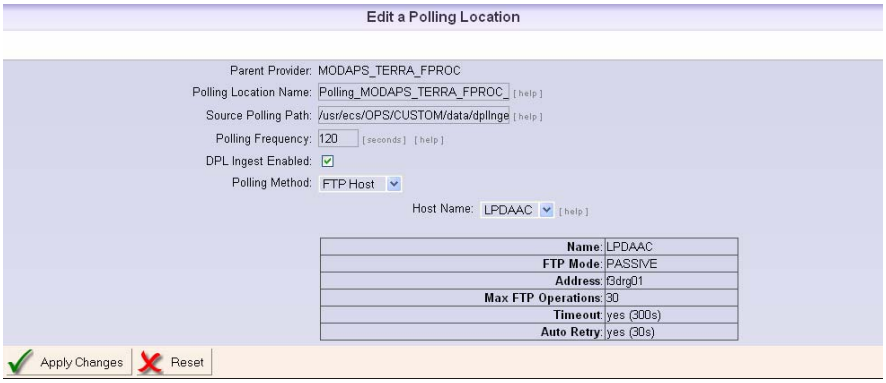


Figure 13.5-3. Existing Polling Location Page

6 Enter the desired changes.

7 Select the **Apply Changes** button.

- A Confirmation prompt is displayed. Select **OK**
- Changes are accepted.

13.5.1.3 Remove a Data Provider

1 Click on the **Configuration** link in the navigation frame of the **DPL Ingest GUI.**

- The **Configuration** menu is expanded.

2 Click on the **Providers** link in the navigation frame of the **DPL Ingest GUI.**

- The **Provider Configuation** page is displayed.

3 Click on the box next to the provider to be removed.

- A checkmark is placed in the box.
- Multiple selections are accepted.

4 At the bottom of the screen, select the **Remove Selected Providers** button.

5 A Confirmation screen is displayed. Select **OK**

- Changes are accepted.

13.5.1.4 Remove a Polling Location

1 Click on the **Configuration** link in the navigation frame of the **DPL Ingest GUI.**

- The **Configuration** menu is expanded.

2 Click on the **Providers** link in the navigation frame of the **DPL Ingest GUI.**

- The **Provider Configuation** page is displayed.

3 Click on the **Provider** name.

- **The Edit a Provider** Configuation page is displayed.

4 Scroll down to the bottom of the information panel.

- The **Existing Polling Locations** are displayed.

5 Click on the box next to the polling location to be removed.

- A checkmark is placed in the box.
- Multiple selections are accepted.

6 At the bottom of the screen, select the **Remove Selected Polling Locations** button.

7 A Confirmation screen is displayed. Select **OK**

- Changes are accepted

13.5.1.5 Add a Data Provider

1 Click on the **Configuration** link in the navigation frame of the **DPL Ingest GUI.**

- The Configuration menu is expanded.

2 Click on the **Providers** link in the navigation frame of the **DPL Ingest GUI.**

- The **Provider Configuation** page is displayed.

3 On the **Provider Configuration** page, click on the **Add Provider** button at the bottom of the existing provider list.

- The **Add a Provider** (see Figure 13.5-4) page is displayed.

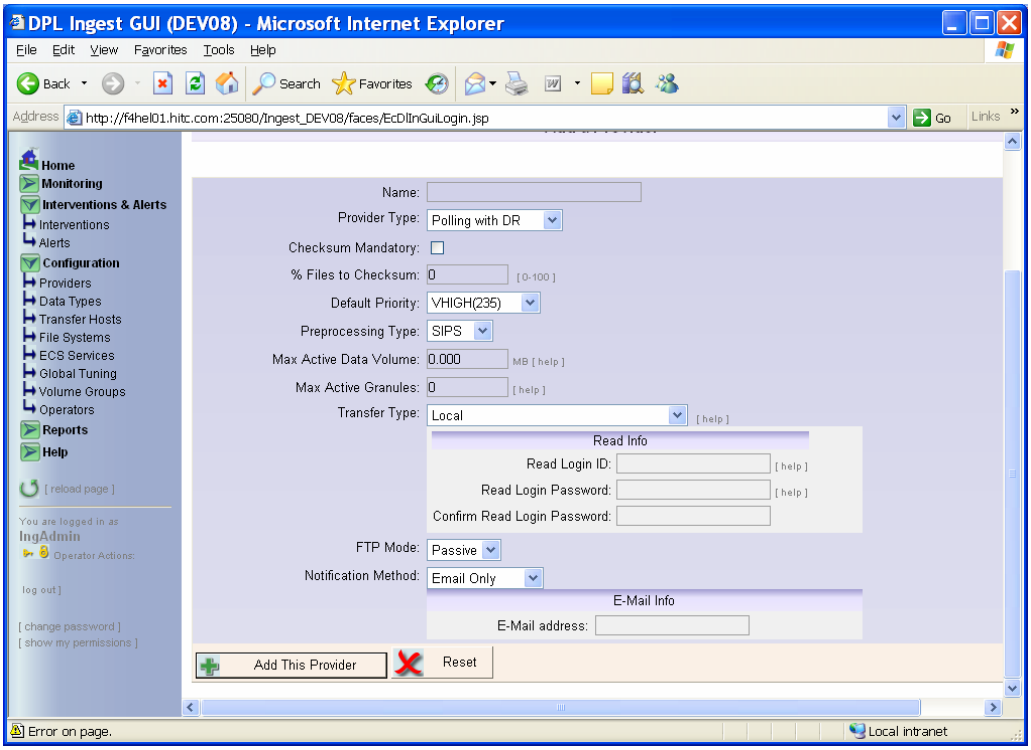


Figure 13.5-4. Add a Provider Page

4 In the **Name** field enter a unique name for this provider.

- Already existing names will be rejected by the database.

5 Select the **Provider Type** from the pull-down window.

- If you select **EDOS**, **Preprocessing Type** will become **NONE**, **Transfer Type** will become **FTP** and **Notification** method will become **FTP Only**. These options cannot be changed.

- If you select **Polling without DR**, a **VersionedDataType** drop-down list will appear on the page for operator to select the ESDT this provider will ingest from a predefined list of polling without DR ESDTs; the **Checksum Mandatory** checkbox will be unchecked and disabled, the **% Files to Checksum** will be set to 0 and disabled; the **Preprocessing Type** will become **NONE** and **Notification Method** will become **NONE**. These options cannot be changed.

6 If applicable, check the box for **Checksum Mandatory**;

- If this box is checked, the DPL Ingest System will perform **100% checksum** regardless of the "% of Checksum Files" entered.
- If you would like to checksum at less than 100%, disable checksum mandatory and enter the desired percent.

7 Select one of the following default priorities from the **Default Prioity** pull-down window.

- **LOW (60), NORMAL (150), HIGH (220), VHIGH (235), XPRESS (255)**.

8 Select **Preprocessing Type** from the pull-down window.

- **NONE, SIPS, DDIST**

9 Enter the maximum active data volume in the **Max Active Data Volume** field that can be processed at the same time on this provider.

- The Ingest Service uses the maximum data volume and number of granules to limit the amount of the work which it will activate for a provider.
- Ingest will activate a new granule for an active ingest request when the amount of work for the provider that is currently in progress reaches one of the configured limits.
- New granules will be activated as granules complete and slots are opened up.

10 Enter the maximum number of granules in the **Max Active Granules** field that can be processed at the same time.

- New granules will be activated as granules complete and slots are opened up.
- Note: There are overall limits on the total amount of work in progress, across all providers, which may further limit how much work is activated.

11 Select the **Transfer Type** from the pull-down window.

a. If data transfer will be **FTP**, the operator must enter "FTP Read Info" at the bottom of the page.

- If this information is not filled out, when a polling location is added, the operator will not be able to select ftp as the transfer method.

12 Select **FTP Mode** (FTP Host only) from the pull-down window.

- **Active** or **Passive**.

13 Select the **Notification Method** from the pull-down window. Depending on your selection, the appropriate boxes will appear below the drop-down list:

a. **Email Only**: enter a valid Email address in the **E-Mail Info** field.

b For **FTP only or SCP only:** enter the login information **(Write Login User ID, Write Login Password, Write Login Password Confirmation,** the directory **Path** and the **Choose Host** information**.**

- c Pick an existing, pre-configured **FTP host** as defined in the FTP Host Configuration page from the pull-down window: When you select the desired host, an information box is displayed, showing the host's login information, IP address, and other details:

d. If you select **Local:** Enter the local disk directory.

e. For **Email and FTP**, or **Email and SCP:** enter the E-Mail imformation and write login information **(Write Login User ID, Write Login Password and Write Login Password Confirmation),** the directory **Path** and the **Choose Host** information**.**

14 Add this provider by clicking the **Add This Provider** button at the bottom of the screen.

Note: Polling locations can not be added until the provider has been added.

- A confirmation screen is displayed. Select **OK**
- The Provider Configuration page is displayed.

15 Select the new provider just entered.

- **Edit a Provider** detail page is displayed.

16 At the bottom of the page, click **Add a Polling Location** button.

Note: The provider will not become active until at least one polling location is added.

- **Add a Polling Location** page is displayed.

17 Enter a unique name for the **Polling Location Name**.

- Names that already exist for this polling location will be rejected by the database.

18 Enter the **Source Polling Path**.

- This is the pathname from which to transfer the PDR files.

19 Enter the **Polling Frequency** in seconds.

- The minimum value is 120 seconds.

20 Select whether or not this Polling Location is **DPL Ingest Enabled**.

- A checkmark is placed in the box.

21 Choose the **Polling Method** from the pull-down list of pre-configured hosts.

- Selecting **FTP Host** will cause an information box to be displayed, showing the host's login information, IP address, and other details.
- Selecting **Local Disk** will not require additional (the directory path is already provided at the top of the page).

22 Click the **Add Polling Location** button at the bottom of the screen.

13.5.2 Data Type Configuration

Any DataPool (DPL) collection is eligible for DPL ingest. DPL collections are added via the **DataPool Maintenance GUI**. Default assumptions are given for each Data Type collection depending on whether or not it is an ECS data type. The default assumptions are:

- ECS collections are archived but not inserted into the public Data Pool upon ingest. The operator can change this so all granules are inserted into the public Data Pool as soon as they complete normal ingest processing for each data type. This would take the place of an unqualified subscription for Data Pool insert and is more efficient.
- The operator can configure a default public and hidden retention time for all datatypes. Adding a public retention period will guarantee that they remain in the Data Pool for ordering purposes after ingest for the specified time. Otherwise, they will be removed immediately after archiving completes. The operator can override the default retention for individual collections.

The Configuration Data Type link allows the DAAC users to change the default assumptions. Table 13.5-4 provides decriptions of the fields found on this page.

Table 13.5-4. Data Type Configuration Page Field Descriptions

Field Name	Entry	Description
Short Name	Not Editable	Data Type identifier
Version ID	Not Editable	Version number of the data type
Public In Data Pool	Optional	Indicates whether or not to “publish data” for this data type in the public Data Pool following successful Ingest.
Ignore Validation Warnings	Optional	Designates the ESDTs configured to ignore XML validation warnings.
Public Retention Time	Required if set for publication	Days to keep granules of this data type in the public data pool. If no value is entered, a default value will be used.
Public Priority	Required if set for publication	The precedence which a data type will have for publication in the data pool. If no value is entered, a default value will be used.
Hidden Retention Time	Required	Days to keep granules of this data type in the hidden data pool

13.5.2.1 Changing Default Retention Times for Data Types

1 Click on the **Configuration** link in the navigation frame of the **DPL Ingest GUI.**

- The **Configuration** menu is expanded.

2 Click on the **Data Type** link in the navigation frame of the **DPL Ingest GUI.**

- **The Data Type Configuation** (see Figure 13.5-5) page is displayed.

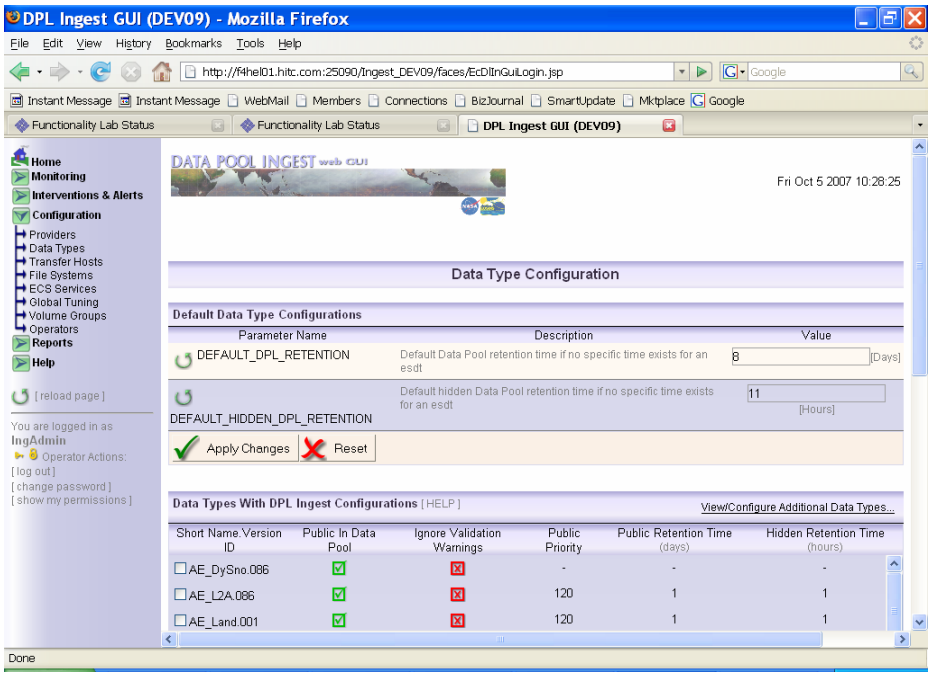


Figure 13.5-5. Data Type Configuration Page

- Displays **Default Data Type Configurations** (top of page)**.**

- Displays **Data Types With DPL Ingest Configurations** whose configurations have been altered to support non-default options (middle of page).

3 **The DEFAULT_DPL_RETENTION Value** for a public retention time can be changed. Adding a public retention period will guarantee that data types will remain in the Data Pool for ordering purposes after ingest for the specified time. Otherwise, they will be removed immediately after archiving completes. Click in the **Value** field and enter the desired number of **Days.**

4 The **DEFAULT_HIDDEN_DPL_RETENTION Value** for a hidden retention time can be changed. Adding a hidden retention period will guarantee that data types will remain in the Data Pool for ordering purposes after ingest for the specified time. Otherwise, they will be removed immediately after archiving completes. Click in the **Value** field and enter the desired number of **Hours.**

5 Click on the **Apply Changes** button.

- A Confirmation prompt is displayed. Select **OK**

- Rentention default values will be change.

13.5.2.2 Changing Data Types Attributes

1 Click on the **Configuration** link in the navigation frame of the **DPL Ingest GUI.**

- The **Configuration** menu is expanded.

2 Click on the **Data Type** link in the navigation frame of the **DPL Ingest GUI.**

- **The Data Type Configuation** page (see Figure 13.5-5) is displayed.
- Displays **Default Data Type Configurations** (top of page)**.**
- Displays **Data Types With DPL Ingest Configurations** whose configurations have been altered to support non-default options (middle of page).
- If the **Data Type** you want to change is not listed, select the **View/Configure Additional Data Types** link**.** This will display the list of **Data Types *without* DPL Ingest Configurations.**

3 Select each **Data Type** by clicking on the box nexy to the **Short Name Version ID**.

- A checkmark with appear in the box.

4 Scroll to the bottom of the screen until the Modify Selected Data Types or the Configure selected Data Types section appears.

5 Make the desired changes for the following parameters.

- **Publish in Public DPL.** Indicates whether or not to "publish data" for this data type in the public Data Pool following successful Ingest. Place a checkmark to select this option by clicking on the box.
- **Ignore Validation warnings.** Selection of this parameter will allow ingest to ignore warnings received from the XML Validation Utility for the selected ESDT. Place a checkmark to select this option by clicking on the box.
- **Public Retention Time.** Determines number of days to keep granules of this data type in the public data pool. Enter a value in the box next to the field name.
- **Public Priority.** The precedence which a data type will have for publication in the data pool. Enter a value in the box next to the field name.
- **Hidden Rentention Time.** Days to keep granules of this data type in the hidden data pool. Enter a value in the box next to the field name.

6 Click on the **Apply Changes** button.

- A Confirmation prompt is displayed. Select **OK.**
- Parameters for the selected **Data Types** are change to the new settings.

13.5.3 Transfer Host Configuration

The **Transfer Host** link allows the operator to manage SCP and FTP hosts for general use in the Data Pool Ingest system. These hosts can be referenced when defining polling locations or notification hosts. Table 13.5-5 provides a list of the parameters.

In addition, if the host ip-addresses are referenced within PDRs as the source locations for granule files, DPL Ingest will automatically refer to their definition to obtain time out and retry parameters.

In cases where a host has not been explicitly defined, the ingest operator will be able to define default time-out and retry parameters for SCP or FTP hosts. If a request is sent through processing with a host configured in the PDR that does not show up on the GUI, a new host will automatically be added to the list of SCP/FTP Hosts with the name UNDEFHOST_[Provider]_[RequestID]. Default host configuration parameters will be applied to the new host until the operator chooses to modify them.

Table 13.5-5. SCP/FTP Host Page Field Descriptions

Field Name	Entry	Description
Label	Required	A unique identifier for the host
Address	Required	An IP address or the canonical name and port (if needed) of an FTP host
Max Operations	Required	Total number of operations that can occur simultaneously on the host. If this field is left empty a default value will be supplied.
Timeout	Optional	Whether or not to allow a host to timeout if operations of a particular size take too much time to complete
Expected Throughput	Required if timeout is flagged	Expected amount of MBs of a granule to be processed during the configured pad time. If this field is left empty a default value will be supplied.
Pad Time	Required if timeout is flagged	Time an configured chunk of data should be processed before raising a timeout alert. If this field is left empty a default value will be supplied.
Auto Retry	Optional	Whether or not to retry an action that failed or generated an error on the host
Retry Interval	Required if Auto Retry is flagged	Time in between retries on the host. If this field is left empty a default value will be supplied.

13.5.3.1 Remove FTP or SCP Transfer Hosts

1 Click on the **Configuration** link in the navigation frame of the **DPL Ingest GUI.**

- The **Configuration** menu is expanded.

2 Click on the **Transfer Hosts** link in the navigation frame of the **DPL Ingest GUI.**

- **The Host Configuration** page (see Figure 13.5-6) is displayed.

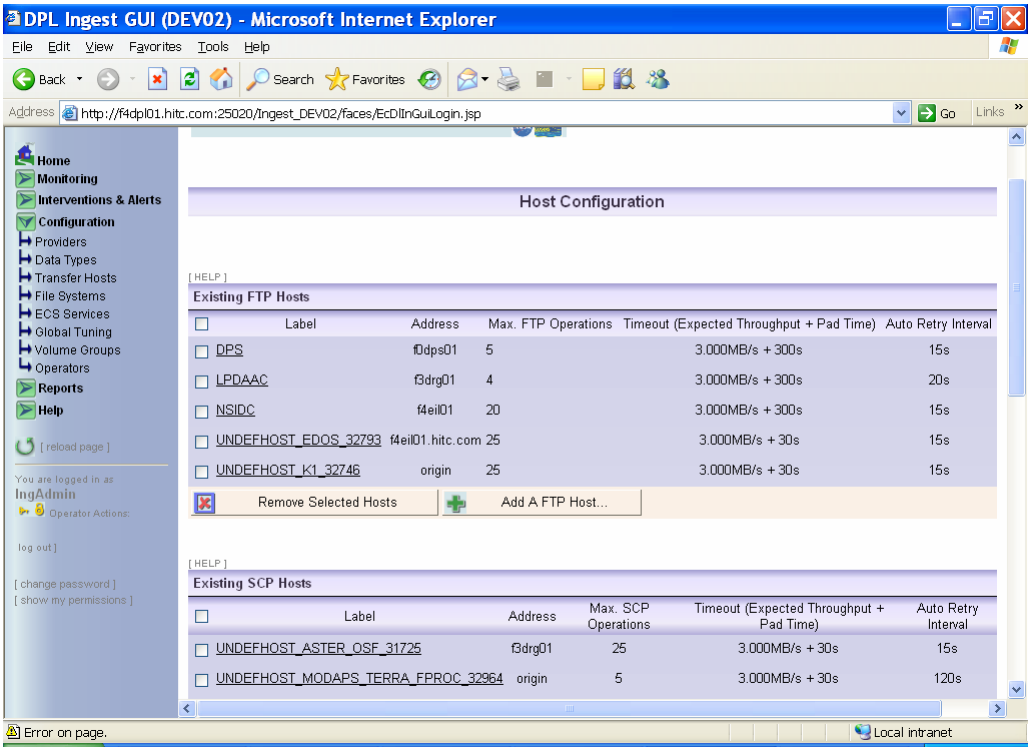


Figure 13.5-6. Host Configuration Page

- Displays **Existing FTP Hosts (first section).**
- Displays **Existing SCP Hosts (second section)**.
- Displays **Default FTP Host Configurations (third section)**.
- Displays **Default SCP Host Configurations (forth section)**.
- Displays **Local Host Configurations (fifth section)**.

3 Click the box next to the host name of the existing FTP or SCP Host(s) to be removed (multiple selections are accepted).

- A checkmark is plsced in the box.

4 Click on the **Remove Selected Hosts** button.

- A Confirmation prompt is displayed. Select **OK**
- The selected hosts are removed.

13.5.3.2 Add an FTP or SCP Transfer Host

1 Click on the **Configuration** link in the navigation frame of the **DPL Ingest GUI.**

- The **Configuration** menu is expanded.

2 Click on the **Transfer Hosts** link in the navigation frame of the **DPL Ingest GUI.**

- **The Host Configuration** page (see Figure 13.5-6) is displayed.
- Displays **Existing FTP Hosts (first section).**
- Displays **Existing SCP Hosts (second section)**.
- Displays **Default FTP Host Configurations (third section)**.
- Displays **Default SCP Host Configurations (forth section)**.
- Displays **Local Host Configurations (fifth section)**.

3 In the **Existing FTP Hosts** or **Existing SCP Hosts** sections**,** click on the **Add a FTP Host** or **Add a SCP Host** button.

- The **FTP Host Configuartion-add a new host** or **SCP Host Configuartion-add a new host** screen is (see Figure 13.5-7) displayed.

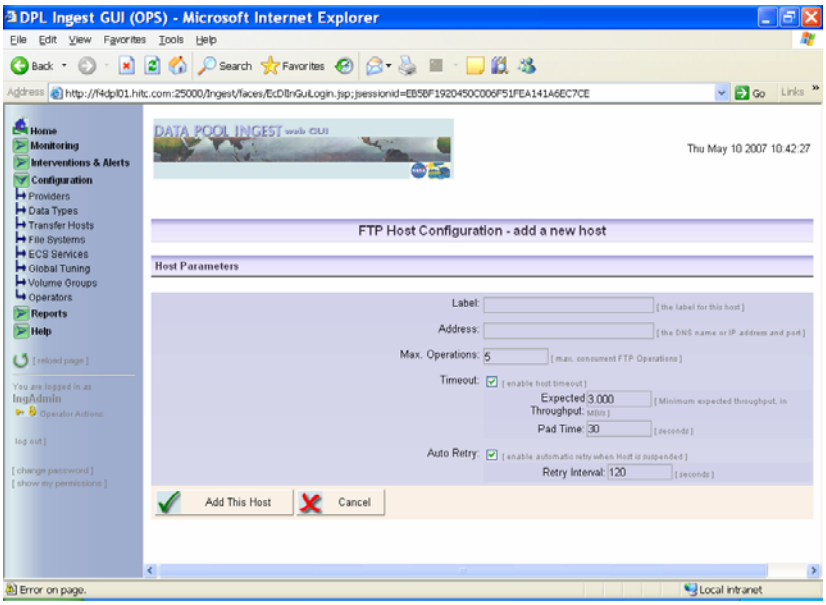


Figure 13.5-7. FTP Host Configuration Add a New Host Page

- FTP and SCP Hosts have similar (but slightly different fields).

4 Enter a unique name in the **Label** field.

- Existing names will be rejected by the database.

5 Enter the IP (e.g., 192.168.2.1) address or the canonical name (e.g., my.ftp.host) in the **Address** field. Place the port number on the same line, separated by a colon.

6 Enter the **Max. Operations** parameter.

- This value represents the maximum number of concurrent FTP or SCP operations this host may initiate.

7 Click on the **Timeout** box (optional field).

- If this box checked, text boxes will be displayed for the **Expected Throughput** (in mb/s) and **Fixed Overhead** (seconds) values: Enter these values.

8 Click on the **Auto Retry flag** (optional field).

- If this box checked, a textbox will be displayed to set the **Retry Interval** value (the number of minutes to wait between retries of this host if it becomes suspended by the server. Enter this value.

9 Select the **Add This Host** button at the bottom of the screen.

- The new entry will be displayed on the **FTP Host Configuration** page.

13.5.3.3 Edit an SCP or FTP Transfer Host

1 Click on the **Configuration** link in the navigation frame of the **DPL Ingest GUI.**

- The **Configuration** menu is expanded.

2 Click on the **Transfer Hosts** link in the navigation frame of the **DPL Ingest GUI.**

- **The Host Configuration** page is displayed.
- Displays **Existing FTP Hosts (first section).**
- Displays **Existing SCP Hosts (second section)**.
- Displays **Default FTP Host Configurations (third section)**.
- Displays **Default SCP Host Configurations (forth section)**.
- Displays **Local Host Configurations (fifth section)**.

3 In the **Existing FTP Hosts** or **Existing SCP Hosts** sections**,** click on the name of the desired host.

- The **Host Configuartion for [*XXXX*]** screen (see Figure 13.5-8) is displayed.

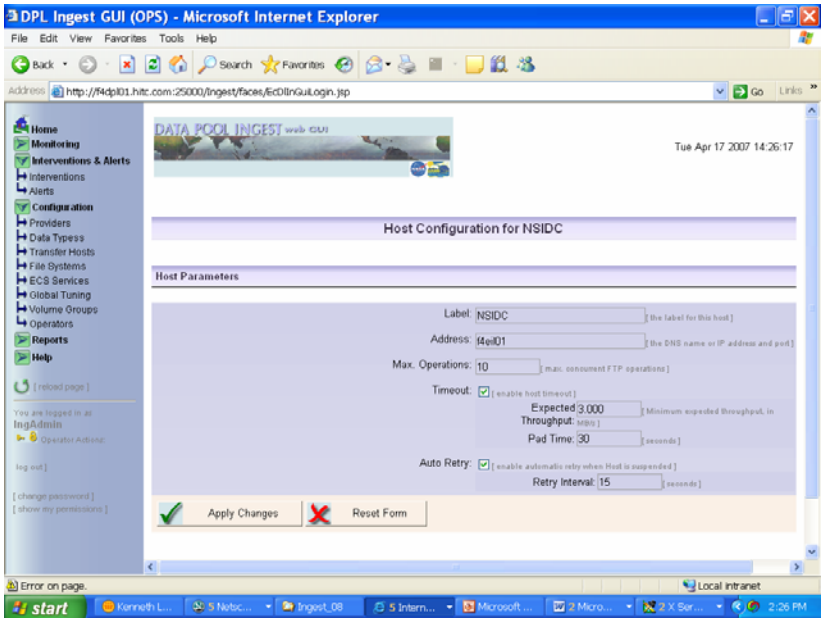


Figure 13.5-8. Host Configuration for DAAC Screen

- FTP and SCP Hosts have similar (but slightly different fields).

4 Edit the desired fields.

5 Select the **Apply Changes** button at the bottom of the screen.

- Changes will be implemented.

Local Host configuration parameters are used during any local transfer operations. The maximum Local operations limits how many local copies will occur concurrently. The timeout values apply to each individual local copy operation.

Default SCP and FTP Host configuration values are used to fill in default values whenever a new SCP or FTP host is added, or if a field is left empty when updating an existing SCP or FTP host.

13.5.3.4 Edit Local and Default [*FTP/SCP*] Host Configuration

1 Click on the **Configuration** link in the navigation frame of the **DPL Ingest GUI.**

- The **Configuration** menu is expanded.

2 Click on the **Transfer Hosts** link in the navigation frame of the **DPL Ingest GUI.**

- **The Host Configuration** page is displayed.
- Displays **Existing FTP Hosts (first section).**
- Displays **Existing SCP Hosts (second section)**.
- Displays **Default FTP Host Configurations (third section)**.

- Displays **Default SCP Host Configurations (forth section)**.
- Displays **Local Host Configurations (fifth section)**.

3 In the **Default FTP Host Configurations (third section)**, **Default SCP Host Configurations (forth section)**, or **Local Host Configurations (fifth section)** sections**,** click on the **Edit** button,

- The **Host Configuartion for FTPDEFAULT** screen is displayed (see Figure 13.5-9).

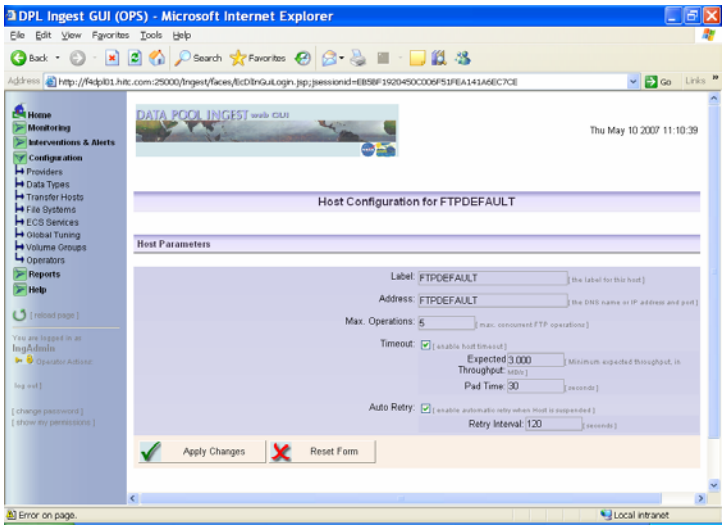


Figure 13.5-9. Host Configuartion for FTPDEFAULT

- FTP and SCP Hosts have similar screens.
- The **LOCAL Host Configuration** screen is displayed.

4 Edit the desired fields.

5 Select the **Apply Changes** button at the bottom of the screen.

- Changes will be implemented.

13.5.4 File System Configuration

The File System Configuration page allows the operator to configure warning and suspension thresholds (see Table 13.5-6 for any configured regular file system or archive file systems.)

Table 13.5-6. Archive File Systems Configuration Page Field Descriptions

Field Name	Entry	Description
Cache Warning Threshold	Required	The percentage of cache used which will trigger an operator alert. This must be below the Cache Full Threshold and above the Cache Warning Low Watermark.
Cache Full Threshold	Required	The percentage of cache used which will trigger an operator alert and suspend the Archive File System. This must be above the other threshold and watermarks.
Cache Warning Low Watermark	Required	The percentage of cache used that will clear the Archive Cache Warning Alert. This must be below the Cache Warning Threshold and the Cache Full Low Watermark.
Cache Full Low Watermark	Required	The percentage of cache used that will clear the Archive Cache Full Alert. This must be below the other watermark and thresholds.

13.5.4.1 Change File System Threshold

1 Click on the **Configuration** link in the navigation frame of the **DPL Ingest GUI.**

- The **Configuration** menu is expanded.

2 Click on the **File Systems** link in the navigation frame of the **DPL Ingest GUI.**

- The **File System Configuration** page is displayed (see Figure 13.5-10).

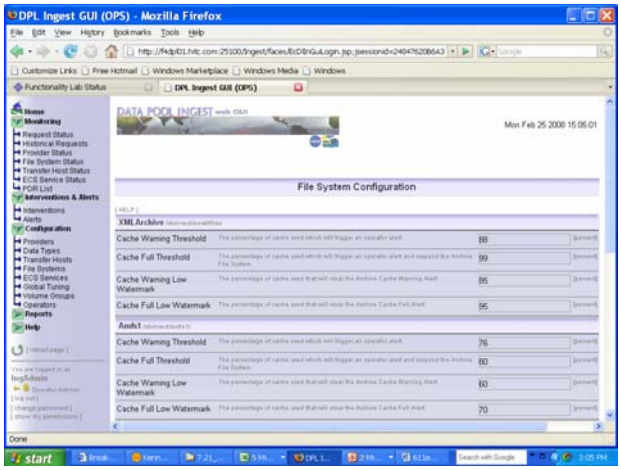


Figure 13.5-10. File System Configuartion

- Displays **Cache Warning Threshold** for each filesystem.
- Displays **Cache Full Threshold** for each filesystem**.**
- Displays **Cache Warning Low Watermark** for each filesystem**.**

- Displays **Cache Full Low Watermark** for each filesystem.

3 Enter the desired changes in the configurable fields.

- The change appears in the field.

4 Select the **Update** button at the bottom of the screen.

- The change is accepted.

13.5.5 ECS Service Configuration

This page allows the operator to configure the parameters of ECS services (see Tables 13.5-7 and 13.5-8) on a specific host. A default checksum type and algorithm can also be set for use by the checksumming service hosts. This page also allows the operator to select the host from which the Science Data Server will be run. This must be configured to ensure proper functionality of the DPL Ingest system.

Note: The **ECS Services Configuration** page is a listing page only. Modifications cannot be made from this page. The list shows which services are enabled for each host.

The **ECS Services Configuration** page contains the following three sections.

- Checksum Types and Algorithm. The operator can add, edit, and delete checksum types and their specific algorithms, and specify if the checksum type will be used as the default type.
- Host Used For SDSRV Operations. The operator can select a host on which Science Data Server operations will take place (such as metadata validation and SDSRV insert).
- Host Used For ECS Services. The operator can view and configure the attributes of the ECS Service host and can configure each of the services that run on that host.

Table 13.5-7. ECS Services Configuration Field Description

Field Name	Entry	Description
Name	n/a	The unique name given for this ECS Service Host
Address	n/a	The IP address and port of the host
Comments	n/a	Any descriptive comment text given for this host.
Max. Insert Checksum Operations	n/a	The maximum number of Insert Checksum Operations that will be performed by this host (checksum performed before archiving)
Max. Insert Copy Operations	n/a	The maximum Insert Copy operations that will be performed by this host.
The following are ECS Services		
Checksum	n/a	Each of these ECS Services are indicated as enabled (green checkmark) or disabled (red x) for each host.
File Transfer	n/a	
SCP	n/a	
Archive	n/a	

Table 13.5-8. ECS Services Configurable Field Descriptions (1 of 2)

Field Name	Entry	Description
Global Parameters		
Label	Required	A unique name for the ECS Service host, preferably based on the actual host name.
Address	Required	The IP address (e.g., 127.5.2.88) or canonical name (e.g., f4eil01.hitc.com) of the host.
Port	Required	The port number associated with this service. Hint: the port can be determined by looking at the Quickserver's configuration file.
Max. CPU Operations	Required	The total maximum number of checksum operations that may be performed on this host.
Auto Retry	Optional	Whether or not to automatically retry processing of actions for all services enabled on this host.
Comment	Optional	The description of the host and its services.
Checksum		
Enable this service	Optional	Whether or not to use this service.
Max. Concurrent Checksum Operations	Required if enabled	The maximum number of concurrent checksum operations that may be performed on this host at any one time.
Expected Throughput	Required if enabled	The expected data throughput for checksum operations. This is to identify stuck operations,
Checksum Timeout Pad Time	Required if enabled	The additional delay for a checksum operation before it is considered timed-out.

Table 13.5-8. ECS Services Configurable Field Descriptions (2 of 2)

Field Name	Entry	Description
File Transfer		
Enable this service	Optional	Whether or not to use this service.
Enable SCP	Optional	Whether or not to use SCP as the file transfer method.
Max. Concurrent File Transfers	Required if enabled	The maximum number of concurrent file transfers that may be executed on this host.
Archiving		
Enable this service	Optional	Whether or not to use this service.
Max. Concurrent Archive Operations	Required if enabled	The maximum number of concurrent archive operations that may be executed on this host.
Expected Throughput	Required if enabled	The expected data throughput for archive operations. This is to identify stuck operations.
Archive Timeout Pad Time	Required if enabled	The additional delay for an archive operation before it is considered timed-out.
Insert Checksum		
Max. Concurrent Insert Checksum Operations:	Optional	The maximum number of concurrent Insert Checksum operations that may be executed on this host.
Insert Copy		
Max. Concurrent Insert Copy Operations	Required	The maximum number of concurrent Insert Copy operations that may be executed on this host.
Expected Throughput	Required	The expected data throughput for Insert Copy operations. This is to identify stuck operations.
Insert Timeout Pad Time	Required	The additional delay for an Insert Copy operation before it is considered timed-out.

13.5.5.1 Remove Checksum Type

1 Click on the **Configuration** link in the navigation frame of the **DPL Ingest GUI.**

- The **Configuration** menu is expanded.

2 Click on the **ECS Services** link in the navigation frame of the **DPL Ingest GUI.**

- The **ECS Services Configuration** page is displayed (see Figure 13.5-11).

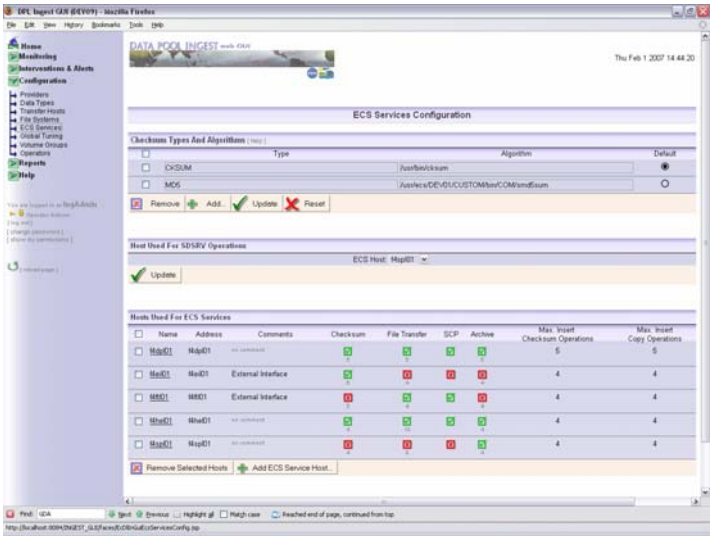


Figure 13.5-11. ECS Services Configuartion Page

- The **ECS Services Configuration** page is a view only page. Individual settings cannot be entered using this page.

3 In the **Checksum Types and Algorithms** section, click the box next to the checksum type to be removed (multiple selections are accepted).

- A checkmark is plsced in the box.

4 Click on the **Remove** button.

- A Confirmation prompt is displayed. Select **OK**
- The selected checksums are removed.

13.5.5.2 Add Checksum Type

1 Click on the **Configuration** link in the navigation frame of the **DPL Ingest GUI.**

- The **Configuration** menu is expanded.

2 Click on the **ECS Services** link in the navigation frame of the **DPL Ingest GUI.**

- The **ECS Services Configuration** page is displayed.
- The **ECS Services Configuration** page is a view only page. Individual settings cannot be entered using this page.

3 In the **Checksum Types and Algorithms** section, click on the **Add** button.

- The **New Checksum Type Information** screen is displayed in the white space portion of the **Checksum Types and Algorithms** section. Two input boxes are provided to enter **Type** and **Algorithm** information.

4 In the **New Checksum Type Information** section, click in the in the input box next to **Type.** Enter the Checksum **Type** information.

- The checksum type entered will be displayed in the **Type** field.

5 In the **New Checksum Type Information** section**,** click in the input box next to **Algorithm.** Enter the **Algorithm** information.

- The Algorithm entered will be displayed in the **Algorithm** field.

6 Click on the **ok** button.

- The new checksum type and algorithm will be added to the **ECS Services Configuration** page**.**

7 In the **Checksum Types and Algorithms** section of **ECS Services Configuration** page**,** review the default setting for the checksum**.** If the setting is not what you want, click on the desired **Checksum Default** button**.**

- The desired default **Checksum Types and Algorithms** will be selected.

13.5.5.3 Update Host Used For SDSVR Operations

1 Click on the **Configuration** link in the navigation frame of the **DPL Ingest GUI.**

- The **Configuration** menu is expanded.

2 Click on the **ECS Services** link in the navigation frame of the **DPL Ingest GUI.**

- The **ECS Services Configuration** page is displayed.
- The **ECS Services Configuration** page is a view only page. Individual settings cannot be entered using this page.

3 In the **Host Used For SDSVR Operations** section, select by highlighting the **ECS Host** from the pull-down window.

- The selected **ECS Host** will be displayed in the ECS Host window.

4 In the **Host Used For SDSVR Operations** section, click on the **Update** button.

- The new **ECS Host** will be displayed in the **ECS Host** window and will be added to the **ECS Services Configuration** page**.**

13.5.5.4 Add an ECS Service Host Type

1 Click on the **Configuration** link in the navigation frame of the **DPL Ingest GUI.**

- The **Configuration** menu is expanded.

2 Click on the **ECS Services** link in the navigation frame of the **DPL Ingest GUI.**

- **The ECS Services Configuration** page is displayed.

- The **ECS Services Configuration** page is a view only page. Individual settings cannot be entered using this page.

3 In the **Hosts Used For ECS Services**, Click on the **Add ECS Service Host** button.

- The **ECS Services Configuration: Add Service Host** page (see Figure 13.5-12 is displayed

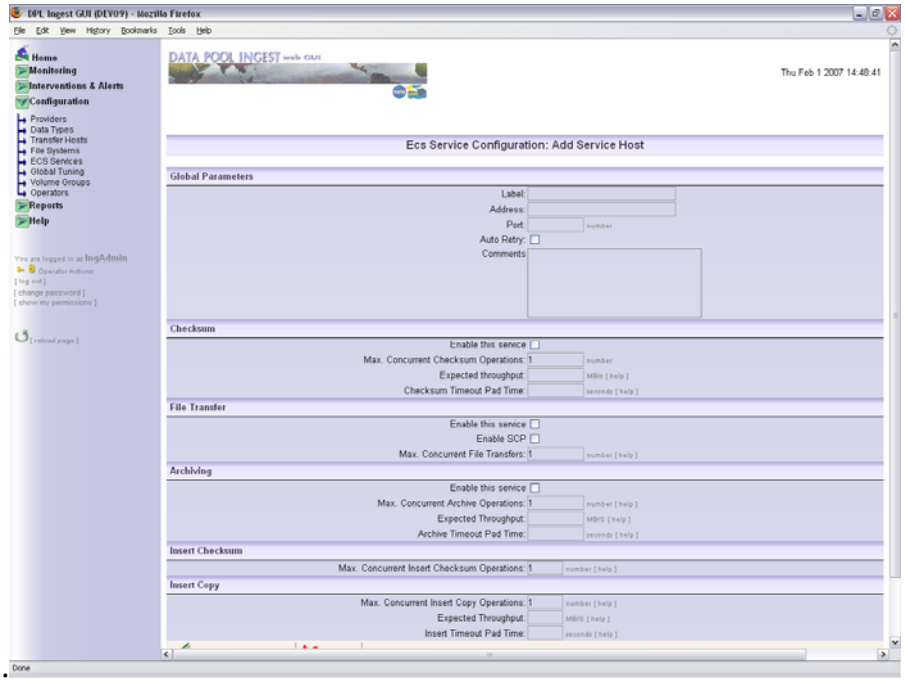


Figure 13.5-12. ECS Services Configuartion Add Service Host Page

4 In the **Global Parameters** section, click in the **Label** field. Enter a unique name for the **ECS Service** host.

- The name entered will be displayed in the **Label** field.

- Existing names will be rejected by the database.

5 In the **Global Parameters** section, click in the **Address** field. Enter the IP Address or the name of the **ECS Service** host.

- The address or name entered will be displayed in the Address field.

6 In the **Global Parameters** section, click in the **Port** field. Enter the port number associated with the service **ECS Service** host.

- The address or name entered will be displayed in the Address field.
- The port can be determind by looking at the quickserver's configuration file.

7 In the **Global Parameters** section, click in the **Auto Retry** (optional) field.

- A checkmark is placed in the box.
- All services enabled will automatically retry processing in the event of failure.

8 In the **Global Parameters** section, click in the **Comments** (optional) box. Enter a description of the host and its services.

- Information is displayed in the box.

9 In the **Checksum** section, click in the box next to the **Enable this service (optional)** field.

- A checkmark will be displayed in the box.

10 If **Checksum** is enabled, complete steps 11-13, otherwise, go to step 14.

11 In the **Checksum** section, click in the **Max. Concurrent Checksum Operations** field. Enter the maximum number of checksum operations that may be performed on this host at any one time.

- A number will be displayed in the box.

12 In the **Checksum** section, click in the **Expected Throughput** field. Enter the expected data throughput (MBs) for checksum operations. This will help to identify stuck operations.

- A number will be displayed in the box.

13 In the **Checksum** section, click in the **Checksum Timeout Pad Time** field. Enter additional delay for a checksum operation before it is considered timed-out.

- A number will be displayed in the box.

14 In the **File Transfer** section, click in the box next to the **Enable this service** (optional) field.

- A checkmark will be displayed in the box.

15 If **File Transfer** is enabled, complete steps 16-17, otherwise, go to step 18.

16 In the **File Transfer** section, click in the box next to the **Enable SCP** field.

- A checkmark will be displayed in the box.

17 In the **File Transfer** section, click in the **Max. Concurrent File Transfers** field. Enter the maximum number of concurrent file transfers that may be executed on this host at any one time.

- A number will be displayed in the box.

18 In the **Archiving** section, click in the box next to the **Enable this service** (optional) field.

- A checkmark will be displayed in the box.

19 If **Archiving** is enabled, complete steps 20-22, otherwise, go to step 23.

20 In the **Archiving** section, click in the **Max. Concurrent Archive Operations** field. Enter the maximum number of archive operations that may be performed on this host at any one time.

- A number will be displayed in the box.

21 In the **Archive** section, click in the **Expected Throughput** field. Enter the expected data throughput (MBs) for archive operations. This will help to identify stuck operations.

- A number will be displayed in the box.

22 In the **Archive** section, click in the **Archive Timeout Pad Time** field. Enter additional delay for an archive operation before it is considered timed-out.

- A number will be displayed in the box.

23 In the **Insert Checksum** section, click in the **Max. Concurrent Insert Checksum Operations** field.

24 Enter the maximum number of concurrent checksum operations that may be executed on this host at any one time.

- A number will be displayed in the box.

25 In the **Insert Copy** section, click in the **Max. Concurrent Insert Copy Operations** field. Enter the maximum number of concurrent Insert copy Operations that may be performed on this host at any one time.

- A number will be displayed in the box.

26 In the **Insert Copy** section, click in the **Expected Throughput** field. Enter the expected data throughput (MBs) for Insert Copy operations. This will help to identify stuck operations.

- A number will be displayed in the box.

27 In the **Insert Copy** section, click in the **Insert Copy Timeout Pad Time** field. Enter additional delay for an Insert Copy operation before it is considered timed-out.

- A number will be displayed in the box.

28 Select the **Add This Service Host** button at the bottom of the screen.

- A Confirmation prompt is displayed. Select **OK.**
- The new entry will be displayed on the **ECS Service Configuration** page.

13.5.5.5 Edit an ECS Service Host Type

1 Click on the **Configuration** link in the navigation frame of the **DPL Ingest GUI.**

- The **Configuration** menu is expanded.

2 Click on the **ECS Services** link in the navigation frame of the **DPL Ingest GUI.**

- **The ECS Services Configuration** page is displayed.

- The **ECS Services Configuration** page is a view only page. Individual settings cannot be entered using this page.

3 In the **Hosts Used For ECS Services**, click on the name of the **ECS Service Host** to be edited.

- The **ECS Services Configuration: Service Host Detail [Name]** page is displayed.

- Any or all parameters listed on the **ECS Services Configuration: Service Host Detail [Name]** page can be edited from this page..

4 In the **Global Parameters** section, enter the desired changes.

5 In the **Checksum** section, enter the desired changes.

6 In the **File Transfer** section, enter the desired changes.

7 In the **Archiving** section, enter the desired changes.

8 In the **Insert Checksum** section, enter the desired changes.

9 In the **Insert Copy** section enter the desired changes.

10 Select the **Apply Changes** button at the bottom of the screen.

- A Confirmation prompt is displayed. Select **OK**.

- Edited changes will be implemented.

13.5.6 Global Tuning Configuration

The Global Tuning link allows the operator to configure the Parameter Name and Value of the global tuning parameters in the Data Pool Ingest database. The parameters are listed along with their descriptions in Table 13.5-9.

There are three sections of the Global Tuning page. The first section titled Global Admin Tuning Parameters, consists of tuning parameters that can be edited by an operator with Ingest Admin Tuning privileges. The second section titled Global Tuning Parameter Configuration requires Tuning privileges. The third section titled Database Connection Configuration requires Tuning privileges. If the logged in operator does not have permission to edit a section, the fields and buttons for that section will be disabled.

Dynamic parameters are those that are applied to the Ingest Service without having to restart the Ingest Service. The Ingest Service will automatically apply these parameters within 1 minute of having been set in the database. Static parameters are those that require the Ingest Service to be restarted.

Table 13.5-9. Global Tuning Parameter Descriptions (1 of 4)

Parameter Name	Dynamic/ Static	Description
Global Admin Tuning Parameters:		
ARCHIVE_CACHE_CHECK_INTERVAL	Dynamic	Number of seconds between checks on archive cache.
DEFAULT_ALERT_RETRY_SECS	Dynamic	Default number of seconds to wait in between retrying a situation that caused a suspensio n
ENABLE_NOTIFICATION_PERFORMANCE_LOG	Dynamic	Indicates whether notification performance log has been enabled
ENABLE_POLLING_PERFORMANCE_LOG	Dynamic	Indicates whether polling performance log has been enabled
ENABLE_PROCESSING_PERFORMANCE_LOG	Dynamic	Indicates whether processing performance log has been enabled
GET_DPL_SPACE_MINS	Dynamic	Number of minutes to wait in between refreshing DPL free space info
MAX_RETRY_CHECKSUM_VERIFY	Dynamic	Maximum number of allowable retries for a checksum verification error
MINS_TO_KEEP_COMPLETED_REQS	Dynamic	Number of minutes before a completed request qualifies for archival
MONTHS_TO_KEEP_HIST_STATS_ALERTS	Dynamic	The retention time in months for keeping historic information for requests, alerts and throughput statistics
NUM_RETRIES_UR_ERROR	Dynamic	Number of times to retry UR Translation due to an error
RETRY_SECS_UR_ERROR	Dynamic	Number of seconds to wait in between retrying UR Translation on error

Table 13.5-9. Global Tuning Parameter Descriptions (2 of 4)

Parameter Name	Dynamic/ Static	Description
Global Tuning Parameter Configuration:		
DEFAULT_NUM_RETRIES	Dynamic	Default number of retries for an error condition where no error-specific number exists
DEFAULT_RETRY_INTERVAL	Dynamic	Default retry interval (seconds) where no error-specific interval exists
EDOS_SUCCESSFUL_PAN_DIR	Dynamic	EDOS directory in which successful PANs are to be stored
FAILED_CHECKSUM_HOLDING_DIR	Dynamic	Location of files that failed checksum verification. This directory needs to be monitored and when necessary, cleaned up.
IGNORE_ARCHIVE_ALERT	Dynamic	Still activate requests independent of archive status.
IGNORE_DPL_FS_DOWN	Dynamic	Indicates whether or not we activate requests that use a suspended file system
IIU_TIMEOUT_VALUE	Dynamic	The amount of time the server will wait for a request to IIU before assuming the request is lost
MAX_CONCURRENT_PREPROCESS	Dynamic	Maximum number of concurrent preprocessing operations
MAX_CONCUR_DPL_INSERT	Static	Maximum number of allowed concurrent DPL Insert processes for Ingest
MAX_CONCUR_IIU_PROCESSES	Dynamic	Maximum number of allowed concurrent IIU processes for Ingest
MAX_CONCUR_SCP_OPS	Dynamic	Max number of concurrent SCP ops
MAX_CONCUR_XVU_PROCESSES	Dynamic	Maximum number of allowed concurrent XVU processes for Ingest
MAX_CONSEC_FS_ERRORS	Dynamic	Maximum number of permissible file system access errors for different granules prior to raising an alert

Table 13.5-9. Global Tuning Parameter Descriptions (3 of 4)

Parameter Name	Dynamic/ Static	Description
MAX_CONSEC_SDSRV_ERRORS	Dynamic	Maximum number of permissible SDSRV errors of the same type for different granules prior to raising an alert
MAX_CONSEC_XFER_ERRORS	Dynamic	Maximum number of permissible transfer errors for different files (including PAN/PDRD) prior to raising an alert
MAX_CONSEC_XFER_ERRORS_PDR	Dynamic	Maximum number of permissible PDR transfer errors for different files prior to raising an alert
MAX_GRANS_WITH_SERV_ERR	Dynamic	Maximum of allowable number of the same type of error for an ECS service for different granules prior to raising an alert for that ECS Service
PROCESSING_MAX_GRANS	Dynamic	Maximum number of granules that can be in processing at once
PROCESSING_MAX_VOLUME	Dynamic	Maximum amount of data in MB that can be in processing at once
SDSRV_TIMEOUT_VALUE	Dynamic	The amount of time the server will wait for a request to the ScienceDataServer before assuming the request is lost
THROUGHPUT_STATS_INTERVAL	Dynamic	Number of minutes at which throughput statistics will be recorded
USE_SDSRV	Dynamic	Flag to determine whether to use the Science Data Server
VALIDATION_WARNINGS_DIR	Dynamic	Directory where we will store metadata files for granules which had validation warnings returned by the XVU. This directory needs to be monitored and when necessary, cleaned up. Note:the directory is not automatically cleaned up it must be maintained manually.
VALIDATION_WARNINGS_EMAIL	Dynamic	Comma separated list of email addresses to send granule validation warnings to.

Table 13.5-9. Global Tuning Parameter Descriptions (3 of 3)

Parameter Name	Dynamic/ Static	Description
XVU_TIMEOUT_VALUE	Dynamic	The amount of time the server will wait for a request to XVU before assuming the request is lost
Database Connection Configuration		
MAX_DPL_DB_CONN	Dynamic	Maximum database connection pool size for DPL DB
MAX_INGEST_DB_CONN	Dynamic	Maximum database connection pool size for Ingest DB
MAX_SDSRV_DB_CONN	Dynamic	Maximum database connection pool size for SDSRV DB
MAX_SSS_DB_CONN	Dynamic	Maximum database connection pool size for SSS DB
MIN_DPL_DB_CONN	Dynamic	Minimum database connection pool size for DPL DB
MIN_INGEST_DB_CONN	Dynamic	Minimum database connection pool size for Ingest DB
MIN_SDSRV_DB_CONN	Dynamic	Minimum database connection pool size for SDSRV DB
MIN_SSS_DB_CONN	Dynamic	Minimum database connection pool size for SSS DB

13.5.6.1 Change Global Tuning Parameters

1 Click on the **Configuration** link in the navigation frame of the **DPL Ingest GUI.**

- The **Configuration** menu is expanded.

2 Click on the **Global Tuning** link in the navigation frame of the **DPL Ingest GUI.**

- The **Global Tuning** page is displayed (see Figure 3.5-13).

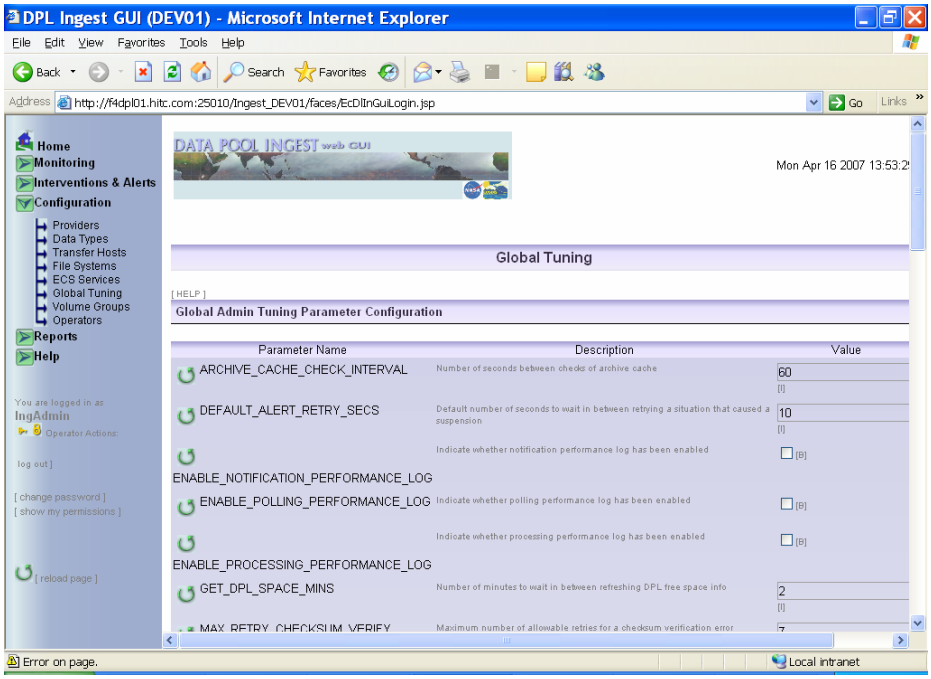


Figure 13.5-13. Global Tuning Page

- The **Global Tuning** page is displayed**.** You must have Ingest Admin or Tuning privileges to make changes on this page.
- The **Global Tuning** page divided into the following three sections:
  - **Global Admin Tuning Parameters.**
  - **Global Tuning Parameter Configuration.**
  - **Database Connection Configuration.**

3 Click in the **Value** field and enter the desired change for the selected parameter.

- The change is diaplayed as entered.

4 Click on the **Apply Changes** button found in each section.

- A Confirmation prompt is displayed. Select **OK.**
- The new changes will be applied.

NOTE: The directories for Failed_Checksum_Holding_DIR and Validation_Warning_DIR need to be routinely monitored. These directories require a manual cleanup.

13.5.7 Configure Volume Groups

The Volume Group configuration in the DPL Ingest GUI is the same functionality that existed in the STMGT GUI tab prior to release 7.20. The **Volume Groups Configuration** page (see Figure 13.5-14) displays the list of currently configured volume groups. This list is displayed all on one page and not broken into chunks. By default, the entries are listed alphabetically by Data Type Shortname. You can search for a desired data type by using the browser's built-in search function.

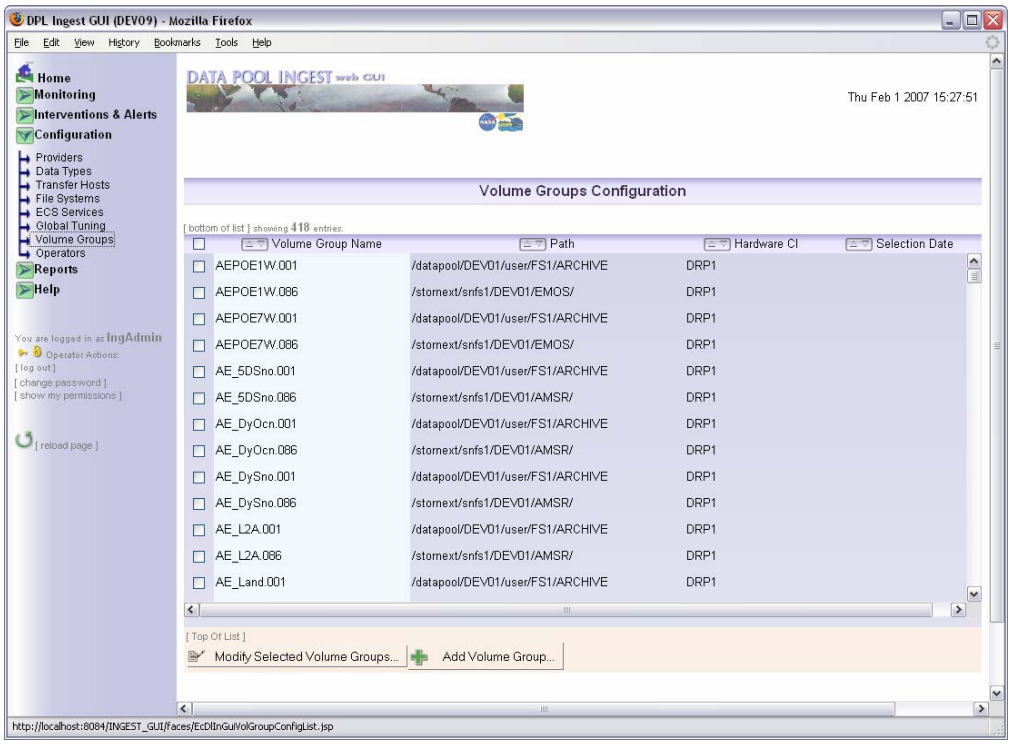


Figure 13.5-14. Volume Groups Configuration (listing page)

The **Volume Groups Configuration page** is divided into four distinct columns that provide important information about a particular **Volume Group**. The first column contains the **Volume Group** name. The second column shows the **Path**, the third column shows the **HWCI**, and the last column shows the **Selection Date** for that volume group, if applicable.

The bottom of the **Volume Groups Configuration** page has buttons to add a new volume group or to modify existing volume groups.

Table 13.5-10 contains a description of the fields contained on the **Volume Groups Configuration** page.

Table 13.5-10. Volume Groups Configuration Page Field Descriptions

Field Name	Entry	Description
Volume Group Name	System Generated	The name of the Volume Group based on a Data Type shortname with version identifier.
Path	System Generated	The fully qualified Unix path to where data is stored for the specified data type.
Hardware CI	System Generated	The label of the Archive silo group instance currently responsible for storing data of the specified data type.
Selection Date	System Generated	Non-NULL selection date defined for the ESDT version of which there are two volume group history sets defined for forward processing and reprocessing data respectively.
New Volume Group Path	Operator	A hidden field that is displayed when the operator clicks "Modify Selected Volume Groups".

An authorized Ingest Admin is authorized to add a Volume Group for a new Data Type version or add a Volume Group to an existing Data Type version. When adding a Volume Group for a new Data Type, the following rules apply:

- The Primary path information must be entered.
- The addition of **Backup Volume Group**, **Offsite Volume Group**, **or Alternative Volume Group History Set**, are optional and may be entered at a later time.

Table 13.5-11 provides a description of the fields contained on the **Volume Groups Configuration: Add a Volume Group** screen**.**

Table 13.5-11. Add Volume Group Page Field Description (1 of 3)

Field Name	Data Type	Size	Entry	Description
Data Type and Version ID	Character	16	Required	A Data Type short name and version identifier.
Alternative VG Options	Check box	N/A	Not Required	Allows operator to enter options for alternative Volume Groups. This can only be checked if an Alternative Volume Group was specified, otherwise, the checkbox is disabled.

Table 13.5-11. Add Volume Group Page Field Description (2 of 3)

Field Name	Data Type	Size	Entry	Description
Selection Date for Alternative Volume Groups	Character	8	Required if adding Alternative Volume Group History Set	When the Alternative VG check box is selected, the Selection Date section is enabled and is required to be filled out by the user. Selection Date is a separate date to guide Archive Server to select a appreciate Volume Group History set for storing / retrieving data. When acquisition date is not null and less than the Selection Date, Reprocessing Volume Group history set will be used, otherwise, forward processing Volume Group history set will be used.
Reprocessing, Forward Processing	Radio Buttons	N/A	Required if adding Alternative Volume Group History Set	Alternative volume groups can be configured either for Reprocessing or even for Forward Processing. Default is for Reprocessing. Although the flexibility to add a new alternative for forward processing is supported, it should be used with a great caution.
Primary Volume Group Path:	Character	Unlimited	Required	The fully-qualified Unix path to where data is currently being stored for the specified data type to the Primary Archive.
Backup Volume Group Path:	Character	Unlimited	Required if Backup enabled	The fully-qualified Unix path to where data is currently being stored for the specified data type to the Backup Archive.
Offsite Volume Group Path:	Character	Unlimited	Required if Offsite enabled	The fully-qualified Unix path to where data is currently being stored for the specified data type to the Offsite Archive.
Primary Alternative Volume Group Path:	Character	Unlimited	Required if Primary Alternative enabled	The fully-qualified Unix path to where reprocessing data is currently being stored for the specified data type to the Primary Alternative Archive.

Table 13.5-11. Add Volume Group Page Field Description (3 of 3)

Field Name	Data Type	Size	Entry	Description
Backup Alternative Volume Group Path:	Character	Unlimited	Required if Backup Alternative enabled	The fully-qualified Unix path to where data is currently being stored for the specified data type to the Backup Alternative Archive.
Offsite Alternative Volume Group Path:	Character	Unlimited	Required if Offsite Alternative enabled	The fully-qualified Unix path to where data is currently being stored for the specified data type to the Offsite Alternative Archive.

When a Volume Group is added, the name will be created based on the type of Volume Group that was added. There are six types, as explained in Table 13.5-12. Note that "R" indicates an alternative Volume Group for reprocessing. There is no explicit suffix for forward processing.

Table 13.5-12. Volume Group Naming

Volume Group Type	Extension	Example
Primary	none	AST_L1B.003
Primary Alternative	R	AST_L1B.003R
Backup	B	AST_L1B.003B
Backup Alternative	BR	AST_L1B.003BR
Offsite	O	AST_L1B.003O
Offsite Alternative	OR	AST_L1B.003OR

13.5.7.1 Add a Volume Group for a New Versioned Data Type

1 Click on the **Configuration** link in the navigation frame of the **DPL Ingest GUI.**

- The **Configuration** menu is expanded.

2 Click on the **Volume Groups** link in the navigation frame of the **DPL Ingest GUI.**

- The **Volume Groups Configuration** page is displayed (see Figure 13.5-15).

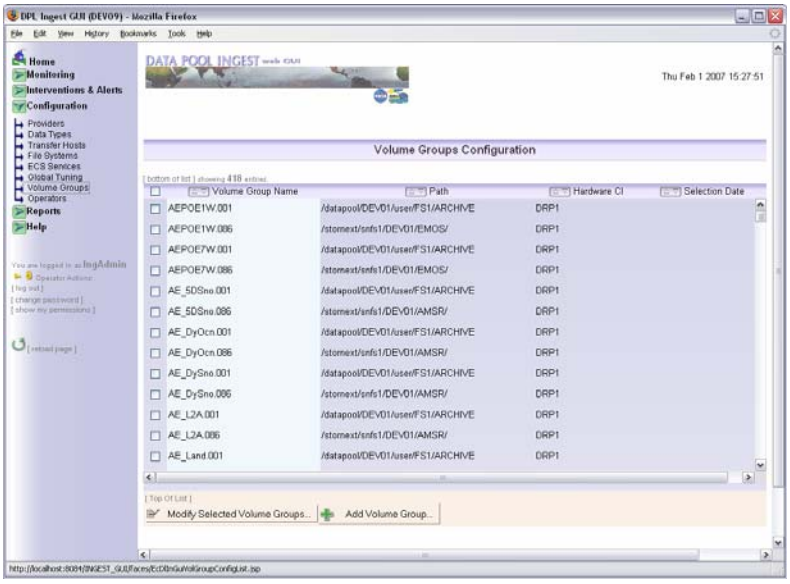


Figure 13.5-15. Volume Group Configuartion Page

3 Scroll to the bottom of the page and select the **Add Volume Groups** button.

- The **Volume Groups Configuration: Add Volume Group** page is displayed (see Figure 13.5.16).

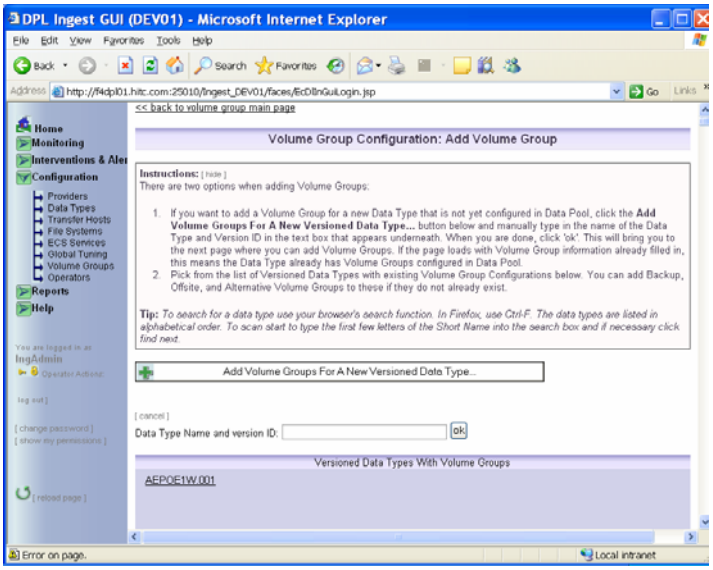


Figure 13.5-16. Volume Group Configuartion Page

4 Click on the **Add Volume Groups For a New Versioned Data Type.**

- The **Data Type and version ID** field is displayed.

5 Click in the **Data Type and version ID** field and enter the new **Data Type and Version ID.** Select **ok.**

- The **Volume Group Configuration: Add a Volume Group** page is displayed (see Figure 13.5-17).

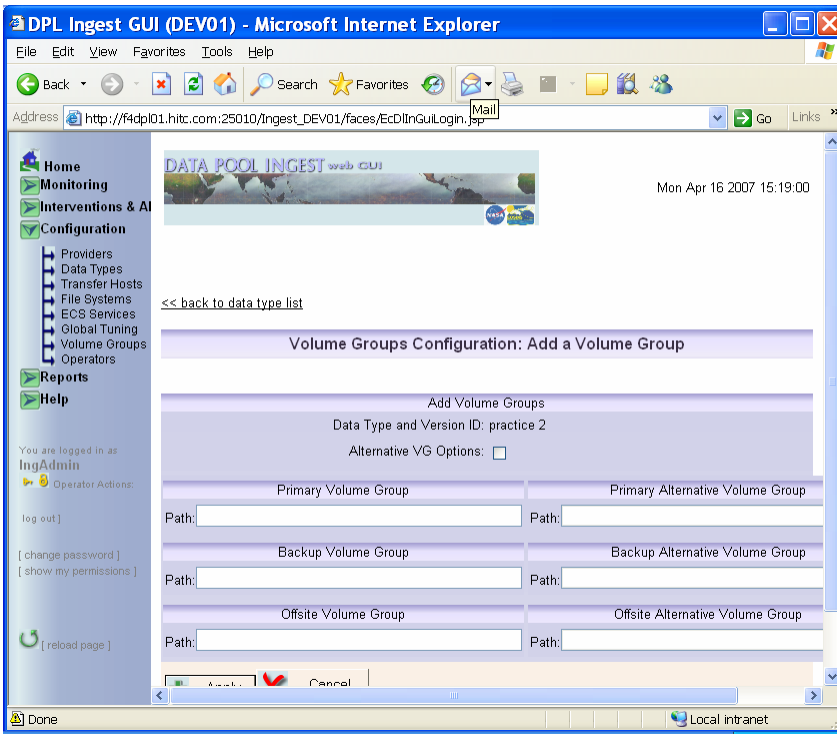


Figure 13.5-17. Add Volume Group Configuartion Page

6 (Optional) Click on the **Alternate VG Options** box.

- A checkmark is displayed in the box.

7 (Optional) Click on the **Reprocessing** or **Forward Processing** button.

- A mark is displayed in the selected button.

8 (Optional) If the **Alternate VG Options**: box is selected, you must enter the date in the provided **Selection Date for alternate Volume Groups** area.

- A date is displayed in the selected area provided.

9 Click in the **Primary Volume Group Path** edit window and enter the fully-qualified Unix path to where data is currently being stored for the specified new data type to the Primary Archive.

- Data is displayed in the edit window.

10 Click in the **Primary Alternative Volume Group Path** edit window and enter the fully-qualified Unix path to where data is currently being stored for the specified new data type. to the Primary Alternate Archive,

- Data is displayed in the edit window.

11 Click in the **Backup Volume Group Path** edit window and enter the fully-qualified Unix path to where data is currently being stored for the specified new data type to the Backup Archive.

- Data is displayed in the edit window.

12 Click in the **Backup Alternative Volume Group Path** edit window and enter the fully-qualified Unix path to where data is currently being stored for the specified new data type. to the Backup Alternate Archive.

- Data is displayed in the edit window.

13 Click in the **Offsite Volume Group Path** edit window and enter the fully-qualified Unix path to where data is currently being stored for the specified new data type to the Offsite Archive.

- Data is displayed in the edit window.

14 Click in the **Offsite Alternative Volume Group Path** edit window and enter the fully-qualified Unix path to where data is currently being stored for the specified new data type. to the Offsite Alternate Archive.

- Data is displayed in the edit window.

15 Click in the **Apply** button.

- A Confirmation prompt is displayed. Select **OK**
- The changes are applied.

The following rules apply when adding Volume Groups to an existing Data Type version (e.g., Backup, Offsite, etc.):

- The **Volume Group** name will be selected from the **Primary Volume Groups** page. When the **Add Volume Group** page is loaded, the Volume Group name will appear at the top.

- Any previously added **Volume Group** will be displayed, but not editable. For example, if a **Backup Volume Group** has already been added, the **Volume Group** path will be shown, but the operator will not be able to edit this path.
- Similarly, if any **Alternative Volume Groups** have been specified, the **Alternative VG** options and **Volume Groups** will be displayed, but not editable.
- If the operator is adding the **Alternative Volume Group History Set** for the first time, the **Alternative Options** must be selected and the operator may choose the processing type (Forward Processing or Reprocessing) for the **Alternative Volume Group History Set**, as well as a selection date to be applied to the **Reprocessing Volume Groups.**

13.5.7.2 Add a Volume Group for an Existing Versioned Data Type

1 Click on the **Configuration** link in the navigation frame of the **DPL Ingest GUI.**

- The **Configuration** menu is expanded.

2 Click on the **Volume Groups** link in the navigation frame of the **DPL Ingest GUI.**

- The **Volume Groups Configuration** page is displayed.

3 Scroll to the bottom of the page and select the **Add Volume Groups** button.

- The **Volume Group Configuration: Add Volume Group** page is displayed (see Figure 13.5-18).

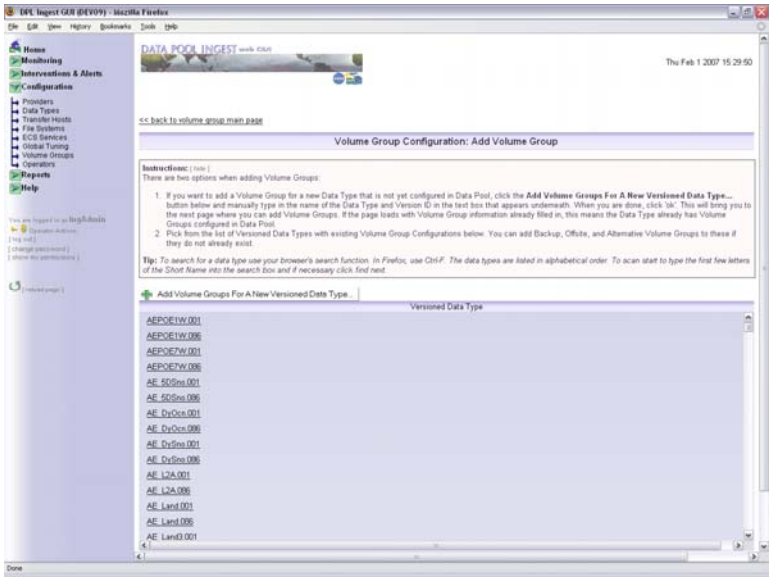


Figure 13.5-18. Volume Group Configuration: Add Volume Group

4 Click on the desired existing **Volume Group Name.**

- The **Volume Groups Configuration: Add a Volume Group** page is displayed for the Data Type selected.

5 (Optional) Click on the **Alternate VG Options**: box.

- A checkmark is displayed in the box.

6 (Optional) Click on the Reprocessing or Forward Processing button.

- A mark is displayed in the selected button.

7 (Optional) If the **Alternate VG Options**: box is selected, you must enter the date in the provided **Selection Date for alternate Volume Groups** area.

- A date is displayed in the selected area provided.

8 Click in the **Primary Volume Group Path** edit window and enter the fully-qualified Unix path to where data is currently being stored for the specified new data type to the Primary Archive.

- Data is displayed in the edit window.

9 Click in the **Primary Alternative Volume Group Path** edit window and enter the fully-qualified Unix path to where data is currently being stored for the specified new data type. to the Primary Alternate Archive,

- Data is displayed in the edit window.

10 Click in the **Backup Volume Group Path** edit window and enter the fully-qualified Unix path to where data is currently being stored for the specified new data type to the Backup Archive.

- Data is displayed in the edit window.

11 Click in the **Backup Alternative Volume Group Path** edit window and enter the fully-qualified Unix path to where data is currently being stored for the specified new data type. to the Backup Alternate Archive.

- Data is displayed in the edit window.

12 Click in the **Offsite Volume Group Path** edit window and enter the fully-qualified Unix path to where data is currently being stored for the specified new data type to the Offsite Archive.

- Data is displayed in the edit window.

13 Click in the **Offsite Alternative Volume Group Path** edit window and enter the fully-qualified Unix path to where data is currently being stored for the specified new data type. to the Offsite Alternate Archive.

- Data is displayed in the edit window.

14 Click in the **Apply** button.

- A Confirmation prompt is displayed. Select **OK**
- The changes are applied.

13.5.7.3 Modify Volume Groups

1 Click on the **Configuration** link in the navigation frame of the **DPL Ingest GUI.**

- The **Configuration** menu is expanded.

2 Click on the **Volume Groups** link in the navigation frame of the **DPL Ingest GUI.**

- The **Volume Groups Configuration** page is displayed.
- All columns on the **Volume Groups Configuration** page can be sorted in ascending or descending order. To sort on a column, click on the up or down arrow at the top of the column.

3 Click in the box next to the **Volume Group Name** of the desired volume group(s) to be change.

- A checkmark is placed in the box.
- Multiple selections may be made.

4 Scroll to the bottom of the page and select the **Modify Selected Volume Groups** button.

- A path input field appears at the bottom of the page.

5 Enter the **New Volume Group Path** in the edit field and click the **ok** button.

- A Confirmation prompt is displayed. Select **OK**
- The changes to the Volume Group Path will be applied.

13.5.8 Operator Configuration

The Operator Configuration page (see Figure 13.5-19) consists of a list of operator names and their current permission settings. The security operator configures authorized users for the Data Pool Ingest GUI. and add, edit, or remove users.

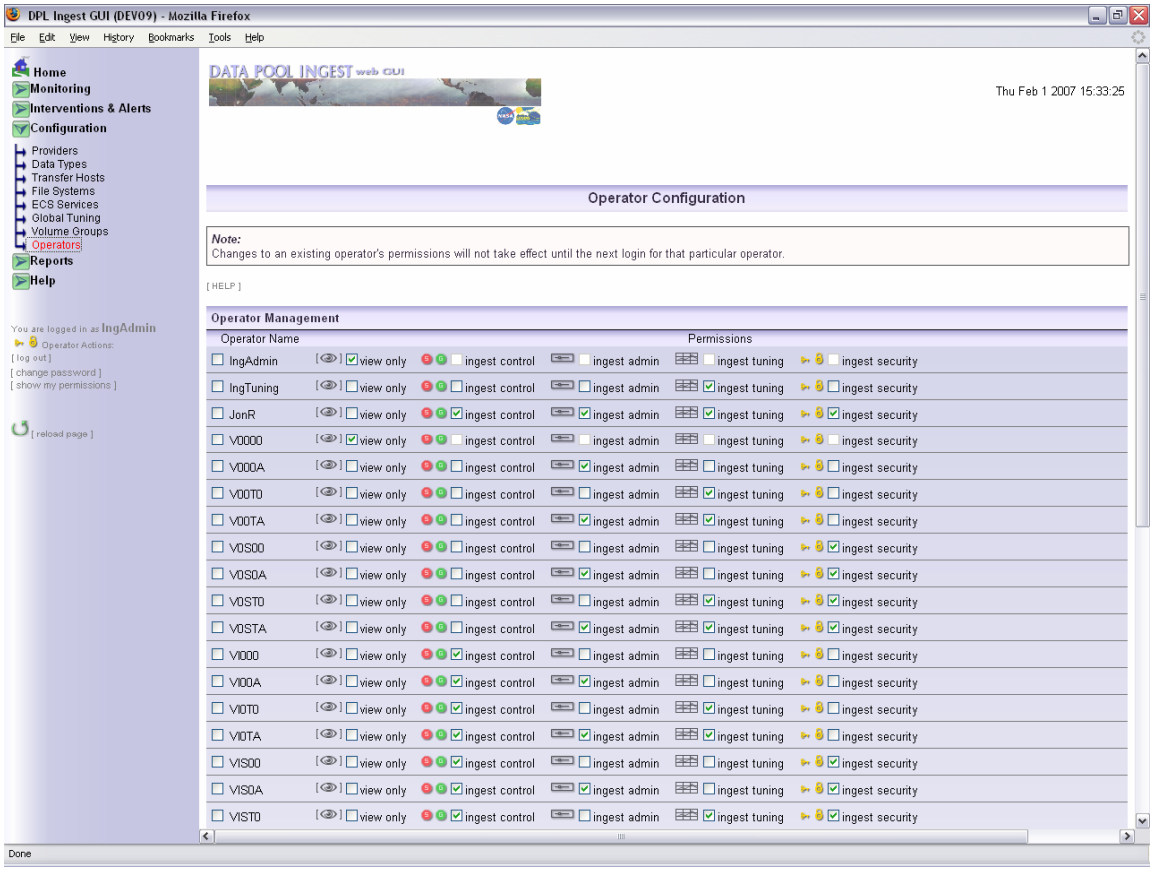


Figure 13.5-19. Operator Configuration Page

There are five different permission levels. An operator assigned the view only permission level, cannot be authorized additional permissions. The other 4 levels can be added together as they represent the ability to manage an exclusive set of properties associated with data pool ingest. An operator may be assigned multiple permissions other than view only. The following list reviews the five permission levels for the Data Pool Ingest GUI.

- **View Only** – The operator cannot alter or modify anything on the GUI, nor can he/she take actions. All textboxes, checkboxes, drop-down lists, etc. are disabled.
- **Ingest Control.** The operator can manage Ingest requests and interventions, i.e., he/she has the ability to suspend or resume requests, place on hold and close interventions, fail or resume granules, etc. This also implies that the operator may suspend and resume services, hosts, file systems, archives, etc.
- **Ingest Admin**. The operator can alter general configuration parameters such has SCP/FTP Host configuration, providers, data types, etc. This level of operator cannot modify tuning parameters.

- **Ingest Tuning**. The operator can alter global and host-specific tuning configuration parameters.
- **Security Admin**. The operator can maintain security-related information like passwords and operators.

13.5.8.1 Modify Operator Permission Settings

1 Click on the **Configuration** link in the navigation frame of the **DPL Ingest GUI.**

- The **Configuration** menu is expanded.

2 Click on the **Operators** link in the navigation frame of the **DPL Ingest GUI.**

- The **Operator Configuration** page is displayed.
- All Operators and their current Permission settings are displayed.
- Changes to this page can only be made if you have **Security Admin** permissions.

3 Click in the box next to the **Operator Name** of the desired operator to change permissions.

- A checkmark is placed in the box.
- Multiple selections may be made.
- A checkmark must be visible befor further changes can be made.

4 Click in the box next to the desired permission.

- A checkmark is placed in the box.

5 Scroll to the bottom of the **Operator Management** section and click on the **Update Operators** button.

- A Confirmation prompt is displayed. Select **OK.**
- Changes to an existing operator's permissions will not take effect until the next login by that particular operator.

13.5.8.2 Add Operator Permissions

1 Click on the **Configuration** link in the navigation frame of the **DPL Ingest GUI.**

- The **Configuration** menu is expanded.

2 Click on the **Operators** link in the navigation frame of the **DPL Ingest GUI.**

- The **Operator Configuration** page is displayed.
- All Operators and their current Permission settings are displayed.

- Changes to this page can only be made if you have **Security Admin** permissions.

3 Scroll to the bottom of the **Operator Configuration** page until you see the **Add Operator** section of the page.

4 Click in the **Operator Name** field and enter the name of the operator**.**

- The entered name is displayed.

5 Click in the **Password** field and enter the password**.**

- The **Password** field will be populated with stars.

6 Click in the **Verify Password** field and re-enter the password**.**

- The **Verify Password** field will be populated with stars.

7 Click in the box next to the desired **Permissions**.

- A checkmark is displayed.
- At least one permission level must be selected.

8 Select the **Add Operator** button.

- A Confirmation prompt is displayed. Select **OK.**
- The new operator name will be added to the list of operators in the **Operator Management** section of the **Operator Configuration Page.**

13.5.8.3 Remove Operator Permission Settings

1 Click on the **Configuration** link in the navigation frame of the **DPL Ingest GUI.**

- The **Configuration** menu is expanded.

2 Click on the **Operators** link in the navigation frame of the **DPL Ingest GUI.**

- The **Operator Configuration** page is displayed.
- All Operators and their current Permission settings are displayed.
- Changes to this page can only be made if you have **Security Admin** permissions.

3 Click in the box next to the **Operator Name** to be removed.

- A checkmark is placed in the box.
- Multiple selections may be made.
- A checkmark must be visible befor further changes can be made.

4 Scroll to the bottom of the Operator Management screen and click on the **Remove Operators** button.

- A Confirmation prompt is displayed. Select **OK**

- Changes to an existing operator's permissions will not take effect until the next login by that particular operator.

13.6 Reports

13.6.1 Reports

The reporting capability of the Ingest GUI offers the ability to view detailed reports on data providers and data types, as well as request summary and granule summary reports. The report pages are located under the **Reports** menu in the navigation pane.

The report pages (**Detailed**, **Request Summary** and **Granule Summary**) display the information across data providers or data types. As with all types of reports, the operator must select a date range (presets are provided for the last 24 and 48 hours), as well as criteria for the search. These include one or more **Data Providers**, one or more **Data Types**, and one or more **Final Request Statuses.** Additionally, **Ingest Method (DPL or NON-DPL)** can also be selected. All Data Criteria fields are optional, but at least one selection of one field must be made to generate the report. Due to the large volume of data that may be in the database, reports can sometime take a while to process and be displayed.

Table 13.6-1 provides an activity Checklist for Reports.

Table 13.6-1. Reports

Order	Role	Task	Section
1	Ingest Technician	Generating a Report	(P) 13.6.1.1
2	Ingest Technician	Viewing Volume Group(s) History	(P) 13.6.2.1

13.6.1.1 Generating a Detailed Report

1 Click on the **Reports** link in the navigation frame of the **DPL Ingest GUI.**

- The **Configuration** menu is expanded.

2 Click on the **Detailed Report** , **Request Summary** or **Granule Summary** in the navigation frame of the **DPL Ingest GUI.**

- **The Detailed Report, Request Summary Report** or **Granule Summary Report** setup page is diaplayed.

3 In the **Date/Time Criteria** portion of the screen, select the desired **Start Date/Time** and **End Date/Time.**

- The date and time information is displayed as selected.

4 In the **Data Criteria** portion of the screen, select one or more **Data Providers**, one or more **Data Types**, one or more **Final Request Statuses** and the **Ingest Method** (DPL or NON-DPL) by clicking on the the desired selections.

- The desired selections are highlighted.

5 Click on the **Generate Report** button.

- The **Processing Your Request** transitional screen is displayed.
- Time to process your request will depend on factors such as time span, number of Data Providers, Data Types and Request Statuses selected.
- Eventually, one of the following report layouts will be displayed.
- Detailed Report Layout (see Figure 13.6-1)

Current Report Criteria:

Data Provider(s): [ALL] Data Type(s): [ALL]
Final Request Status: [ALL] Start Date/Time: 1/11/2006 15:18
End Date/Time: 31/10/2006 15:18

Report Output

Req.ID	Data Provider	Ingest Type	Ingest Method	Start Date/Time	End Date/Time	Tot.# grans.	# Succ. grans.	Vol (MB)	File Count	Time to xfer (mins)	Time to preproc (min)	Time to Archive (min)	Priority	Restart Flag
21807	S4P00	Polling_w/DR	DPL	2006-11-01 08:21:07	2006-11-01 08:21:19	1	1	6.311	2	0	8	0	VHIGH	
21808	S4P00	Polling_w/DR	DPL		2006-11-01 08:29:06	1		6.245	2	0	0	0	VHIGH	
21809	ASTER_OSF	Polling_w/DR	DPL	2006-11-01 10:03:09	2006-11-01 10:03:10	1		0.473	2	0	0	0	NORMAL	
21810	ASTER_OSF	Polling_w/DR	DPL	2006-11-01 10:03:09	2006-11-01 10:03:10	1		0.473	2	0	0	0	NORMAL	
21811	ASTER_OSF	Polling_w/DR	DPL	2006-11-01 10:03:09	2006-11-01 10:03:10	1		0.473	2	0	0	0	NORMAL	

Figure 13.6-1. Detailed Report Layout

- Request Summary Report Layout (see Figure 13.6-2)

Current Report Criteria:

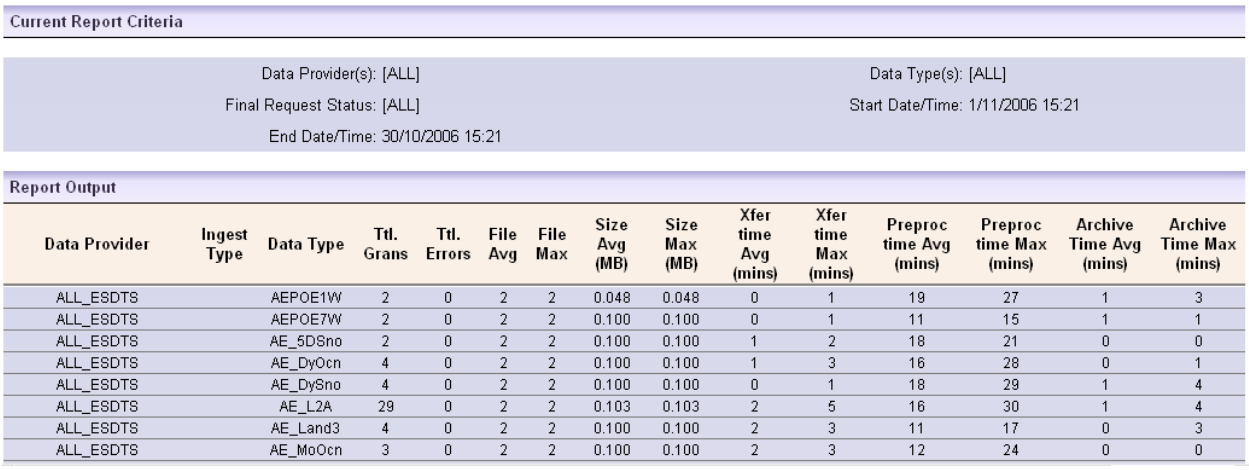
Data Type(s): [ALL] Final Request Status: [ALL]
Start Date/Time: 1/11/2006 15:21 End Date/Time:31/10/2006 15:21

Report Output

Data Provider	Ingest Type	Ttl. Reqs	Ttl. Errors	Gran Avg	Gran Max	File Avg	File Max	Size Avg (MB)	Size Max (MB)	Xfer time Avg (mins)	Xfer time Max (mins)	Preproc time Avg (mins)	Preproc time Max (mins)	Archive Time Avg (mins)	Archive Time Max (mins)
ASTER_OSF		12	0	1	1	2	2	0.473	0.473	10	126	3	14	0	1
CRIT_4150_2		10	0	1	1	2	2	0.473	0.473	0	1	1	2	0	0
MODAPS_TERRA_FPROC		1	0	1	1	2	2	0.473	0.473	0	0	2	2	0	0
S4P00		57	0	1	1	2	3	16.157	71.236	0	2	2	8	0	1

Figure 13.6-2. Request Summary Report Layout

- Granule Summary Report Layout (see Figure 13.6-3)



Current Report Criteria

Data Provider(s): [ALL]
Data Type(s): [ALL]
Final Request Status: [ALL]
Start Date/Time: 1/11/2006 15:21
End Date/Time: 30/10/2006 15:21

Report Output

Data Provider	Ingest Type	Data Type	Ttl. Grans	Ttl. Errors	File Avg	File Max	Size Avg (MB)	Size Max (MB)	Xfer time Avg (mins)	Xfer time Max (mins)	Preproc time Avg (mins)	Preproc time Max (mins)	Archive Time Avg (mins)	Archive Time Max (mins)
ALL_ESDTS		AEPOE1W	2	0	2	2	0.048	0.048	0	1	19	27	1	3
ALL_ESDTS		AEPOE7W	2	0	2	2	0.100	0.100	0	1	11	15	1	1
ALL_ESDTS		AE_5DSno	2	0	2	2	0.100	0.100	1	2	18	21	0	0
ALL_ESDTS		AE_DyOcn	4	0	2	2	0.100	0.100	1	3	16	28	0	1
ALL_ESDTS		AE_DySno	4	0	2	2	0.100	0.100	0	1	18	29	1	4
ALL_ESDTS		AE_L2A	29	0	2	2	0.103	0.103	2	5	16	30	1	4
ALL_ESDTS		AE_Land3	4	0	2	2	0.100	0.100	2	3	11	17	0	3
ALL_ESDTS		AE_MoOcn	3	0	2	2	0.100	0.100	2	3	12	24	0	0

Figure 13.6-3. Granule Summary Report Layout

13.6.2 Viewing the Volume Groups History Page

The **Volume Groups History** page displays the history of the configuration changes that have occurred to volume groups. Table 13.6-2. displays the information contained on the **Volume Groups History** page.

Table 13.6-2. Volume Groups History Page Field Description

Field Name	Data Type	Size	Entry	Description
Volume Group (Data Type. Version ID + Volume Group Type Suffix)	Character	16	Required	The name of a Volume Group.
Path History	Character	Unlimited	System Generated	In reverse chronological order, the fully qualified Unix paths to where data has been stored for the specified data type. The current path is listed first.
Hardware CI History	Character	16	System Generated	The label of the Archive silo group instance that was responsible for storing data of the specified data type.
Start Date	Date	16	System Generated	The date on which this configuration became active for the listed data type.
End Date	Date	16	System Generated	The date on which this configuration was superseded by new configuration information. If blank, this row reflects the current configuration for the volume group. If any row has a blank end date, the volume group is closed, and no further data is accepted for that volume group.

13.6.2.1 Viewing Volume Group(s) History

1 Click on the **Reports** link in the navigation frame of the **DPL Ingest GUI.**

- The **Reports** menu is expanded.

2 Click on the **Volume Group History** link in the navigation frame of the **DPL Ingest GUI.**

- The **Volume Group History** page is displayed (see Figure 13.6-4).

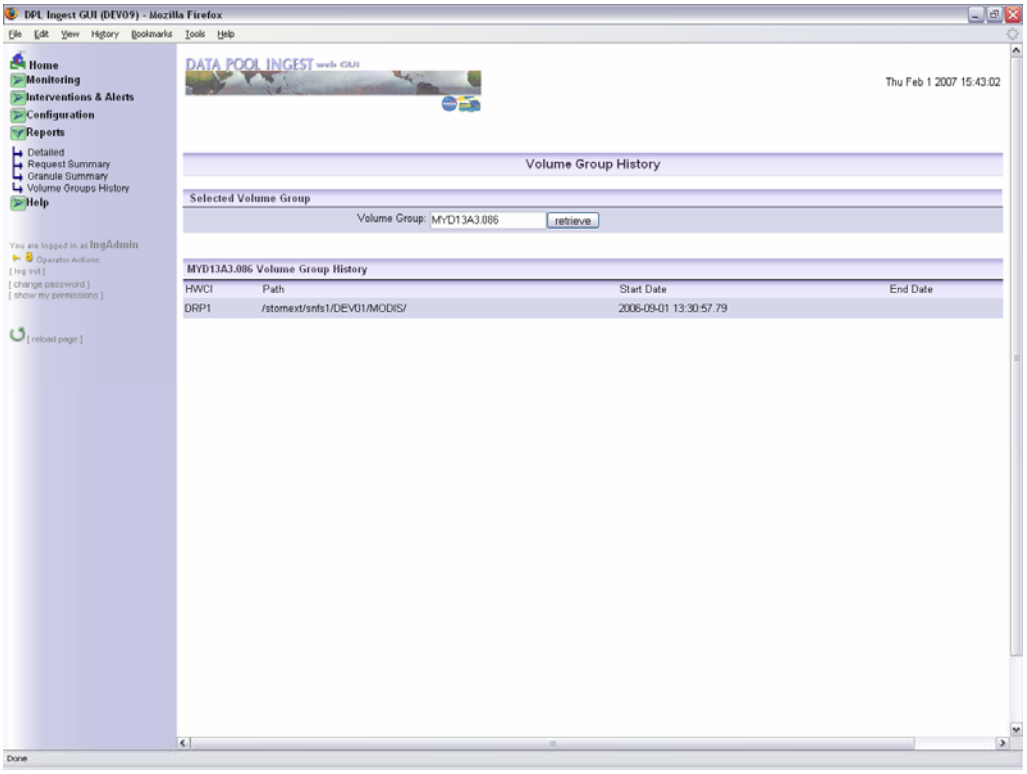


Figure 13.6-4. Group History Page

3 In the **Selected Volume Group** section, click in the field next to the **Volume Group** and enter the name of the **Volume Group** to be reviewed.

- The **Volume Group** name is displayed as entered.

4 Select the **retrieve** button.

- The **HWCI**, **Path**, **Start Date** and **End Date**, are displayed in **[Name] Volume Group History** section.

13.7 Help Pages and Context Help

The **Help** section, contains information the operator can have readily available while operating the Data Pool Ingest GUI. Included in this section are three pages: General Topics, Context Help and About.

The **General Topics** page includes an index of topics that should be useful to the operator in understanding how the GUI and Data Pool Ingest system work (see Figure13.7-1). The operator can press on the name of a section from the index in order to jump to the section text.

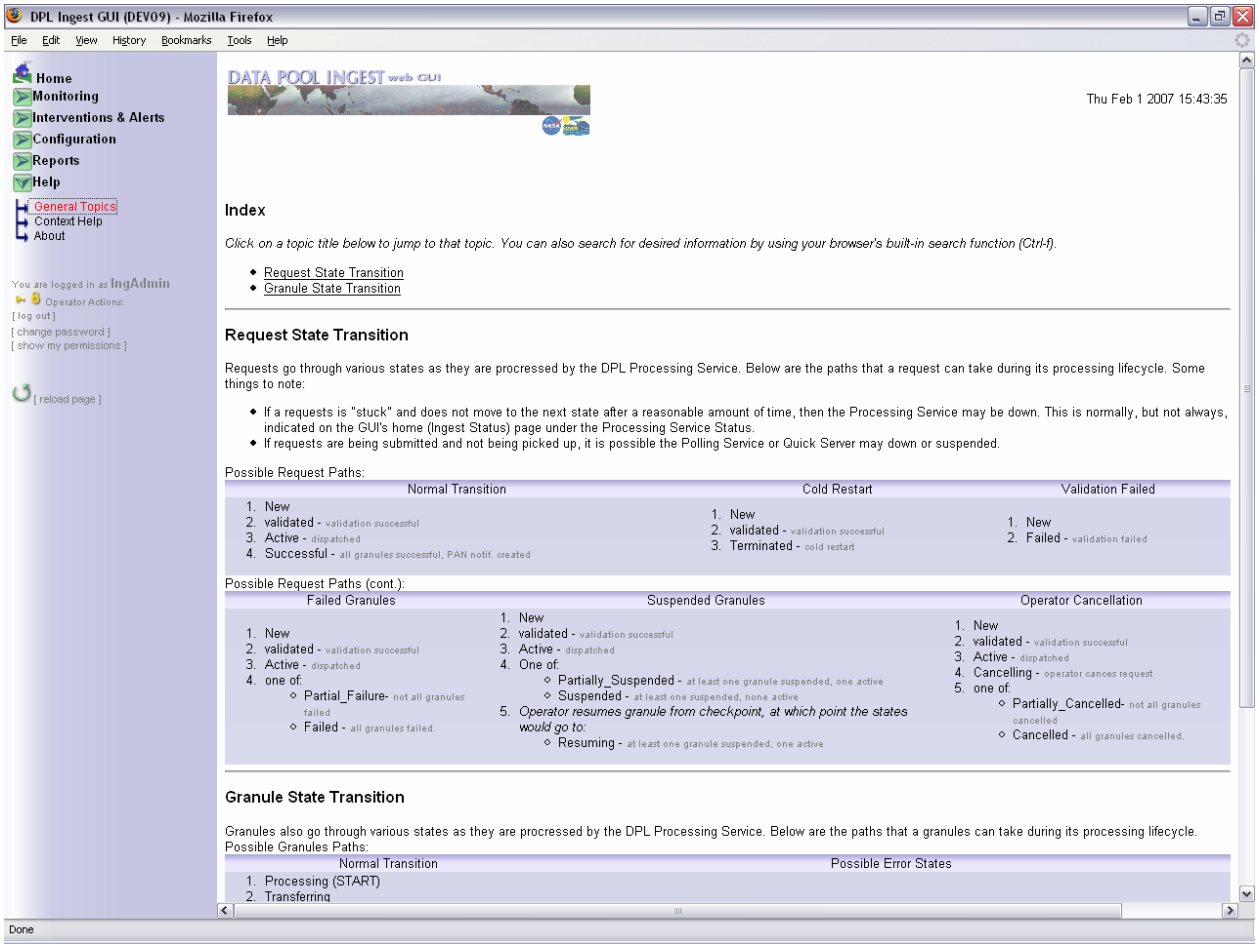


Figure 13.7-1. Help – General Topics

The **Context** page explains another tool provided by the operator to assist him in effectively using the Data Pool Ingest GUI. Throughout most pages on the DPL GUI, you can get relevant, context-sensitive help by hovering your mouse (no need to click) over the **[help]** text. In many

cases this is to explain the significance of a parameter or to provide instructions on what to do on the page. A blue pop-over window will appear and disappears as soon as the mouse is moved.

13.8 Data Pool Maintenance GUI

The Data Pool Maintenance GUI is responsible for monitoring and maintaining the Data Pool. Data Pool Ingest requires the use of the Data Pool Maintenance (DPM) GUI to perform the following tasks:

- Managing Data Pool Collection Groups
- Managing Data Pool Collections within Collection Groups

Table 13.8-1 provides an activity Checklist for Data Pool Maintenance.

Table 13.8-1. Data Pool Maintenance

Order	Role	Task	Section
1	Ingest Technician	Launching the DPM GUI	(P) 13.8.1.1
4	Ingest Technician	View Collection Groups	(P) 13.8.2.1
5	Ingest Technician	Modify Collection Groups	(P) 13.8.2.2
6	Ingest Technician	Add a Collection Group	(P) 13.8.2.3
7	Ingest Technician	Add an ECS Collection to a Collection Group	(P) 13.8.2.4
8	Ingest Technician	Modify an ECS Collection	(P) 13.8.2.5

13.8.1 Data Pool Maintenance GUI

The **DPM GUI** is used for Data Pool maintenance tasks. Of course, the first thing to do is launch the GUI. The procedure for launching the GUI is provided separately here and is referenced in other procedures. It applies to both full-capability and limited-capability operators.

13.8.1.1 Launch the DPM GUI

1 At the UNIX command line prompt enter**: setenv DISPLAY <client name>:0.0**

setenv DISPLAY <client name>:0.0

- Use either the X terminal/workstation IP address or the machine-name for the client name.
- When using secure shell, the DISPLAY variable is set just once, before logging in to remote hosts. If it were to be reset after logging in to a remote host, the security features would be compromised.

2 In the terminal window (at the command line prompt) start the log-in to the appropriate host by entering:

/tools/bin/ssh <host name>

- The **-l** option can be used with the ssh command to allow logging in to the remote host (or the local host for that matter) with a different user ID. For example, to log in to x0acs03 as user cmops enter:

  /tools/bin/ssh -l cmops x4oml01

- Depending on the set-up it may or may not be necessary to include the path (i.e., /tools/bin/) with the ssh command. Using ssh alone is often adequate. For example:

  ssh x4oml01

  - or -

  ssh -l cmops x4oml01

- Examples of Data Pool Maintenance GUI host names include **x4hel01.**
- If you receive the message, "Host key not found from the list of known hosts. Are you sure you want to continue connecting (yes/no)?" enter **yes** ("y" alone will not work).
- If you have previously set up a secure shell passphrase and executed sshremote, a prompt to Enter passphrase for RSA key '<user@localhost>' appears; continue with Step 3.
- If you have not previously set up a secure shell passphrase, go to Step 4.

3 If a prompt to **Enter passphrase for RSA key '<user@localhost>'** appears, enter:

<passphrase>.

- If a command line prompt is displayed, log-in is complete.
- If the passphrase is unknown, press **Return/Enter**, which should cause a **<user@remotehost>'s password:** prompt to appear (after the second or third try if not after the first one), then go to Step 4.
- If the passphrase is entered improperly, a **<user@remotehost>'s password:** prompt should appear (after the second or third try if not after the first one); go to Step 4.

4 If a prompt for **<user@remotehost>'s password:** appears, enter:

<password>

- A command line prompt is displayed.
- Log-in is complete.

5 Type **firefox &** then press **Return/Enter**.

- It may be necessary to respond to dialogue boxes, especially if the browser is already being used by someone else who has logged in with the same user ID.
- The Mozilla Firefox web browser is displayed.

6 If a bookmark has been created for the **DPM GUI**, select the appropriate bookmark from those listed on the browser's Bookmarks pull-down window.

- The **Login:** prompt is displayed.

7 If no bookmark has been created for the **DPM GUI**, type **http://*host*:*port*** in the browser's **Location (Go To)** field then press **Return/Enter**.

- For example: http:// f4hel01.hitc.com:22181/.

8 Type the appropriate user name in the **User** box of the security **Login** prompt.

9 Type the appropriate password in the **Password** box of the security **Login** prompt.

10 Click on the **Login** button:

- The dialogue box is dismissed.
- The **DPM GUI** ["Home" Page] is displayed (see Figure 13.8-1).

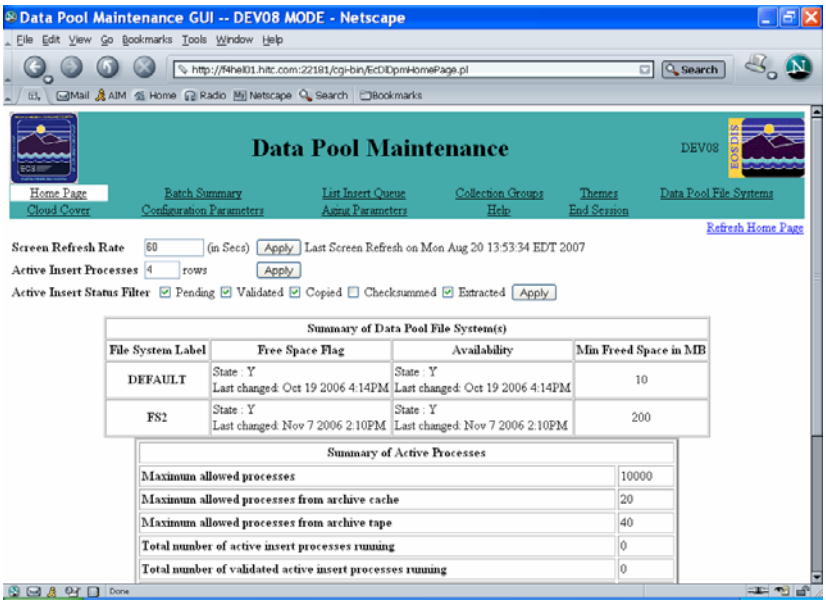


Figure 13.8-1. DPM GUI Home Page

13.8.2 Managing Data Pool Collection Groups

The conceptual structure of the data pool is set up for each DAAC based on the collections and granules archived at the DAAC. Related collections are grouped in **Collection Groups** (e.g., ASTER collections and granules from the Terra mission, MODIS Oceans collections and granules from the Terra Mission, MISR collections and granules from the Terra mission, MODIS Snow and Ice collections and granules from the Terra mission). Each collection group initially consists of a number of collections that have been specified as valid for Data Pool insertion (i.e., granules of the data types in the collection may be inserted into the Data Pool).

The Collection Group page of the **DPM GUI** allows both full-capability operators and limited-capability operators to view collection groups. It also provides access to pages for viewing collections within a collection group. In addition, the page has links that allow a full-capability operator to modify or add a collection group or collection in the Data Pool database.

Both full-capability operators and limited-capability operators can use the procedure that follows to display the list of collection groups that have collections specified as valid for Data Pool insertion and to view information about those collections.

13.8.2.1 View Collection Groups

1 Launch the **DPM GUI**.

- For detailed instructions refer to the **Launch the DPM GUI** procedure (previous procedure of this lesson).
- The **Home Page** is the default display, offering links for access to Data Pool maintenance function pages (i.e., **Data Pool File Systems**, **Cloud Cover**, **List Insert Queue**, **Batch Summary**, **Collection Groups**, **Themes, Configuration Parameters, Aging Parameters, Help** and **End Session**).

2 Click on the **Collection Groups** link.

- The Collection Group page is displayed (see Figure 13.8-2).

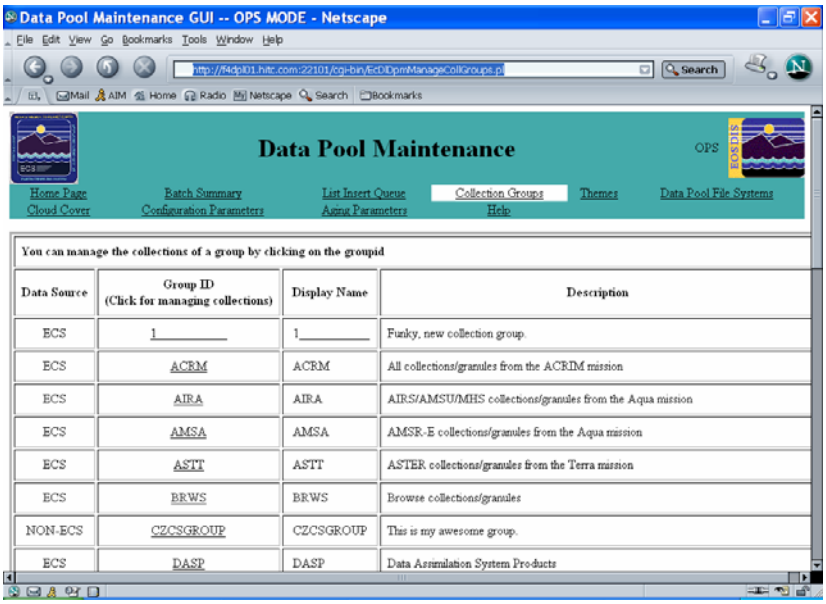


Figure 13.8-2. Collection Group Page

3 Observe data displayed on the Collection Group page.

- The table on the Collection Group page has columns containing the following types of collection group information:
  - **Data Source.**
  - **Group ID (Click for managing collections).**
  - **Display Name.**
  - **Description.**
- The following links are available on the Collection Groups page:
  - **Group ID (Click for managing collections)** - Links to a **List of Collections** contained in that group.
  - **Add Collection Group.**
  - **Modify Collection Group.**

4 To obtain more information about the collections in one of the groups, click on its link in the **Group ID (Click for managing collections)** column.

- The **List of Collection** page is displayed (see Figure 13.8-3).

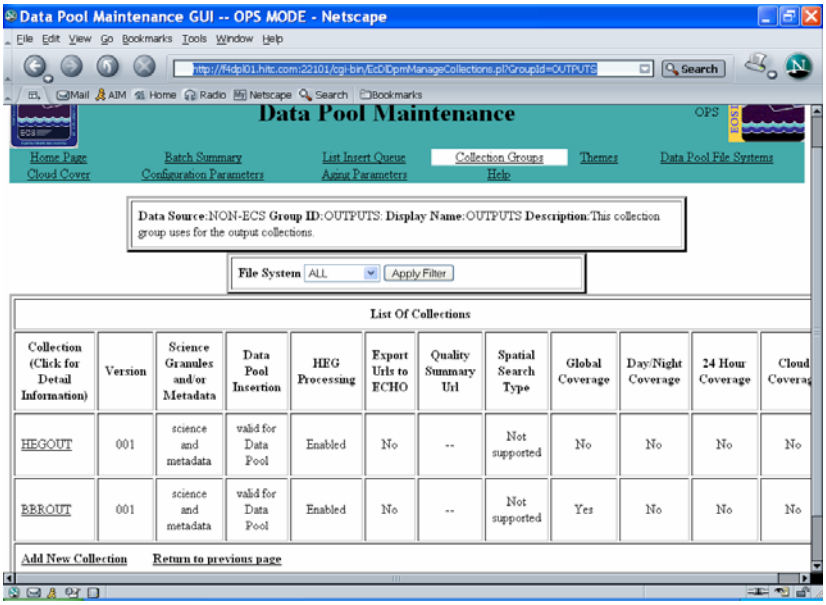


Figure 13.8-3. List of Collection

5 Observe data displayed on the **List of Collections** page.

- Near the top of the **List of Collections** page contains the following basic collection group information:
  - **Data Source.**
  - **Group ID.**
  - **Display Name.**
  - **Description.**
- There is a file system filter (and associated **Apply Filter** button) for displaying data on the **Collection Detail Information** page for all file systems or by individual file system.
- The **List of Collection** page has columns containing the following types of collection group information:
  - **Collection (Click for Detail Information)** link.
  - **Version.**
  - **Science Granules and/or Metadata.**
  - **Data Pool Insertion.**
  - **HEG Processing.**

- **Export Urls to ECHO.**
- **Quality Summary Url.**
- **Spatial Search Type.**
- **Global Coverage.**
- **Day/Night Coverage.**
- **24 Hour Coverage.**
- **Cloud Coverage.**

- The following links are available on the **List of Collection** page:
  - Each collection listed in the **Collection** column links to a **Collection Detail** page.
  - **Add New Collection.**
  - **Return to previous page.**

6 To filter data displayed on the **List of Collections** page, click on the **File System** filter option button.

- Options are displayed.

7 Select a file system filter option click on the appropriate choice from the option list.

8 To implement the filtering of data displayed on the **Collection Detail** link, click on the **Apply Filter** button.

- The **Collection (Click for Detail Information)** column is displayed with the filtered collection group information.

9 To obtain more information about one of the collections in the collection group, click on its link in the **Collection (Click for Detail Information**) column.

- The **Detail Information** page for the selected collection is displayed (see Figure 13.8-4).

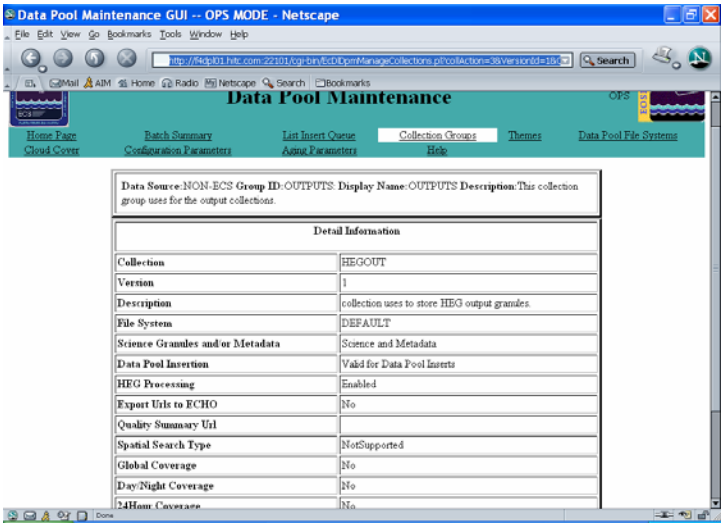


Figure 13.8-4. Detail Information

10 Observe data displayed on the **Detail Information** page.

- Near the top of the **Detail Information** page is the following basic collection group information:
  - **Data Source.**
  - **Group ID.**
  - **Display Name.**
  - **Description.**
- The **Detail Information** page has rows containing the following types of collection information:
  - **Collection**.
  - **Version.**
  - **Description.**
  - **File System.**
  - **Science Granules and/or Metadata.**
  - **Data Pool Insertion.**
  - **HEG Processing.**
  - **Export Urls to ECHO.**
  - **Quality Summary Url.**

- **Spatial Search Type.**
- **Global Coverage.**
- **Day/Night Coverage.**
- **24 Hour Coverage.**
- **Cloud Cover Type.**
- **Cloud Cover Source.**
- **Cloud Cover Description.**

- The following links are available on the Collection **Information Detail** page:
  - **Modify Collection.**
  - **Return to previous page.**

11 To view a description for another collection in the same group first click on the **Return to previous page** link.

- The **List of Collections** page is displayed again.

12 To view a description for another collection in the same group return to Step 9.

13 To view a description for another collection in another group return to Step 2.

It may be desirable to modify the description of one or more of the collection groups listed on the **Collection Groups** page. If there is a need to modify a collection group description, there is a link at the bottom of the list on that page providing access to a page that permits the descriptions to be modified. Full-capability operators (only) can use the following procedure to modify collection groups:

13.8.2.2 Modify Collection Groups

1 Launch the **DPM GUI**.

- For detailed instructions refer to the **Launch the DPM GUI** procedure (previous section of this lesson).
- The **Home Page** is the default display, offering links for access to Data Pool maintenance function pages (i.e., **Data Pool File Systems**, **Cloud Cover**, **List Insert Queue**, **Batch Summary**, **Collection Groups**, **Themes, Configuration Parameters, Aging Parameters, and End Session**).

2 Click on the **Collection Groups** link.

- The Collection Groups page is displayed, providing a table listing collection group information; i.e., **Data Source, Group ID (Click to Manage Collections)**, **Display Name**, and **Description**.

3 Click on the **Modify Collection Group** link at the bottom of the page.

- The **Modify Collection Group** page is displayed (see Figure 13.8-5), providing a table of collection group information showing five columns: **Data Source**, **Group ID (Click to Manage Collections)**, **Display Name**, **Description**, and **Check box to Modify** (containing a check box to mark the collection group for change).

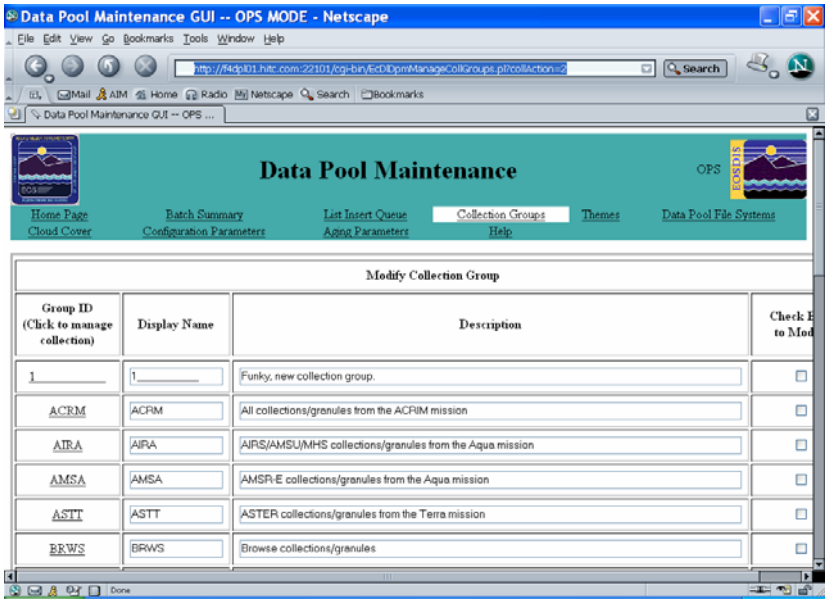


Figure 13.8-5. Modify Collection Group

- There is an **Apply Change** button at the bottom of the page for implementing changes.

4 To change the display name for the collection group, type the desired name in the **Display Name** field for the group ID.

- The **Display Name** may have no more than 12 characters.
  - Valid characters include A-Z, 0-9, underscore and space.

5 To change the description of the collection group, type the desired description in the **Description** field for the group ID.

- The **Description** may have no more than 255 characters.

6 Click in the check box at the end of the row containing collection group information to be modified.

- The selected collection group information is marked for modification.

7 Repeat Steps 4 through 6 for any additional collection groups to be modified.

8 Click on the **Apply Change** button.

- The revised collection group information is entered in the Data Pool database.

- The Collection Group page is displayed with the modified collection group information.

From time to time, it may be necessary to add a collection group (e.g., if a DAAC begins archiving data from a new instrument). If a collection group is to be added to the list of collection groups, it is necessary to use the **Add Collection Group** link at the bottom of the **Manage Collection Groups** page. Full-capability operators (only) can use the procedure that follows to modify collection groups:

NOTE: Although the following procedure is applicable, most of the time new collection groups will be added only during releases of new software versions and you will not use this procedure often.

Caution

The Add Collection Group function is to be exercised judiciously because the **DPM GUI** does not provide any means of deleting collection groups.

13.8.2.3 Add a Collection Group

---

1 Launch the **DPM GUI**.

- For detailed instructions refer to the **Launch the DPM GUI** procedure (previous procedure of this lesson).
- The **Home Page** is the default display, offering links for access to Data Pool maintenance function pages (i.e., **Data Pool File Systems**, **Cloud Cover**, **List Insert Queue**, **Batch Summary**, **Collection Groups**, **Themes, Configuration Parameters, Aging Parameters, and End Session**).

2 Click on the **Collection Groups** link.

- The Collection Groups page is displayed, providing a table listing collection group information; i.e., **Data Source, Group ID (Click to Manage Collections)**, **Display Name**, and **Description**.

3 Click on the **Add Collection Group** link at the bottom of the page.

- The **Add Collection Group** page is displayed (see Figure 13.8-6) providing a page with three columns of text-entry fields, **Data Source**, **Group ID**, **Display Name**, and **Description**.

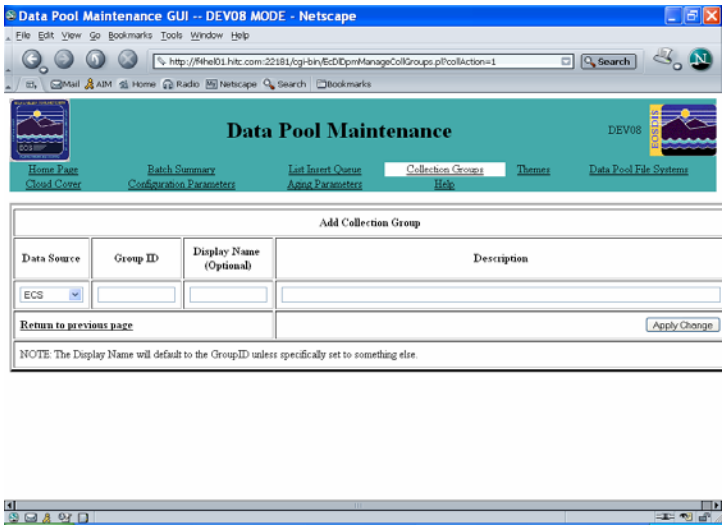


Figure 13.8-6. Add Collection Group

4 Select the **Data Source** from the pull-down window**.**

- **ECS** or **NONECS** is displayed in the field.

5 Type a unique identifier for the new collection group in the **Group ID** field.

- The **Group ID** may have no more than 12 characters.
  - Valid characters include A-Z, 0-9, and underscore.
- The **Group ID** will be compared with the existing **Group IDs** to ensure that it is not a duplicate of another ID.

6 To provide a display name that is different from the **Group ID** type a name in the **Display Name** field.

- The **Display Name** is the name for the collection as displayed on the **Data Pool Web Access GUI**.
- If no **Display Name** is entered, the **Group ID** will be used as the **Display Name**.
- The **Display Name** may have no more than 12 characters.
  - Valid characters include A-Z, 0-9, underscore and space.

7 Type the description for the new collection group in the **Description** field.

- The **Description** may have no more than 255 characters.

8 Click on the **Apply Change** button.

- The new collection group information is entered in the Data Pool database.
- The Collection Group page is displayed with the new collection group information.

Although an initial Data Pool structure is provided, not all collections are necessarily specified as eligible for Data Pool insertion. Based on experience, or on changes in demand, a DAAC may wish to add one or more collections to a data group. Full-capability operators (only) can use the following procedure to add an ECS collection to an existing collection group:

13.8.2.4 Add an ECS Collection to a Collection Group

1 Launch the **DPM GUI**.

- For detailed instructions refer to the **Launch the DPM GUI** procedure (previous section of this lesson).
- The **Home Page** is the default display, offering links for access to Data Pool maintenance function pages (i.e., **Data Pool File Systems**, **Cloud Cover**, **List Insert Queue**, **Batch Summary**, **Collection Groups**, **Themes, Configuration Parameters, Aging Parameters, and End Session**).

2 Click on the **Collection Groups** link.

- The Collection Group page is displayed, providing a table listing collection group information; i.e., **Data Source, Group ID (Click to Manage Collections)**, **Display Name**, and **Description**.

3 Click on the **Group ID** link for the ECS collection group to which the collection is to be added.

- The **List of Collections** page is displayed (see Figure 13.8-7)**.**

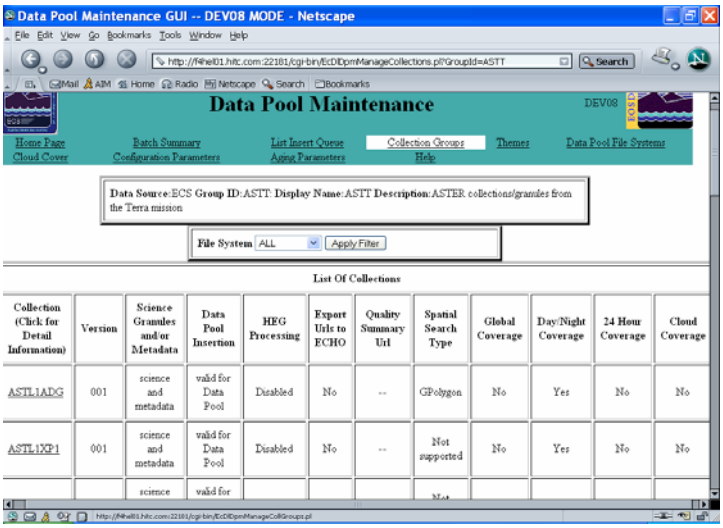


Figure 13.8-7. List of Collections

4 Click on the **Add New Collection** link at the bottom of the **List of Collections** page.

- The **Collections Not in Data Pool** page is displayed (see Figure 13.8-8).

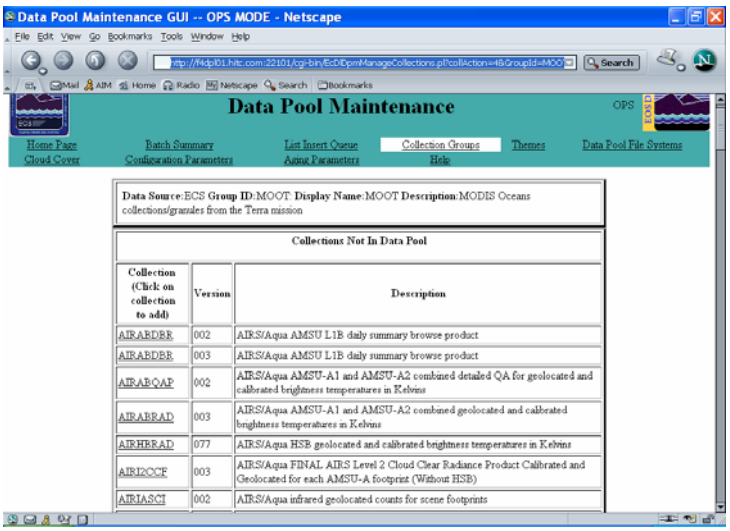


Figure 13.8-8. Collections Not In Data Pool Page

5 Click on the link in the **Collection (Click on collection to add)** column of the collection to be added to the collection group.

- The **Add New Collection** page is displayed (see Figure 13.8-9).

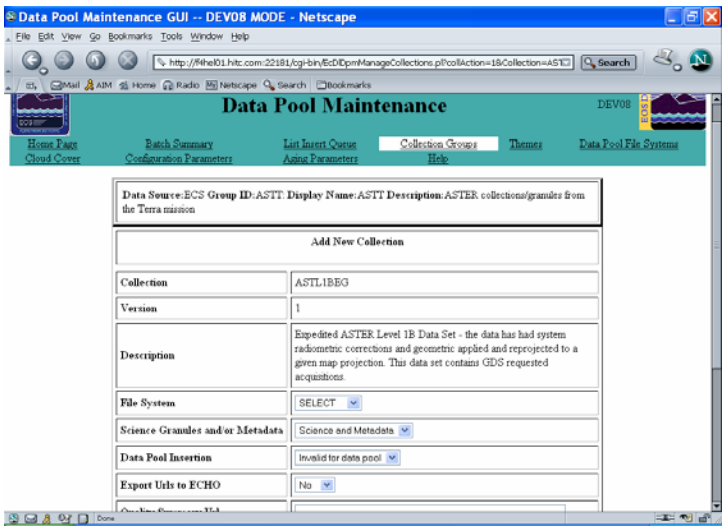


Figure 13.8-9. Add New Collection Page

NOTE: On the ECS collection version of the **Add New Collection** page the **Collection**, **Version**, **Description**, and **Spatial Search Type** fields are already filled in using information from the Data Pool database.

6 To select a file system option (if applicable), click on the appropriate choice from the **File System** option list.

7 To select a Science Granules and/or Metadata option, click on the appropriate choice from the Science Granules and/or Metadata option list.

- **Science and Metadata** is the default option.

8 To select a data pool insertion option, click on the appropriate choice from the Data Pool Insertion option list.

- **Invalid for data pool** is the default option.
- **Valid for data pool** must be selected if the collection is to be eligible for insertion into the Data Pool.

9 To select an ECHO export option, click on the appropriate choice from the **Export Urls to ECHO** option list.

- **No** is the default option.
- **Yes** must be selected if collection URLs are to be eligible for export to ECHO.

10 If the collection is to be linked to a quality summary web site, enter the URL in the **Quality Summary Url** text entry field.

- Ensure that **http://** is included in the **Quality Summary Url** text entry field.

11 To select a global coverage option, click on the appropriate choice from the **Global Coverage** option list.

- **Yes** indicates no spatial searches for the collection.
- **No** indicates that spatial searches are allowed for the collection.

12 To select a day/night coverage option, click on the appropriate choice from the **Day/Night Coverage** option list.

- **Yes** indicates that day/night searches are allowed for the collection.
- **No** indicates that the collection is excluded from day/night searches.

13 To select a 24-hour coverage option, click on the appropriate choice from the **24 Hour Coverage** option list.

- **Yes** indicates that the collection is excluded from time of day searches.
- **No** indicates that time of day searches are allowed for the collection.

14 To select a cloud cover type and source option, click on the appropriate choice from the **Cloud Cover Type & Source** option list.

- All cloud cover information in the Data Pool database is listed.
- If the desired cloud cover type/source is not listed, it can be entered using the procedure **Add New Cloud Cover Information Using the DPM GUI** (previous section of this lesson).

15 To view details of cloud cover type and source, click on the **View Details** link adjacent to the **Cloud Cover Type & Source** option list.

16 Click on the **Apply Change** button.

- The new collection information is entered in the Data Pool database.
- The **List of Collection** page is displayed with the new collection information.

As part of managing the Data Pool storage and retention of data, making adjustments based on experience and/or changes in demand, it may be desirable to modify a collection. The modification may mean specifying that metadata only may continue to be inserted and science granules may no longer be inserted, or declaring the collection no longer valid for data pool insertion at all.

13.8.2.5 Modify an ECS Collection

---

1 Launch the **DPM GUI**.

- For detailed instructions refer to the **Launch the DPM GUI** procedure (previous section of this lesson).
- The **Home Page** is the default display, offering links for access to Data Pool maintenance function pages (i.e., **Data Pool File Systems**, **Cloud Cover**, **List Insert Queue**, **Batch Summary**, **Collection Groups**, **Themes, Configuration Parameters, Aging Parameters, and End Session**).

2 Click on the **Collection Groups** link.

- The Collection Group page is displayed, providing a table listing collection group information; i.e., **Data Source, Group ID (Click to Manage Collections)**, **Display Name**, and **Description**.

3 Click on the **Group ID** link for the ECS collection group to which the collection is to be added.

- The **List of Collections** page is displayed**.**

4 Click on the desired link found in the **Collection (Click on collection to add)** column.

- The **Detail Information** page is displayed.

5 Click on the **Modify Collection** link.

- The **Modify Collection** page is displayed (see Figure 13.8-10).

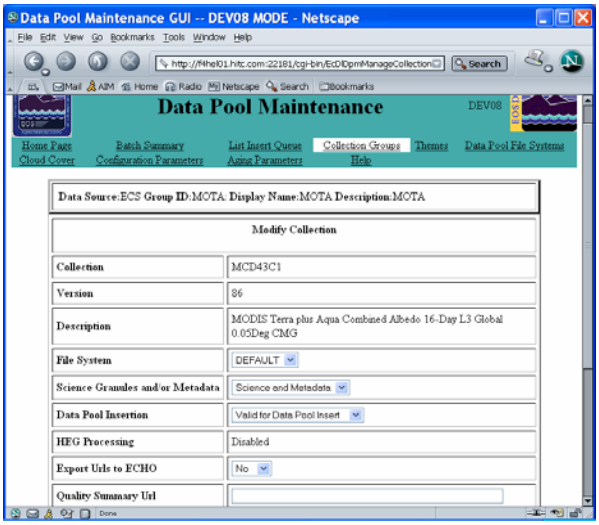


Figure 13.8-10. Modify Collection Page

NOTE: On the ECS collection version of the **Modify Collection** page, the **Collection**, **Version**, **Description**, **Spatial Search Type, and HEG Processing** fields cannot be edited.

6 To select a file system option (if applicable), click on the appropriate choice from the **File System** option list.

7 To select a Science Granules and/or Metadata option, click on the appropriate choice from the Science Granules and/or Metadata option list.

- **Science and Metadata** is the default option.

8 To select a data pool insertion option, click on the appropriate choice from the Data Pool Insertion option list.

- **Invalid for data pool** is the default option.
- **Valid for data pool** must be selected if the collection is to be eligible for insertion into the Data Pool.

9 To select an ECHO export option, click on the appropriate choice from the **Export Urls to ECHO** option list.

- **No** is the default option.
- **Yes** must be selected if collection URLs are to be eligible for export to ECHO.

10 If the collection is to be linked to a quality summary web site, enter the URL in the **Quality Summary Url** text entry field.

- Ensure that **http://** is included in the **Quality Summary Url** text entry field.

11 To select a global coverage option, click on the appropriate choice from the **Global Coverage** option list.

- **Yes** indicates no spatial searches for the collection.
- **No** indicates that spatial searches are allowed for the collection.

12 To select a day/night coverage option, click on the appropriate choice from the **Day/Night Coverage** option list.

- **Yes** indicates that day/night searches are allowed for the collection.
- **No** indicates that the collection is excluded from day/night searches.

13 To select a 24-hour coverage option, click on the appropriate choice from the **24 Hour Coverage** option list.

- **Yes** indicates that the collection is excluded from time of day searches.
- **No** indicates that time of day searches are allowed for the collection.

14 To select a cloud cover type and source option, click on the appropriate choice from the **Cloud Cover Type & Source** option list.

- All cloud cover information in the Data Pool database is listed.
- If the desired cloud cover type/source is not listed, it can be entered using the procedure **Add New Cloud Cover Information Using the DPM GUI** (previous section of this lesson).

15 To view details of cloud cover type and source, click on the **View Details** link adjacent to the **Cloud Cover Type & Source** option list.

16 Click on the **Apply Change** button.

- The new collection information is entered in the Data Pool database.
- The **List of Collection** page is displayed with the new collection information.
