625-EMD-010

EOSDIS Maintenance and Development Project

Training Material for the EMD Project
Volume 10: Archive Processing

Revision 02

July 2006

Raytheon Company
Upper Marlboro, Maryland

Training Material for the EMD Project
Volume 10: Archive Processing

Revision 02

July 2006

Prepared Under Contract NAS5-03098
CDRL Item 23

RESPONSIBLE AUTHOR

Debra Stepp July 28 2006

Lay'wan Gamble and Debra Stepp Date
EOSDIS Maintenance and Development Project

RESPONSIBLE OFFICE

7/28/06

Mary Armstrong, Program Manager Date
EOSDIS Maintenance and Development Project

Raytheon Company
Upper Marlboro, Maryland

This page intentionally left blank.

Preface

This document is a formal contract deliverable. It requires Government review and approval within 45 business days. Changes to this document will be made by document change notice (DCN) or by complete revision.

Any questions should be addressed to:

Data Management Office
The EMD Project Office
Raytheon Company
1616 McCormick Drive
Upper Marlboro, Maryland 20774-5301

Revision History

Document Number	Status/Issue	Publication Date	CCR Number
625-EMD-010	Original	April 2004	04-384
625-EMD-010	Revision 01	July 2005	05-320
625-EMD-010	Revision 02	July 2006	06-377

This page intentionally left blank.

Abstract

This is Volume 10 of a series of lessons containing the training material for the Earth Observing System Data and Information System (EOSDIS) Maintenance and Development (EMD) Project. This lesson provides a detailed description of the process required to perform the tasks associated with archive functions.

***Keywords**:* training, instructional design, course objective, archive, AMASS, ACSLS, AAWin, granule deletion tool, Spatial Subscription Server, Data Pool, Data Pool maintenance, Release 7.

This page intentionally left blank.

Contents

Preface

Abstract

Introduction

Identification 1

Scope 1

Purpose 1

Status and Schedule 1

Organization 1

Related Documentation

Parent Documents 3

Applicable Documents 3

Information Documents 3

Information Documents Referenced 3

Information Documents Not Referenced 4

Archive Processing Overview

Lesson Overview 5

Lesson Objectives 5

Importance 9

Archive Processing

Archive Manager 11
Hardware 12
Software 14
ACSLS 16
Data Sources and Uses 17

Starting and Shutting Down AMASS

Starting the AMASS Tape Archive System 19
Shutting Down AMASS Tape Archive System 22
Rebooting AMASS 22
Entering the Archive After AMASS Is Started 23

Logging in to System Hosts

Logging in to System Hosts 25

Archive Storage Structures

Storage Element Relationships 29
Launching DSS GUIs 29
Archive Resource Management 35

Insert Data Into the Archive

Archive Insert Scenario 39

Monitor Archive Requests

Request Status Window 43
Distrib’n Requests Window 46

Retrieve Data from the Archive

Data Pool Insert 49

Monitoring Distribution Requests 51

Deleting Granules

Phases of Granule Deletion 53

Phase 1, Logical Deletion 54

Generating a GeoID File to Be Used as Input to the Bulk Delete Utility or the Bulk Undelete Utility 54

Deleting Granules, Phase 1: Running the Bulk Delete Utility 59

"Undeleting" Granules from the Archive and Inventory 65

Phase 2, Physical Deletion 69

Deleting Granules, Phase 2: Running the Deletion Cleanup Utility 69

Phase 3, Deletion from the Archive 78

Deleting Granules, Phase 3: Running the Archive Batch Delete Utility 78

Loading Archive Media

Automatically Loading Archive Media 85

Manually Loading Archive Media 86

Formatting a Volume 87

Remove Media 88

Backup Archived Data

Creating Offsite Backups 89

Creating a Backup for AMASS 93

Replacing the AMASS Database Backup Volume (Volume 1) 94

Create Replacement Backups Manually from Existing Archives 96

Restore Archive Data

Use of Backup Data for Recovery 97
Manual Data Recovery from Local or Offsite Backup Tapes 97
Reprocessing 98
Requesting Replacement Data from Provider 98
Restoring the AMASS Database 111
Restoring the ACSLS Database 112

AMASS Graphical User Interface

Using the AMASS GUI 115
Modify a Volume Group 116
Modify a Volume 119

Archive Monitoring and Troubleshooting

Trouble Symptoms 123
Fault Recovery 123
Fault Handling 124
Client Crash and Restart 126
Server Crash and Restart 128
Request Resubmission 130
Troubleshooting Archive Problems 131
AMASS Commands, Utilities, and Scripts for Monitoring and Fault Response 134
Recovery from Failure to Store or Retrieve Data 142
Checking Server Log Files 142
A Special Case: Checking the Request Manager Server Debug Log 150
Checking the *tac* Log 162
Handling a Data Insertion Failure 163
Handling a Data Acquire Failure 167
Checksum Verification 172
Checksum De-activation 174

Data Pool Management

Features of the Data Pool Maintenance GUI 177

Home Page 182

Data Pool File Systems 182

Compression Algorithms 185

Cloud Cover 189

Batch Summary 192

List Insert Queue 193

Configuration Parameters 195

Aging Parameters 196

Collection Groups 197

Themes 206

Help 208

Procedures for Using the Data Pool Maintenance GUI 210

Launching and Shutting Down the DPM GUI 210

Monitoring Data Pool Active Insert Processes and Insert Actions 213

Managing Data Pool File Systems Using the DPM GUI 215

Enabling/Disabling Data Compression Using the DPM GUI 219

Managing Compression Algorithms Using the DPM GUI 221

Managing Cloud Cover Information Using the DPM GUI 226

Checking the Status of Batch Inserts 230

Checking the Data Pool Insert Queue and Canceling a Data Pool Insert Action 231

Managing Data Pool Configuration Parameters and Data Pool Tuning 233

Managing Data Pool Collection Groups 241

Managing Data Pool Collections within Collection Groups 247

Managing Themes Using the DPM GUI 263

Working with Data Pool Scripts 270

Extending the Period of Retention for Granules in the Data Pool 271

Running the Data Pool Cleanup Utility 275

Running the Data Pool Access Statistics Utility 284

Using the Batch Insert Utility 293

Running the Bulk Metadata Generation Tool (BMGT) 295

Running the Bulk URL Utility 301

Running the Most Recent Data Pool Inserts Utility 304

Running the Data Pool Collection-to-Group Remapping Utility 306
Running the Data Pool QA Update Utility 307
Running the Data Pool Move Collections Utility 308
Running the Data Pool Hidden Scrambler Utility in Rename Mode 311

Tuning Data Pool and Order Manager Subsystem Configuration Parameters

Staging Policies 318
Preventing a Set of Distribution Requests from Monopolizing Archive Tape Drives 319
Adjusting the Pace of Staging 322
Throttling Archive Staging for Output Devices and FTP Connections 326
Ensuring the Staging of Low-Priority Requests at a Reasonable Pace 326
Ensuring That High-Priority Requests Are Expedited 328
Reserving Enough Tape Drives for Non-OMS/Non-DPL-Related Activities 330
Managing Data Pool In-Cache and From-Tape Insert Processes Efficiently 331
Optimizing Tape Archive Resources 332
Summary 332

Troubleshooting Data Pool Problems

Trouble Symptoms 335
Checking Data Pool Log Files 348
Checking GeoID Files 350
Using ISQL to Check Data Pool Database Tables 353
Notes Concerning the Data Pool Database 355

Open Geospatial Consortium (OGC) Web Services (OWS) Support

Deployment of Open Geospatial Consortium (OGC) Web Services (DOWS) 357
Working with DOWS Scripts 360
Manually Invoking the DOWS Synchronization Utility 360
Manually Invoking the DOWS GeoTIFF Converter Utility 363
Using the DOWS OWS Binding Utility (OBU) 365
Standalone OGC (WCS/WMS) Archive 367
Using the Open Geospatial Consortium (OGC) Web Services (OWS) Registration Utility 367

Using the Spatial Subscription Server (NBSRV) GUI

Spatial Subscription Server (NBSRV) GUI 371

Procedures for Using the Spatial Subscription Server GUI 392

Launching and Shutting Down the Spatial Subscription Server GUI 392

Listing Subscribable Events and Subscriptions 395

Adding a Subscription 398

Working with Existing Subscriptions 403

Managing Bundling Orders Using the Spatial Subscription Server GUI 414

Monitoring the Spatial Subscription Server Using the Spatial Subscription Server GUI .427

Practical Exercise

Perform Activities Related to Archive Processing 433

Slide Presentation

Slide Presentation Description 437

This page intentionally left blank.

Introduction

Identification

Training Material Volume 10 is part of Contract Data Requirements List (CDRL) Item 23, which is a required deliverable under the Earth Observing System Data and Information System (EOSDIS) Maintenance and Development (EMD) Contract (NAS5-03098).

Scope

Training Material Volume 10 describes the process and procedures associated with Archive Processing. It describes archive hardware, software, and data. In addition, it addresses starting and shutting down the tape archive control software, monitoring archive requests, and performing archive management tasks. This lesson is designed to provide the operations staff with sufficient knowledge and information to satisfy all lesson objectives.

Purpose

The purpose of this Student Guide is to provide a detailed course of instruction that forms the basis for understanding data archiving. Lesson objectives are developed and will be used to guide the flow of instruction for this lesson. The lesson objectives will serve as the basis for verifying that all lesson topics are contained within this Student Guide and slide presentation material.

Status and Schedule

This lesson module provides detailed information about training for the current baseline of the system. Revisions are submitted as needed.

Organization

This document is organized as follows:

Introduction:	The Introduction presents the document identification, scope, purpose, and organization.
Related Documentation:	Related Documentation identifies parent, applicable and information documents associated with this document.
Student Guide:	The Student Guide identifies the core elements of this lesson. All Lesson Objectives and associated topics are included.
Slide Presentation:	Slide Presentation is reserved for all slides used by the instructor during the presentation of this lesson.

This page intentionally left blank.

Related Documentation

Parent Documents

The parent documents are the documents from which the EMD Training Material's scope and content are derived.

423-41-01	Goddard Space Flight Center, EOSDIS Core System (ECS) Statement of Work
423-46-03	EMD Task 101 Statement of Work For ECS SDPS Maintenance
423-46-02	Contract Data Requirements Document for EMD Task 101 ECS SDPS Maintenance

Applicable Documents

The following documents are referenced within this EMD Training Material, or are directly applicable, or contain policies or other directive matters that are binding upon the content of this document:

420-05-03	Goddard Space Flight Center, Earth Observing System (EOS) Performance Assurance Requirements for the EOSDIS Core System (ECS)
423-41-02	Goddard Space Flight Center, Functional and Performance Requirements Specification for the Earth Observing System Data and Information System (EOSDIS) Core System (ECS) (ECS F&PRS)
423-46-01	Goddard Space Flight Center, Functional and Performance Requirements Specification for the Earth Observing System Data and Information System (EOSDIS) Core System (ECS) Science Data Processing System (EMD F&PRS)

Information Documents

Information Documents Referenced

The following documents are referenced herein and amplify or clarify the information presented in this document. These documents are not binding on the content of the EMD Training Material.

609-EMD-001	Release 7.11 Operations Tools Manual for the EMD Project
611-EMD-001	Release 7.11 Mission Operation Procedures for the EMD Project

Information Documents Not Referenced

The following documents, although not referenced herein and/or not directly applicable, do amplify or clarify the information presented in this document. These documents are not binding on the content of the EMD Training Material.

305-EMD-001	Release 7.11 Segment/Design Specification for the EMD Project
313-EMD-001	Release 7.11 Internal Interface Control Document for the EMD Project
910-TDA-022	Custom Code Configuration Parameters for ECS
508-EMD-001	ACRONYMS for the EOSDIS Maintenance and Development (EMD) Project
152-TP-003	Glossary of Terms for the EOSDIS Core System (ECS) Project

Archive Processing Overview

Lesson Overview

This lesson reviews the process of archiving data, including a description of processing for working with the File Storage Management System (FSMS) software, monitoring the ingest/archiving/distribution performance, managing archive content and capacity, maintaining configuration of peripherals and data servers, backing up and restoring archived data, documenting and troubleshooting archive errors, maintaining the archive processing queue (storage and retrieval), and providing archive status.

Lesson Objectives

Overall Objective - The overall objective of this lesson is proficiency in the methodology and procedures for archive processing in the Earth Observing System Data and Information System (EOSDIS) Core System (ECS) during maintenance and operations. The lesson includes a description of processing for monitoring the ingest/archival/distribution performance, maintaining configuration of peripherals and data servers, documenting archive errors, maintaining the archive processing queue (both storing and retrieval), managing archive content and capacity, submitting new data archive requests to the Science Coordinator, and providing archive status.

Specific Objective 1 - The student will list DAAC operator positions for Archive Manager personnel interfaces and identify responsibilities associated with each interface.

Condition - The student will be given a list of DAAC operators.

Standard - The student will select four personnel positions with which the Archive Manager interfaces and list at least one major area of responsibility for each selected position.

Specific Objective 2 - The student will identify the major hardware facility for archival storage and its associated storage cartridges.

Condition - The student will be given a copy of 611-EMD-001, *Release* 7.11 *Mission Operation Procedures for the EMD Project*, a copy of 609-EMD-001, *Release* 7.11 *Operations Tools Manual for the EMD Project*, a working system archive facility, a 9940 tape cartridge, and a 9840 tape cartridge.

Standard - The student will correctly point out the StorageTek Library Storage Module (LSM) and its associated 9940 and 9840 tape cartridges.

Specific Objective 3 - The student will describe the File Storage Management System (FSMS) software.

Condition - The student will be given a copy of 611-EMD-001, *Release* 7.11 *Mission Operation Procedures for the EMD Project* and a copy of 609-EMD-001, *Release* 7.11 *Operations Tools Manual for the EMD Project*.

Standard - The student will identify the FSMS software as the Archival Management and Storage System (AMASS), correctly describe AMASS by stating its nature as a UNIX file system installed on an SGI XL computer, and state the five steps in the AMASS control path without error.

Specific Objective 4 - The student will start the AMASS tape archive system.

Condition - The student will be given a copy of 611-EMD-001, *Release* 7.11 *Mission Operation Procedures for the EMD Project*, a copy of 609-EMD-001, *Release* 7.11 *Operations Tools Manual for the EMD Project*, and a working system archive facility.

Standard - The student will operate the STK control panels, power up the archive hardware, and then boot AMASS host and start AMASS without error and in accordance with documented procedures.

Specific Objective 5 - The student will shut down the AMASS tape archive system.

Condition - The student will be given a copy of 611-EMD-001, *Release* 7.11 *Mission Operation Procedures for the EMD Project*, a copy of 609-EMD-001, *Release* 7.11 *Operations Tools Manual for the EMD Project*, and a working system archive facility with AMASS started.

Standard - The student will terminate AMASS and shut down the LSM without error and in accordance with documented procedures.

Specific Objective 6 - The student will use manual mode to enter the LSM.

Condition - The student will be given a copy of 611-EMD-001, *Release* 7.11 *Mission Operation Procedures for the EMD Project*, a copy of 609-EMD-001, *Release* 7.11 *Operations Tools Manual for the EMD Project*, and a working system archive facility with AMASS started.

Standard - The student will vary the LSM offline, enter the LSM, leave the LSM, and vary the LSM back online, without error, in accordance with documented procedures, and following all required safety precautions.

Specific Objective 7 - The student will describe the relationships between Earth Science Data Types (ESDTs), Logical Volume Groups (LVGs) in the Archive, and physical archive volume groups.

Condition - The student will be given a diagram depicting the relationships.

Standard - The student will correctly explain the logical and physical structure of archive storage for the system.

Specific Objective 8 - The student will describe the process of, and monitor the progress of, inserting new data into the archive.

Condition - The student will be given a copy of 611-EMD-001, *Release* 7.11 *Mission Operation Procedures for the EMD Project*, a copy of 609-EMD-001, *Release* 7.11 *Operations Tools Manual for the EMD Project*, and a working system archive facility.

Standard - The student will describe without error the process of inserting new data into the archive.

Specific Objective 9 - The student will launch available Data Server Subsystem (DSS) Graphical User Interfaces (GUIs) and monitor retrieval of data from the archive.

Condition - The student will be given a copy of 611-EMD-001, *Release* 7.11 *Mission Operation Procedures for the EMD Project*, a copy of 609-EMD-001, *Release* 7.11 *Operations Tools Manual for the EMD Project*, and a working system archive facility.

Standard - The student will perform without error and in accordance with documented procedures the procedures relating to monitoring retrieval of data from the archive.

Specific Objective 10 - The student will use the granule deletion capability to delete granules from the archive and inventory.

Condition - The student will be given a copy of 611-EMD-001, *Release* 7.11 *Mission Operation Procedures for the EMD Project*, a copy of 609-EMD-001, *Release* 7.11 *Operations Tools Manual for the EMD Project*, and a working system archive facility.

Standard - The student will perform without error and in accordance with documented procedures the procedure for deleting granules from the archive and inventory.

Specific Objective 11 - The student will perform automatic and manual loading of archive storage cartridges.

Condition - The student will be given a copy of 611-EMD-001, *Release* 7.11 *Mission Operation Procedures for the EMD Project*, a copy of 609-EMD-001, *Release* 7.11 *Operations Tools Manual for the EMD Project*, a working system archive facility, 9940 tape cartridges, and 9840 tape cartridges.

Standard - The student will perform procedures for automatic and manual loading of the proper cartridges for the LSM without error and in accordance with documented procedures.

Specific Objective 12 - The student will create a backup for AMASS.

Condition - The student will be given a copy of 611-EMD-001, *Release* 7.11 *Mission Operation Procedures for the EMD Project*, a copy of 609-EMD-001, *Release* 7.11 *Operations Tools Manual for the EMD Project*, and a working system archive facility.

Standard - The student will correctly and in accordance with documented procedures use the *vgexport -q* command to create the AMASS backup.

Specific Objective 13 - The student will replace a full Backup Volume.

Condition - The student will be given a copy of 611-EMD-001, *Release* 7.11 *Mission Operation Procedures for the EMD Project*, a copy of 609-EMD-001, *Release* 7.11 *Operations Tools Manual for the EMD Project*, and a working system archive facility.

Standard - The student will perform without error and in accordance with documented procedures the procedures for replacing the Backup Volume (Volume 1).

Specific Objective 14 - The student will manually create a replacement backup for an archive data tape.

Condition - The student will be given a copy of 611-EMD-001, *Release* 7.11 *Mission Operation Procedures for the EMD Project*, a copy of 609-EMD-001, *Release* 7.11 *Operations Tools Manual for the EMD Project*, and a working system archive facility.

Standard - The student will perform without error and in accordance with documented procedures the procedures for manually creating a replacement backup for an archive data tape.

Specific Objective 15 - The student will "restore" archive data by inserting a backup copy cartridge.

Condition - The student will be given a copy of 611-EMD-001, *Release* 7.11 *Mission Operation Procedures for the EMD Project*, a copy of 609-EMD-001, *Release* 7.11 *Operations Tools Manual for the EMD Project*, and a working system archive facility.

Standard - The student will perform without error the procedure for inserting a backup copy cartridge to replace a lost archive data tape.

Specific Objective 16 - The student will conduct basic checks on AMASS status and functioning appropriate for initial troubleshooting.

Condition - The student will be given a copy of 611-EMD-001, *Release* 7.11 *Mission Operation Procedures for the EMD Project*, a copy of 609-EMD-001, *Release 7.11 Operations Tools Manual for the EMD Project*, and a working system archive facility.

Standard - The student will successfully check for active processes reflecting required AMASS daemons, run commands to check the basic health of the running AMASS applications, and run an AMASS command to display current AMASS input/output activity without error and in accordance with documented procedures.

Specific Objective 17 - The student will use the *quedisplay* command to display what is in the AMASS queue.

Condition - The student will be given a copy of 611-EMD-001, *Release 7.11 Mission Operation Procedures for the EMD Project*, a copy of 609-EMD-001, *Release 7.11 Operations Tools Manual for the EMD Project*, and a working system archive facility.

Standard - The student will perform without error and in accordance with documented procedures the procedure for viewing what is in the AMASS queue.

Specific Objective 18 - The student will use the *amass_log* script to display AMASS errors.

Condition - The student will be given a copy of 611-EMD-001, *Release 7.11 Mission Operation Procedures for the EMD Project*, a copy of 609-EMD-001, *Release 7.11 Operations Tools Manual for the EMD Project*, and a working system archive facility.

Standard - The student will perform without error and in accordance with documented procedures the procedure for using the *amass_log* script to display AMASS messages from the system log file.

Specific Objective 19 - The student will launch and use the AMASS Graphical User Interface (GUI).

Condition - The student will be given a copy of 611-EMD-001, *Release 7.11 Mission Operation Procedures for the EMD Project*, a copy of 609-EMD-001, *Release 7.11 Operations Tools Manual for the EMD Project*, and a working system archive facility.

Standard - The student will successfully start the AMASS GUI, display icons and data for volume groups and volumes, and execute procedures for modifying volume groups and volumes without error and in accordance with documented procedures.

Specific Objective 20 - The student will perform the functions required to maintain the Data Pool, including tasks with the Data Pool Maintenance (DPM) GUI (monitor Data Pool active insert processes; monitor/cancel data pool insert actions; suspend and resume Data Pool actions; check the Data Pool insert queue; toggle the state of the NoFreeSpace flag; configure the number of allowed active insert processes; configure the default retention period and the default retention priority; view and update collection groups in the Data Pool database; list/add/delete a theme), tasks with utilities and scripts (extend the period of retention for selected science granules already in the Data Pool; set up a schedule and cron job for Data Pool cleanup; manually invoke Data Pool cleanup; set up a schedule and cron job for Data Pool access statistics accumulation; manually invoke the Data Pool access statistics utility), and tasks with the Spatial Subscription Server (NBSRV) GUI (extend the period of retention in a Data Pool insert subscription; update a subscription; view/add/cancel a bundling order; view statistics on processing of events and actions by the Spatial Subscription Server).

Condition - The student will be given a copy of 611-EMD-001, *Release 7.11 Mission Operation Procedures for the EMD Project*, a copy of 609-EMD-001, *Release 7.11 Operations Tools Manual for the EMD Project*, and a working Data Pool facility.

Standard - The student will use the GUI tools, scripts, and utilities without error in accordance with applicable procedures to perform the required Data Pool maintenance functions.

Importance

The Archive Manager's role in maintaining the archive data is key to the successful implementation and operation of the system. Ensuring the smooth operation of the archive is crucial for system core functionality.

This page intentionally left blank.

Archive Processing

Archive Manager

Archive processing is at the heart of the system. Through archive processing, data that have been ingested into the system are archived to tape for permanent storage and distributed to users via hard media (tape or disk) or electronic means.

The DAAC Archive Manager's job entails working with the Science Data Specialist, the Science Coordinator, and the Resource Manager, as well as providing direction for the Data Ingest Technician. These personnel interfaces are illustrated in Figure 1.

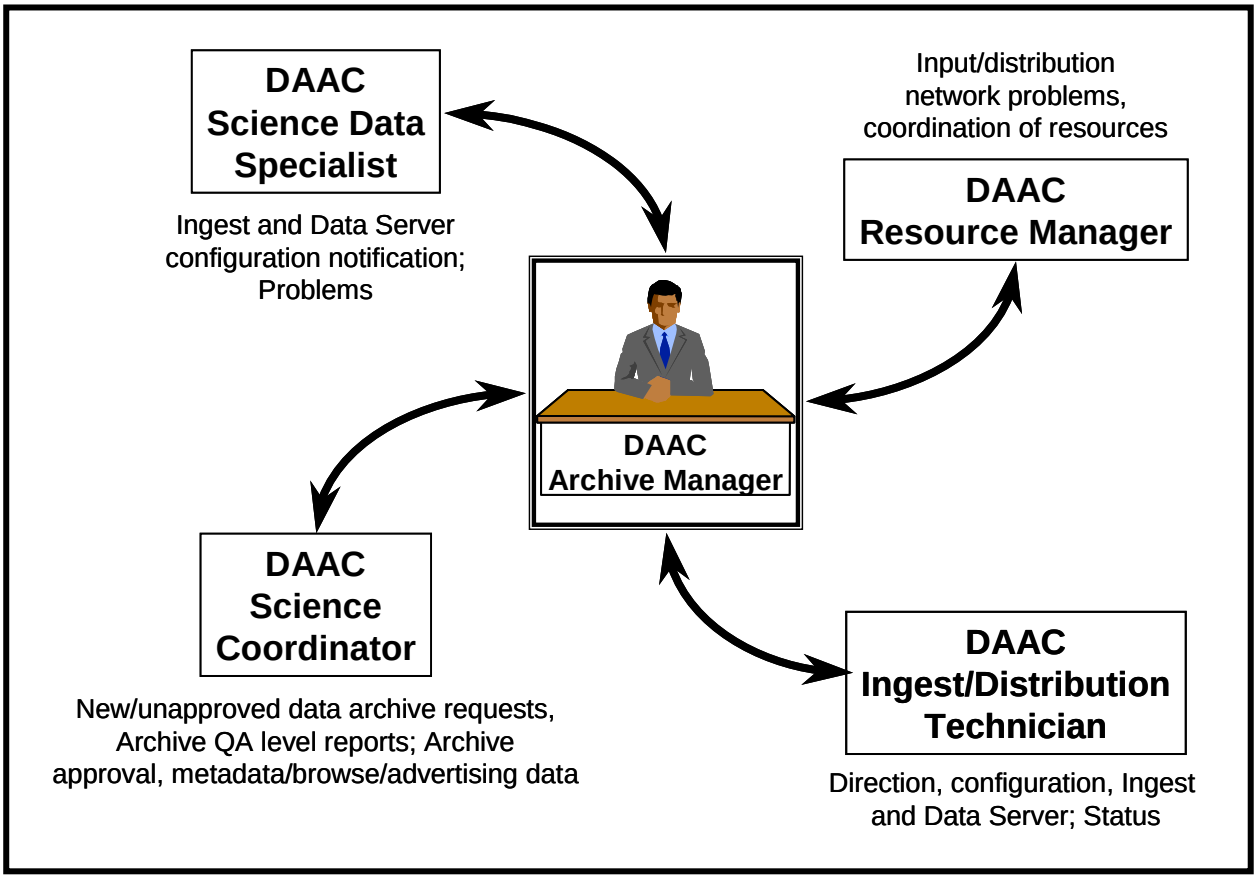


Figure 1. Archive Manager Personnel Interfaces

The Data Server Subsystem (DSS) is responsible for persistent storage of earth science and related data, for providing search and retrieval access to this data, and for supporting the

administration of the data, hardware devices, and software products. As part of its retrieval function, the subsystem also provides for distribution of data electronically or on physical media.

Hardware

The archive uses one major type of archival storage hardware for storing science data, browse data, and other data. The StorageTek (STK) Powderhorn Model 9310 Automated Cartridge System tape storage tower, illustrated in Figure 2, is a mass storage system of removable media jukeboxes. The software that manages the storage in the system architecture is hosted on a Silicon Graphics Inc. (SGI) Origin 2000.

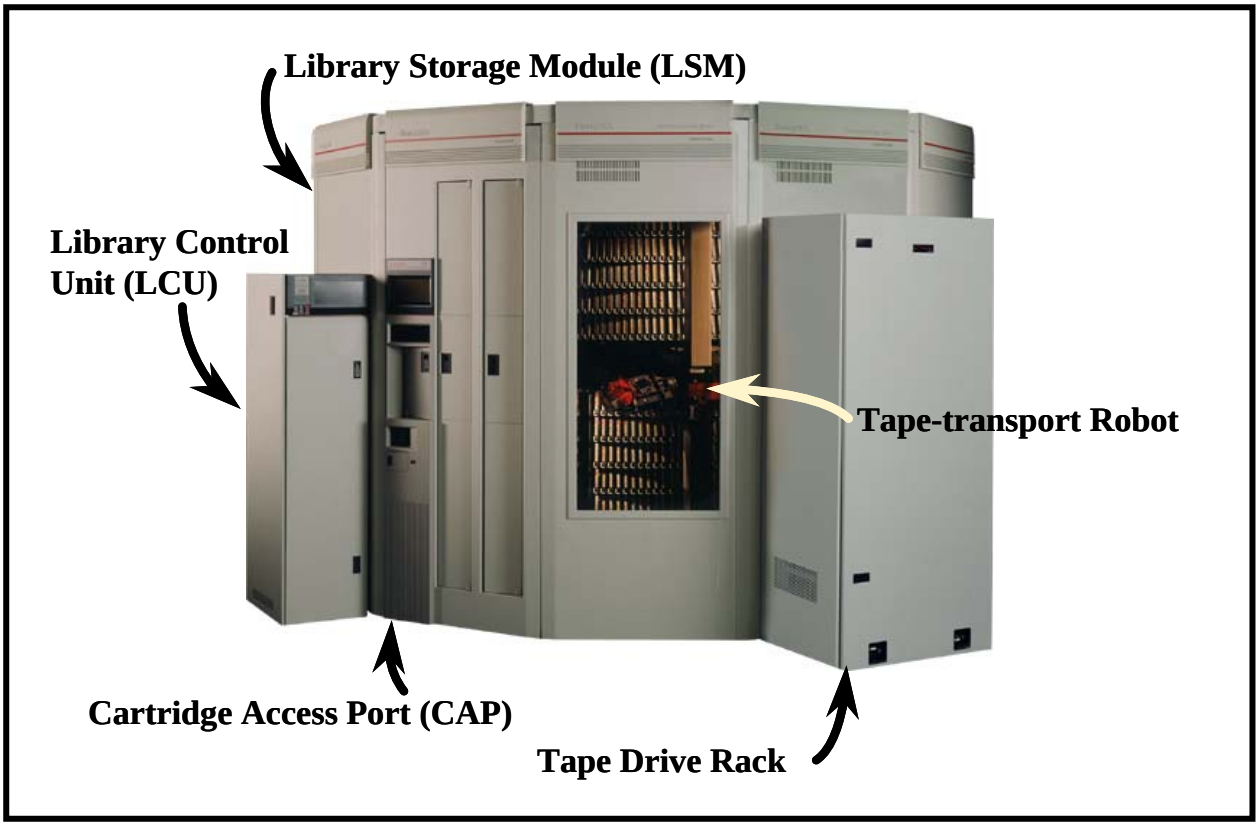


Figure 2. StorageTek (STK) Powderhorn (Model 9310) Data Storage

The typical data storage archive consists of the following major elements:

- Library Storage Module (LSM), Powderhorn Model 9310, Automated Cartridge System (ACS) tape storage tower.
- Cartridge Access Port (CAP), where media are inserted or ejected from the LSM; standard capacity is 21 cartridges.
- Dual tape-transport "robots" for moving cartridges from the tower to a tape drive or CAP and from the tape drive or CAP to the tower.

- Tape drive rack with eight 9940 cartridge tape drives (rack capacity is 20 drives).
- Library Management Unit (LMU), Model 9330, a serial port for the ACS Library Software (ACSLS) that controls and monitors the ACS.
- Library Control Unit (LCU), Model 9311, a hardware interface for managing LSM intercommunications.
- Browse tape drive rack with eight 9840 tape drives (rack capacity is 20 drives).

The LSM tape archive can store thousands of tapes. The archive stores science data on STK 9940 tapes, each of which can store 60 gigabytes of data (up to 200 gigabytes compressed). Browse data are stored on 9840 tapes, each capable of storing 80 gigabytes (compressed) of data. Each 9940 tape cartridge is identified by a colored bar code label that shows the media number (see Figure 3). The 9840 tapes are of the same physical dimensions and use bar code labels, as shown in Figure 4. An archive catalog or database tracks the location of each cartridge within the library, based on information provided by the laser bar code reader.

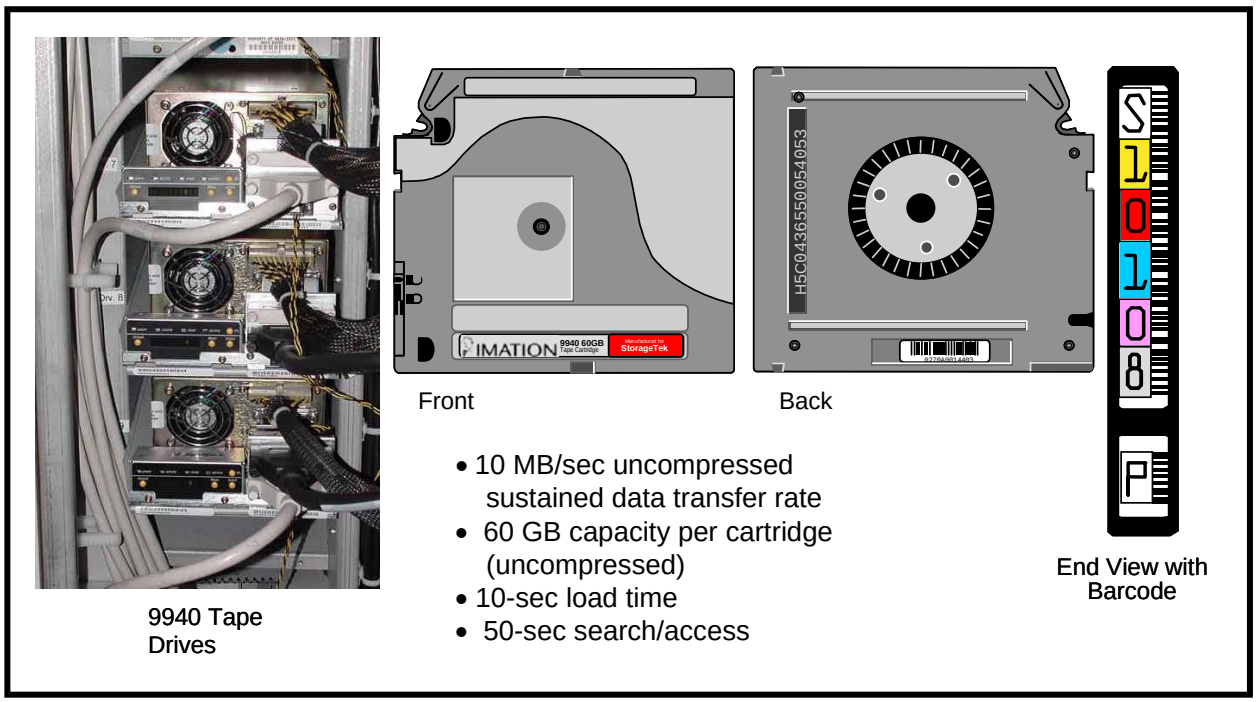


Figure 3. STK 9940 Tape System

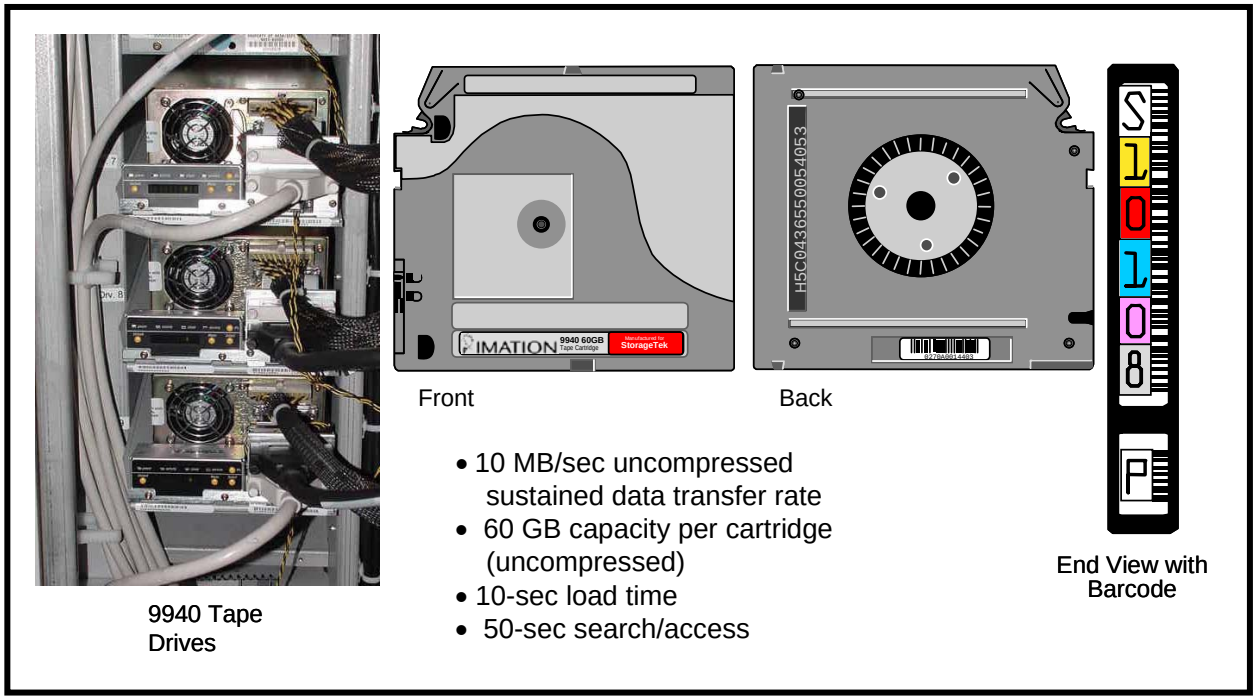


Figure 4. 9840 Tape System for Browse Data

Software

Archive operations rely on both custom and commercial off the shelf (COTS) software for complete mass storage archive management, providing the capability to accept Graphical User Interface (GUI) and command line interface inputs, and to interpret them to the appropriate level needed to control and monitor archive operations. The File Storage Management Service (FSMS) software is the Archival Management and Storage System (AMASS), a product of Advanced Digital Information Corporation (ADIC). The purpose of AMASS in the system is to provide an easy-to-use interface to a large tape archive. AMASS is a UNIX file system that manages files, volumes (media), drives and jukeboxes. It allows UNIX File System (UFS) access methods to be employed (e.g., ftp, rcp, uucp, nfs, RPC, native) while removing some of the limitations of the UFS. Primary among these is reliance on UNIX Index Node (inode) structures. AMASS maintains all inode information in database files rather than in associated disk structures. This minimizes or eliminates many of the file search problems inherent in searching large numbers of files in multiple directories. In addition, AMASS organizes files as groups of blocks which can be individually retrieved. This differs from UFS resident systems that require staging the entire file.

The custom software for the archive provides a standardized format and content for logging of data access and staging activity, which may assist in system troubleshooting. The system also provides parallel AMASS input/output capability for optimum system throughput. Furthermore, the system uses a logical archive ID capability for complete separation of the physical location of data in the archive from its logical reference in the inventory. This means that client requests for data do not change based on a change of the physical location of the data, and as a result there can be improved load balancing, cross-archive fault recovery, and archive upgrades.

AMASS is installed on an SGI Origin 2000 computer. Control information is communicated from the SGI to the LMU using TCP/IP protocols via the FDDI network. Figure 5 shows the basic route taken by control information in the process of sending a file to AMASS.

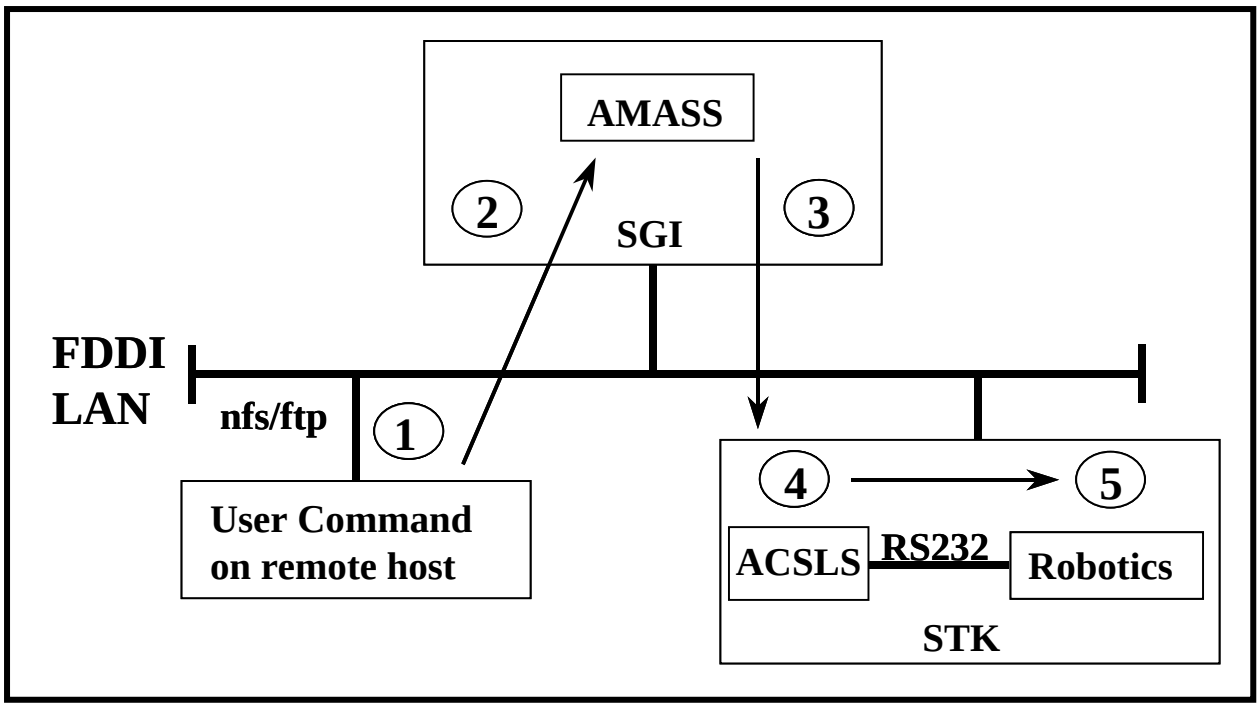


Figure 5. AMASS Control Path

As the figure suggests, there are five elements in the path:

1. The user or application initiates transfer of the file to AMASS.
2. AMASS receives the file over the network via nfs, ftp, dd, or cp, or locally via dd or cp to its cache.
3. AMASS sends information to ACSLS to specify the tape to load.
4. The ACSLS application sends the LMU robot to retrieve the tape.

5. The robot grips the tape, retrieves it, and inserts it into the tape drive to complete the mount.

ACSLS

For the StorageTek Powderhorn, Steps 4 and 5 in the AMASS control path are managed by the Automated Cartridge System Library Software (ACSLS), running on a Sun computer under Solaris 2.6. Full guidance for using ACSLS is provided in the *Automated Cartridge System Library Software System Administrator's Guide, Version 5.3*. Table 1 lists the commands covered in that *Guide*.

Table 1. ACSLS Command Reference

Command	Function
audit	Creates or updates the database inventory of the volumes in a library component.
cancel	Cancels a current or pending request.
clear lock	Removes all active and pending locks on transports or volumes.
define pool	Creates or modifies scratch pools.
delete pool	Deletes empty scratch pools.
dismount	Dismounts a volume.
eject	Ejects one or more volumes from the Automated Cartridge System (ACS).
enter	Sets a Cartridge Access Port (CAP) to enter mode.
idle	Stops ACSLS from processing new requests.
lock	Locks (dedicates) a volume or transport to a user.
logoff	Exits the command processor.
mount	Mounts a data or scratch volume.
query	Displays the status of a library component.
set	Sets various attributes of different library components.
show	Displays your lock ID or user ID.
start	Starts ACSLS request processing.
unlock	Removes active locks on volumes or transports.
vary	Changes the state of an ACS, LSM, CAP, transport, or port.
venter	Enters one or more volumes with missing or unreadable labels into the ACS.

ACSLS commands use the following general syntax:

command type_identifier state [options]

where **type_identifier** is the ACS component and its identifier (these are listed in the *System Administrator's Guide*), **state** is a device state for the **vary** command only, and **options** are command options (these are specified for each command in the *System Administrator's Guide*. The two most useful commands in ACSLS are **query** and **vary**. Other frequently used commands are **enter** and **eject**, for inserting and removing cartridges, respectively. ACSLS does

not have an online help facility, but if you enter a command (e.g., **vary**), it will prompt you for the parameters.

There are also several utilities provided with ACSLS. These are listed in Table 2.

Table 2. ACSLS Utilities

Utility	Function
bdb.acsss	Backs up the ACSLS database.
kill.acsss	Terminates ACSLS.
rc.acsss	Starts and recovers ACSLS.
rdb.acsss	Restores the ACSLS database.
volrpt	Creates a volume report.
db_command	Starts or stops the ACSLS database.

To control and interact with ACSLS, you use the following user IDs:

- **acssa** lets you enter ACSLS commands from a command processor window.
- **acsss** lets you run ACSLS utilities from the UNIX command line prompt.

It is typical to log in as both user IDs to permit entering both ACSLS utilities and commands. You can, however, open a command processor window from the **acsss** user ID if you prefer to work from a single user ID. The *System Administrator's Guide* provides full details.

Data Sources and Uses

Data that are inserted into the archive are managed by the Data Server Subsystem (DSS) and can be received from such sources as the ingest subsystem, processing subsystem, other DAACs, and authorized users (Figure 6). Uses of data from these sources include:

- from ingest – any system function that uses data (e.g., production).
- from processing – various system functions (e.g., further processing, distribution to users).
- from other DAACs – various system functions (e.g., may be needed as inputs for production of other products).
- from authorized users (via ingest) – typically for distribution or processing.

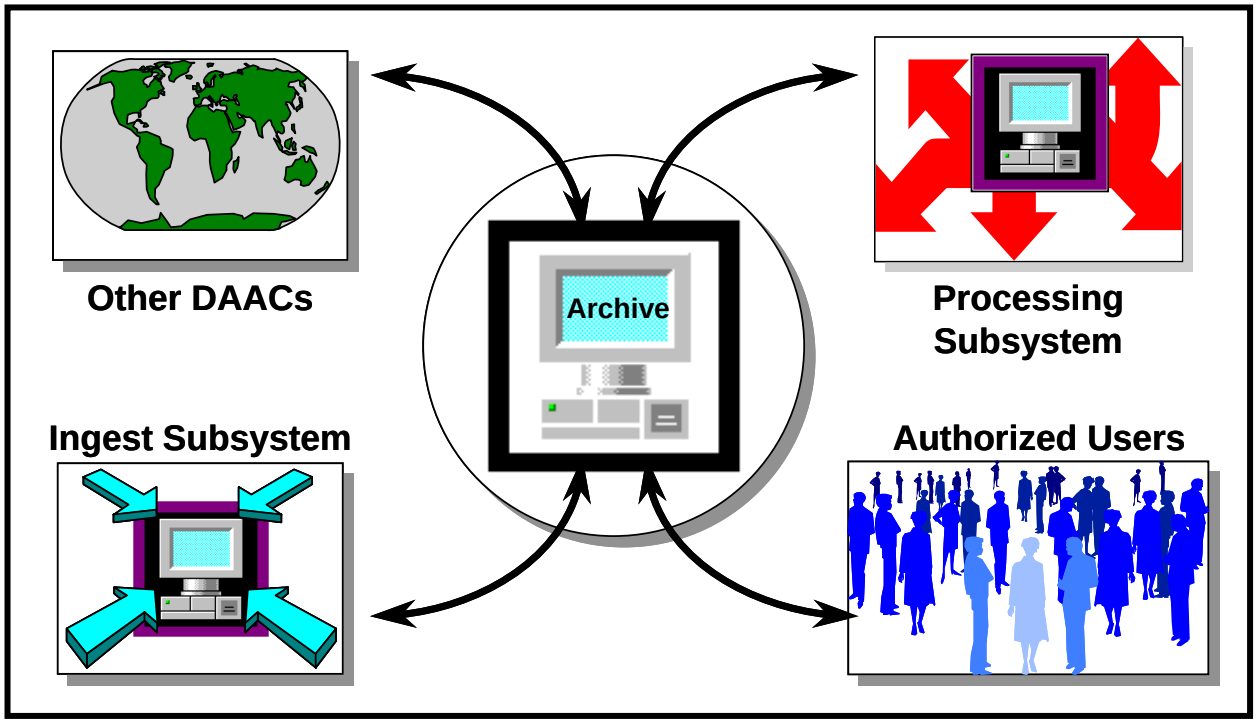


Figure 6. Sources and Uses of Archive Data

Starting and Shutting Down AMASS

The AMASS Tape Archive System can be started and shutdown with little or no impact on the rest of the system.

Starting the AMASS Tape Archive System

Starting the AMASS FSMS requires actions to ensure that the STK Powderhorn storage system is powered up as well as actions at the SGI FSMS host. Powering up the STK will require actions at its control panels, including the Library Management Unit (LMU) and Library Control Unit (LCU) [the Library Storage Module (LSM) is powered through the LCU]. Figures 7 through 9 illustrate major locations of relevant controls and displays.

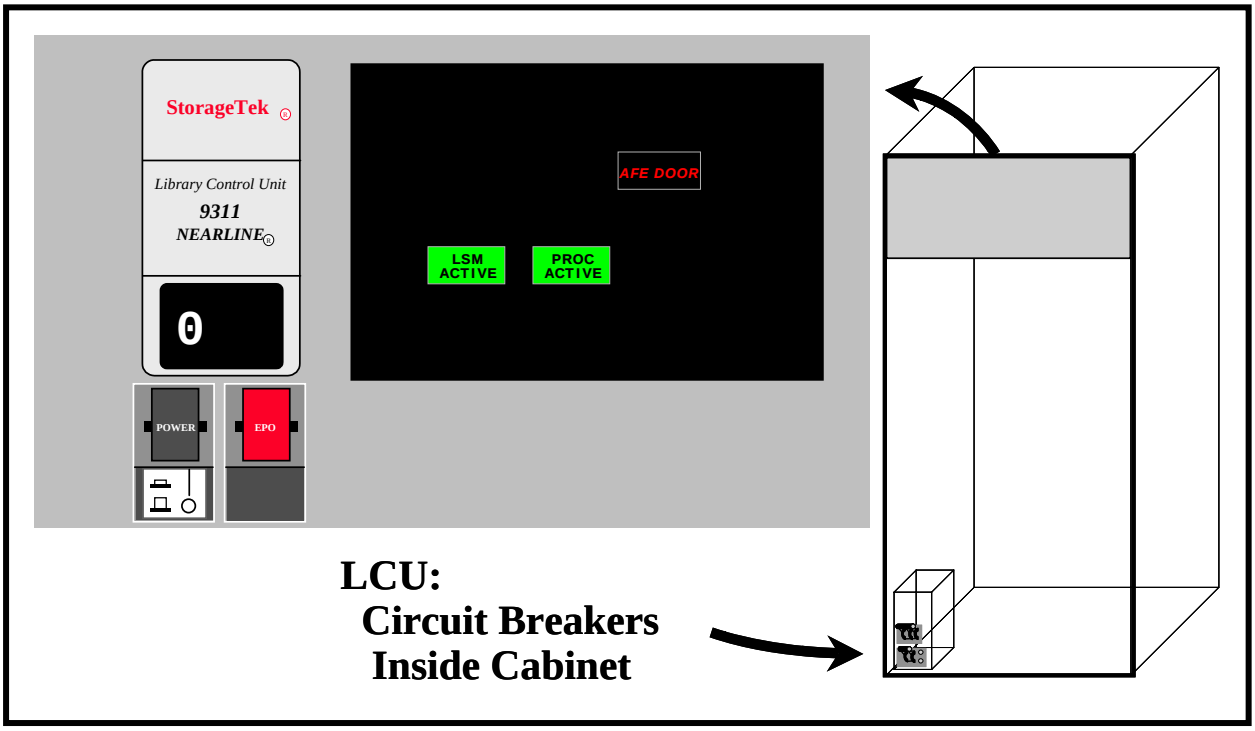


Figure 7. Controls/Displays for the STK Library Control Unit (LCU)

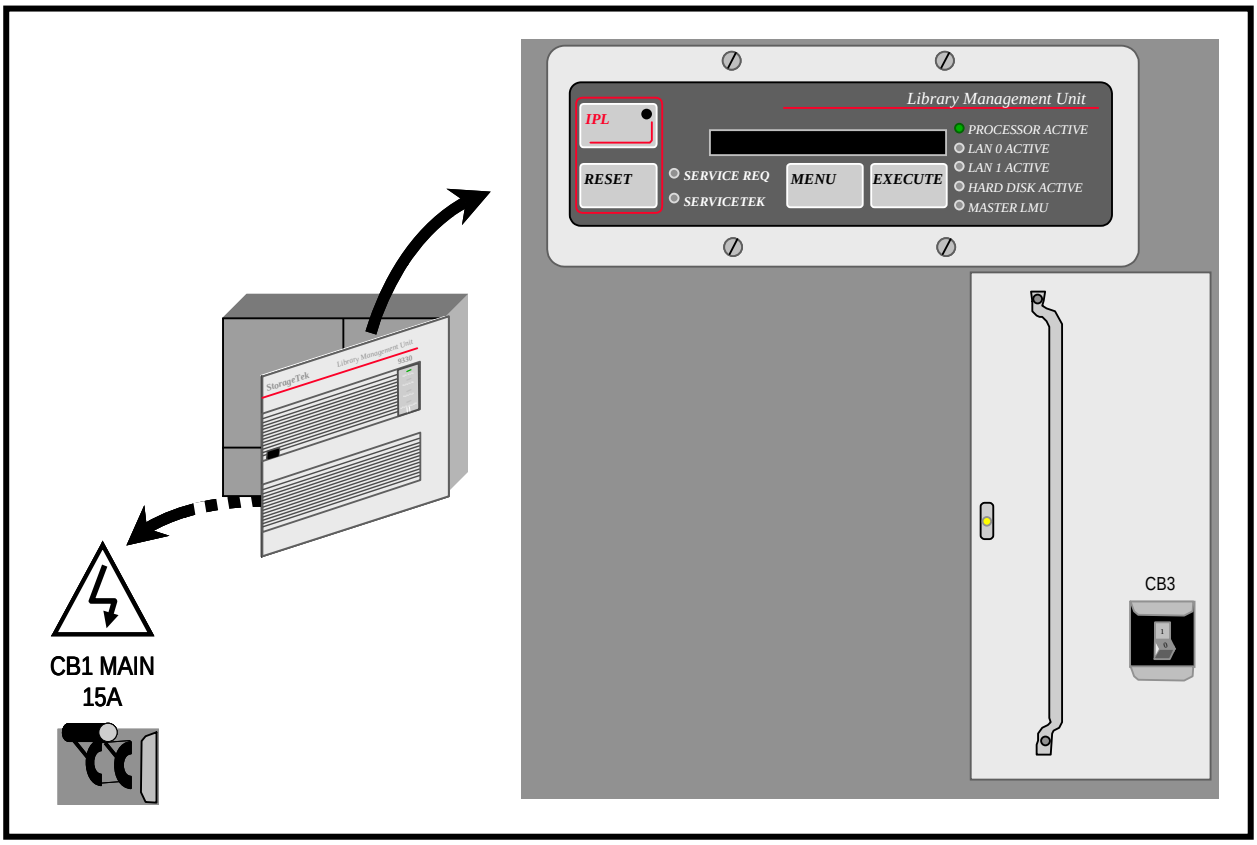


Figure 8. Controls/Displays for the STK Library Management Unit (LMU)

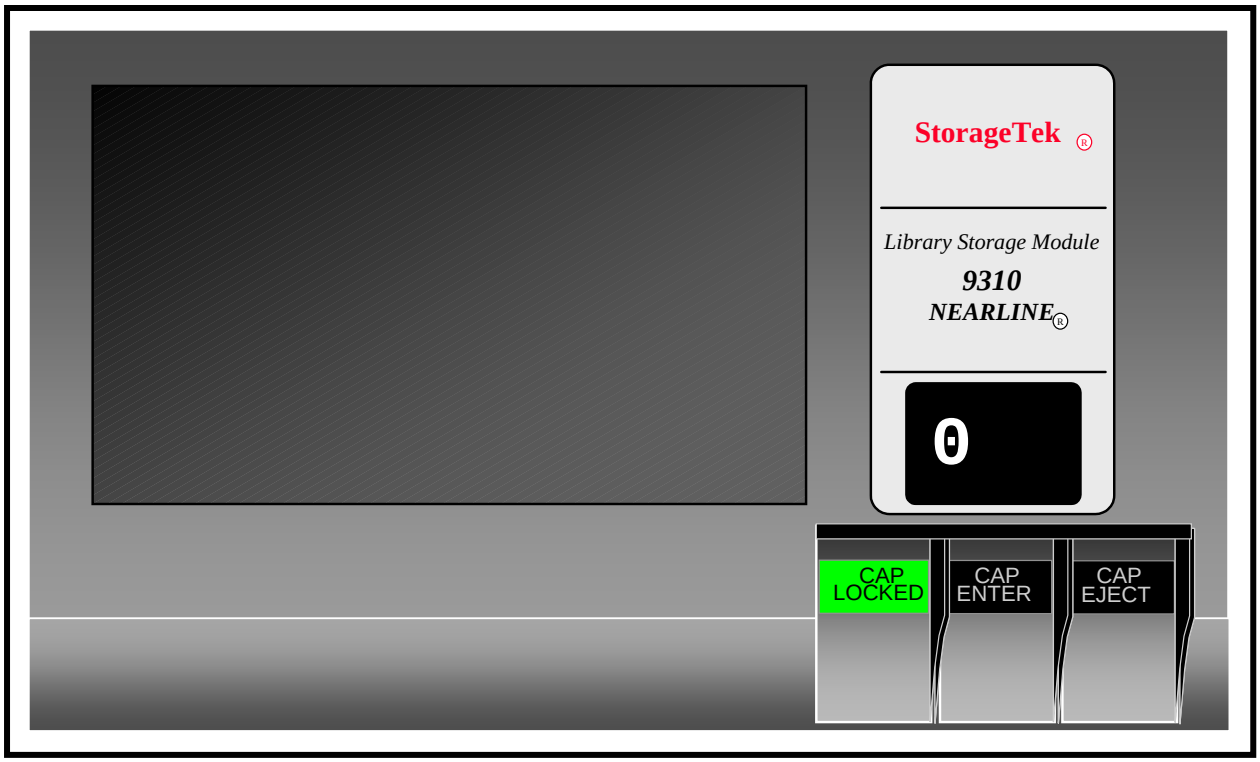


Figure 9. Controls/Displays for the STK Library Storage Module (LSM)

Use the following procedure to start the AMASS software.

Start the AMASS Application

NOTE: Preconditions include that 1) the FDDI network is up and running and 2) power to all units is functional and available.

1 Make sure power switches for the StorageTek LCU and LMU are **ON**.

NOTE: The LCU should be the last unit powered up, but otherwise there are no dependencies within the group.

2 If it is not already running, boot the FSMS SGI host normally.

- There are no dependencies on other hosts, COTS or custom software.
- AMASS normally starts automatically on bootup. If it does, go to **Step 4**. If it does not, or if you are restarting AMASS after a shutdown, go to **Step 3**.

3 At the FSMS SGI host (workstation ***x*0drg##**), as a system administrator (logged in as **root**), type **/usr/amass/tools/amass_start** and then press the **Return/Enter** key.

- The ***x*** in the workstation name will be a letter designating your site: **g** = GSFC, **m** = SMC, **l** = LaRC, **e** = LP DAAC, **n** = NSIDC, **o** = ORNL, **a** = ASF, **j** = JPL; the **##** will be an identifying two-digit number (e.g., **n0drg01** indicates an FSMS SGI server at NSIDC).
- The AMASS application starts.

4 To verify that AMASS has started correctly, type **/usr/amass/bin/amassstat -c** and then press the **Return/Enter** key.

- The message **FILESYSTEM IS ACTIVE** is displayed.

Shutting Down AMASS Tape Archive System

If it is necessary to shut down AMASS, use the following procedure.

Shut Down the AMASS Application

1 Log in as **root** (system administrator) at the FSMS SGI host (workstation ***x*0drg##**,).

- The ***x*** in the workstation name will be a letter designating your site: **g** = GSFC, **m** = SMC, **l** = LaRC, **e** = LP DAAC, **n** = NSIDC, **o** = ORNL, **a** = ASF, **j** = JPL; the **##** will be an identifying two-digit number (e.g., **n0drg01** indicates an FSMS SGI server at NSIDC).

2 Type **/usr/amass/tools/killdaemons** then press the **Return/Enter** key.

- A message is displayed indicating that all daemons have been terminated.

Rebooting AMASS

The AMASS file system may need to be rebooted during certain anomalous conditions (e.g., system "hang," interruption of communication between AMASS and ACSLS, a required daemon is down). AMASS needs to have the following daemons running at all times: amassmain, daemons/lm_ip -a fslock, klogd, amass_iocomp, qset, libsched, and libio_tape. To verify they are running, simply search for the AMASS processes (refer to Procedure 17.7.1.1, **Checking Daemons and Using *healthcheck*** in Document 611-EMD-001, *Release 7.11 Mission Operation Procedures for the EMD Project*). To check the health of AMASS while it is still running, execute the **healthcheck** command (refer to Procedure 17.7.1.1 in Document 611-EMD-001, *Release 7.11 Mission Operation Procedures for the EMD Project*).

In order to reboot AMASS you must have root privileges. Use the following procedure.

Reboot AMASS

---

1 Log in as **root** (system administrator) at the FSMS SGI host (workstation ***x*0drg##**).

- The ***x*** in the workstation name will be a letter designating your site: **g** = GSFC, **m** = SMC, **l** = LaRC, **e** = LP DAAC, **n** = NSIDC, **o** = ORNL, **a** = ASF, **j** = JPL; the **##** will be an identifying two-digit number (e.g., **n0drg01** indicates an FSMS SGI server at NSIDC).

2 To kill the daemons, type **killdaemons** and then press the **Return/Enter** key.

- A message is displayed indicating that all daemons have been terminated.

3 If you want to test AMASS before restarting, go to step 4; otherwise, type **amass_start** and then press the **Return/Enter** key.

- The AMASS application starts.

4 To test the AMASS file system prior to starting AMASS type: **install_tests** then press the **Return/Enter** key.

- Tests the jukebox operation and cache partitions, then restarts AMASS.

Entering the Archive After AMASS Is Started

If it is necessary to enter the STK Powderhorn after AMASS is started, use the following procedure.

Enter the STK Powderhorn

---

1 At the host for ACSLS (e.g., **e0drs03**, **g0drs04**, **l0drs02**, **n0drs03**), log in using the **acssa** user ID and password.

- The **acssa** command-process window is displayed with the **ACSSA>** prompt.

2 Type **vary lsm 0,0 offline** and then press the **Return/Enter** key.

- The access port is unlocked (audible unlatching sound).

3 Use the key to unlatch and open the access door.

- A red **DO NOT ENTER** warning is visible inside the enclosure.

Warning

If it is necessary to enter the STK Powderhorn after AMASS is started, it is necessary to perform the following step to avoid hazard and ensure safety of personnel and equipment.

4 Remove the key from the door to ensure that no one inadvertently locks the enclosure with someone inside.

- The red **DO NOT ENTER** warning is extinguished and a green **ENTER** message is displayed inside the enclosure.

5 Upon leaving the enclosed area, insert the key in the access door and latch the door.

- The LED display indicates that the door is locked.

6 At the ACSLS host, type **vary lsm 0,0 online** and then press the **Return/Enter** key.

- After a few seconds, the archive robots execute an initialization sequence and the LSM is back online.

Logging in to System Hosts

Logging in to System Hosts

Logging in to system hosts is accomplished from a UNIX command line prompt. It is an initial set of steps that is performed when accomplishing many other tasks.

Logging in to system hosts starts with the assumption that the applicable hosts are operational and the operator has logged in to a workstation or X-term that has access to the applicable network in the system.

Logging in to System Hosts

NOTE: Commands in Steps 1 and 2 are typed at a UNIX system prompt.

1 At the UNIX command line prompt type **setenv DISPLAY *clientname*:0.0** then press the **Return/Enter** key.

- 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 typing **/tools/bin/ssh *hostname*** then press **Return/Enter**.

- 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 x0acs06 as user cmops type **/tools/bin/ssh -l cmops x0acs06** then press **Return/Enter**.
- An alternative method of logging in with a different user ID is to type the command in the format **/tools/bin/ssh *userid@hostname***. For example:

  /tools/bin/ssh cmops@x0acs06

- 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 x0acs06

  - or -

  ssh -l cmops x0acs06

- or -

ssh cmops@x0acs06

- Examples of Data Pool Server host names include **e0dps01**, **g0dps01**, **l0dps01**, and **n0dps01**.
- Examples of Sun Consolidation Internal Server host names include **e0acs11**, **g0acs11**, **l0acs03**, and **n0acs04**.
- Examples of Sun Consolidation External Server host names include **e0ins01, g0ins01, l0ins01**, and **n0ins02**.
- Examples of APC Server host names include **e0acg11**, **g0acg01**, **l0acg02**, and **n0acg01**.
- Examples of ACSLS Workstation host names include **e0drs03**, **g0drs04**, **l0drs02**, and **n0drs03**.
- Examples of FSMS Server host names include **e0drg11**, **g0drg01**, **l0drg01**, and **n0drg01**.
- Examples of Operations Workstation host names include **e0acs12**, **g0acs02**, **l0acs01**, and **n0acs03**.
- Examples of Ingest Server host names include **e0icg11, g0icg01, l0acg02**, and **n0acg01**.
- 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, type your ***passphrase*** then press **Return/Enter**.

- 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, type your ***password*** then press **Return/Enter**.

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

This page intentionally left blank.

Archive Storage Structures

Although the physical archive stores data on cartridges (referred to as volumes), it is treated in the system as a large UNIX directory. As such, it is prudent to apply good space management practices in managing the archive just as you would for any computer disk. However, there are other factors that must be considered in archive management, addressing unique system requirements (e.g., on-site backups, off-site backups, desirability of physical consolidation of related data).

Storage Element Relationships

It is important for the Archive Manager to know the relationship between physical storage archives (Library Storage Modules, or LSMs) and the Archive Server software applications at the site. For example, a data repository identified as DRP1 is served by the software application EcDsStArchiveServerDRP1.

Subdivisions within LSMs (e.g., for storage of different data types) are reflected in the Storage Management database, where each Volume Group (a logical group of volumes in the archive) has its own path. As suggested in Figure 10, each path maps to an AMASS volume group, and thus to a physical volume group in the archive.

Information concerning archive servers and the logical volume groups served may be obtained from the Storage Management Control Graphical User Interface (GUI). Specifically, as illustrated in Figure 11, the Storage Configuration tab on the Storage Management GUI permits display of server information and access to related status information.

The Vol Grp Config. tab, illustrated in Figure 12, permits display of volume group information and history.

Launching DSS GUIs

The following software applications are associated with DSS:

- Science Data Server (SDSRV).
- Storage Management (STMGT) Servers.
  - Request Manager Server.
  - Staging Disk Server.
  - Cache Manager Server.
  - Archive Server.
  - FTP Server.

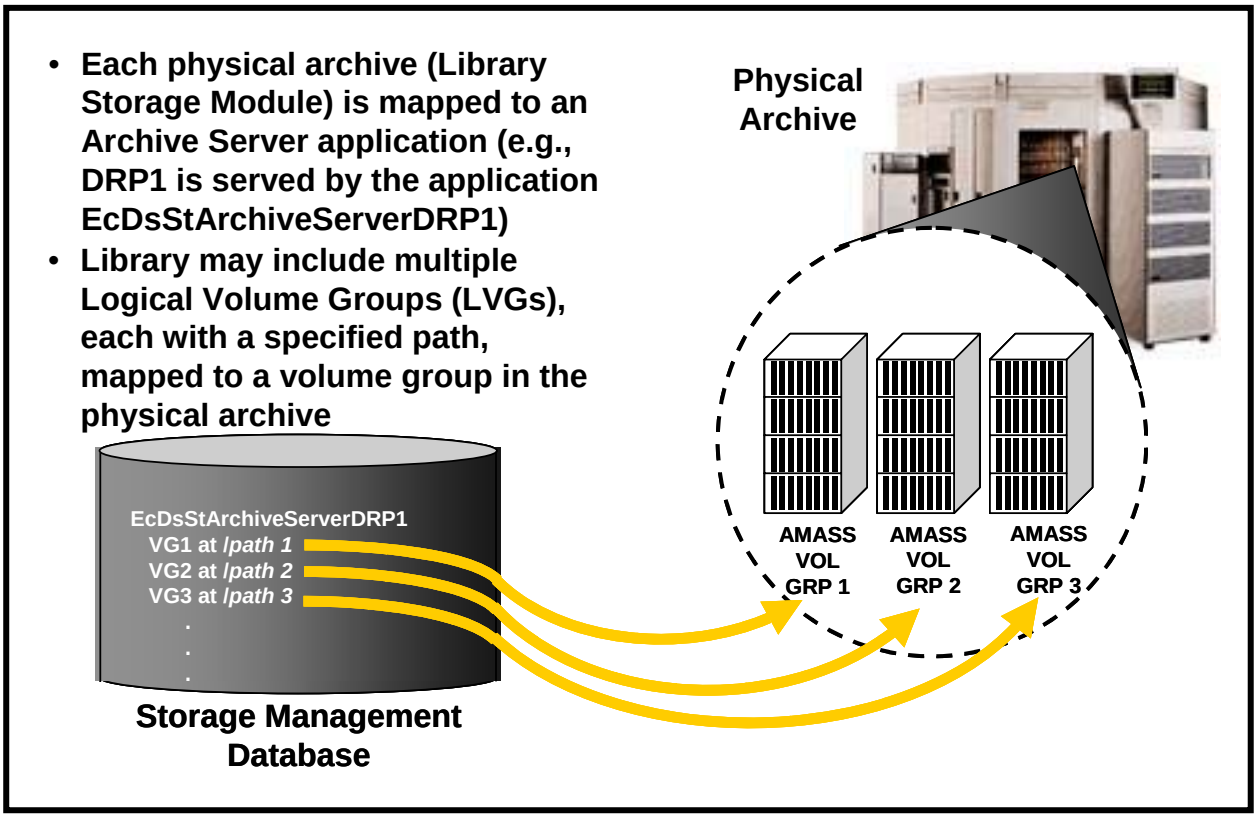


Figure 10. Archive Storage Element Relationships

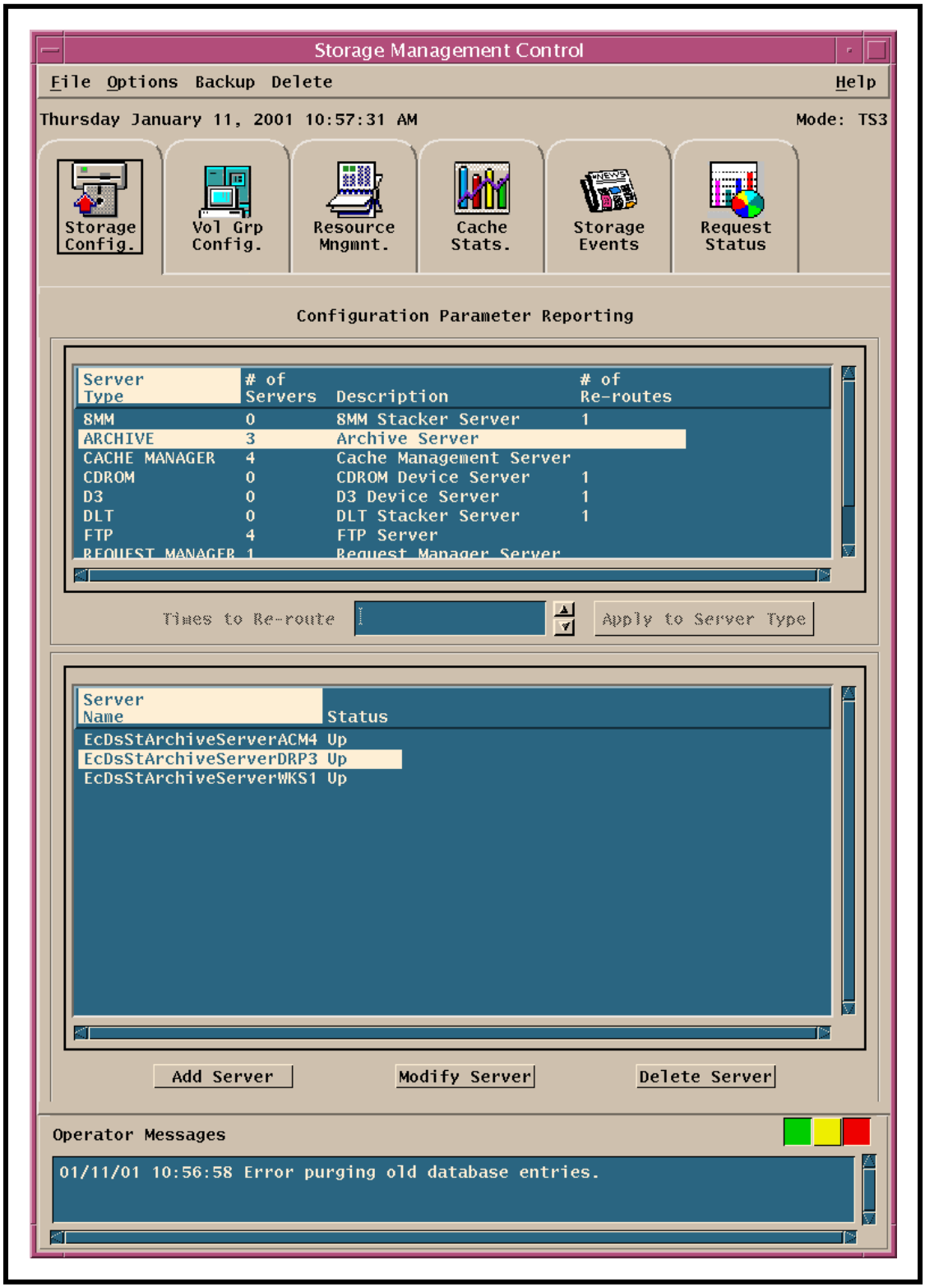


Figure 11. Storage Management Storage Configuration Tab

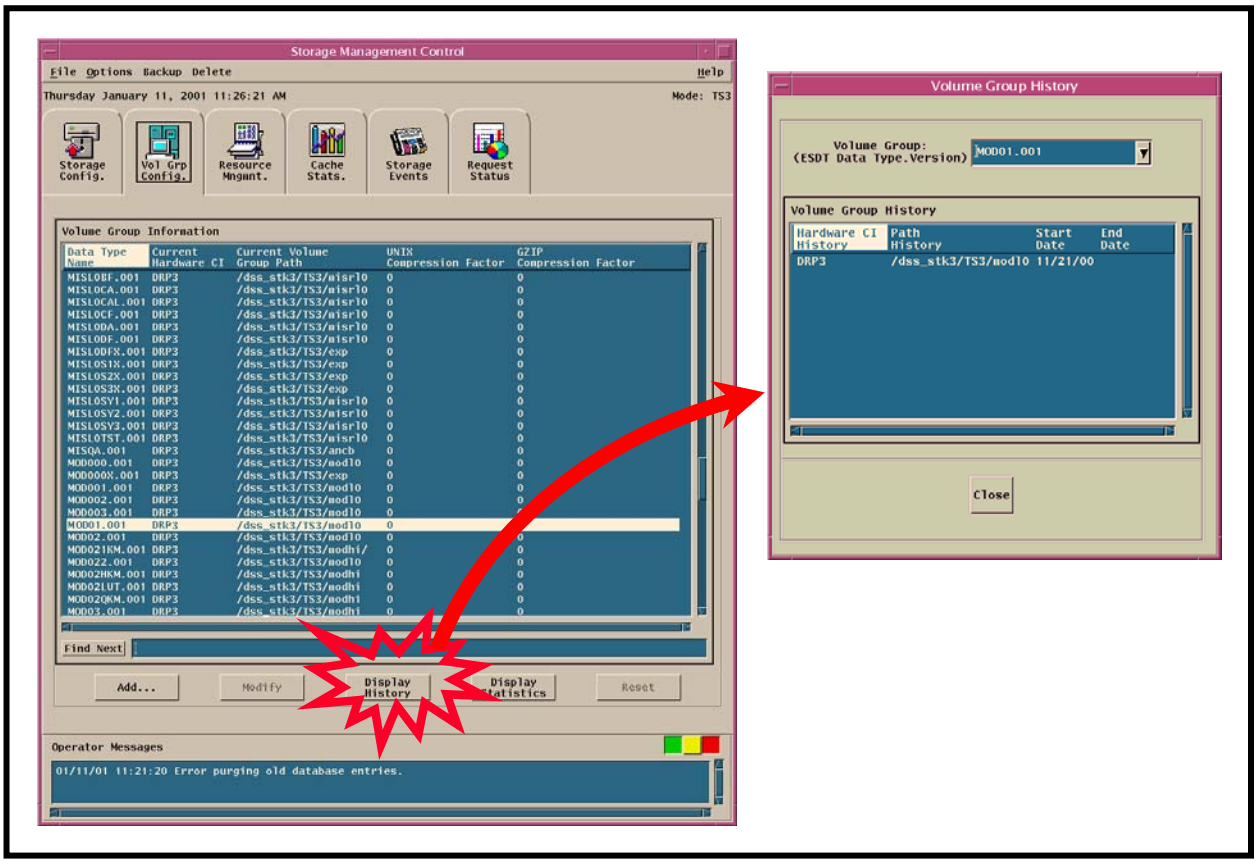


Figure 12. Storage Management, Vol Grp. Config Tab

  - 9940 Tape Server.
  - 8mm Tape Stacker Server.
- Data Distribution (DDIST) Server.
- DDIST Graphical User Interface (GUI).
- STMGT GUI.
- Science Data Server GUI.

Access to Storage Management, Data Distribution (DDIST), and other GUIs is gained through the use of UNIX commands. The procedure for launching the GUIs begins with the assumption that the applicable servers are running and that the operator (Archive Manager or System Administrator) has logged in.

Launch DSS GUIs Using UNIX Commands

1 Access the command shell.

- The command shell prompt is displayed.

NOTE: Commands in Steps 2 through 9 are typed at a UNIX system prompt.

2 Type **setenv DISPLAY *clientname*:0.0** and then press the **Return/Enter** key.

- Use either the terminal/workstation IP address or the machine-name for the ***clientname***.

3 Start the log-in to the appropriate operations workstation (e.g., **e0acs03**, **g0acs02**, **l0acs01**, or **n0acs03**) by typing **/tools/bin/ssh *hostname*** and then press the **Return/Enter** key.

- 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)?** type **yes** ("**y**" alone does 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 4.
- If you have not previously set up a secure shell passphrase, go to Step 5.

4 If a prompt to **Enter passphrase for RSA key '<*user@localhost*>'** appears, type your ***Passphrase*** and then press the **Return/Enter** key. Go to Step 6.

5 At the **<*user@remotehost*>'s password:** prompt, type your ***Password*** and then press the **Return/Enter** key.

6 To change directory to the directory containing the startup scripts for DSS, type **cd /usr/ecs/<*MODE*>/CUSTOM/utilities** and then press the **Return/Enter** key.

- The <***MODE***> will most likely be one of the following operating modes:
  - OPS (for normal operation).
  - TS1 or TS2 (for testing).
- Note that the separate subdirectories under /usr/ecs apply to different operating modes.

7 To launch the Storage Management Control GUI, type **EcDsStmgtGuiStart <*MODE*>** then press the **Return/Enter** key.

- The Storage Management Control GUI, used for review of storage events and status **of devices, is displayed.**

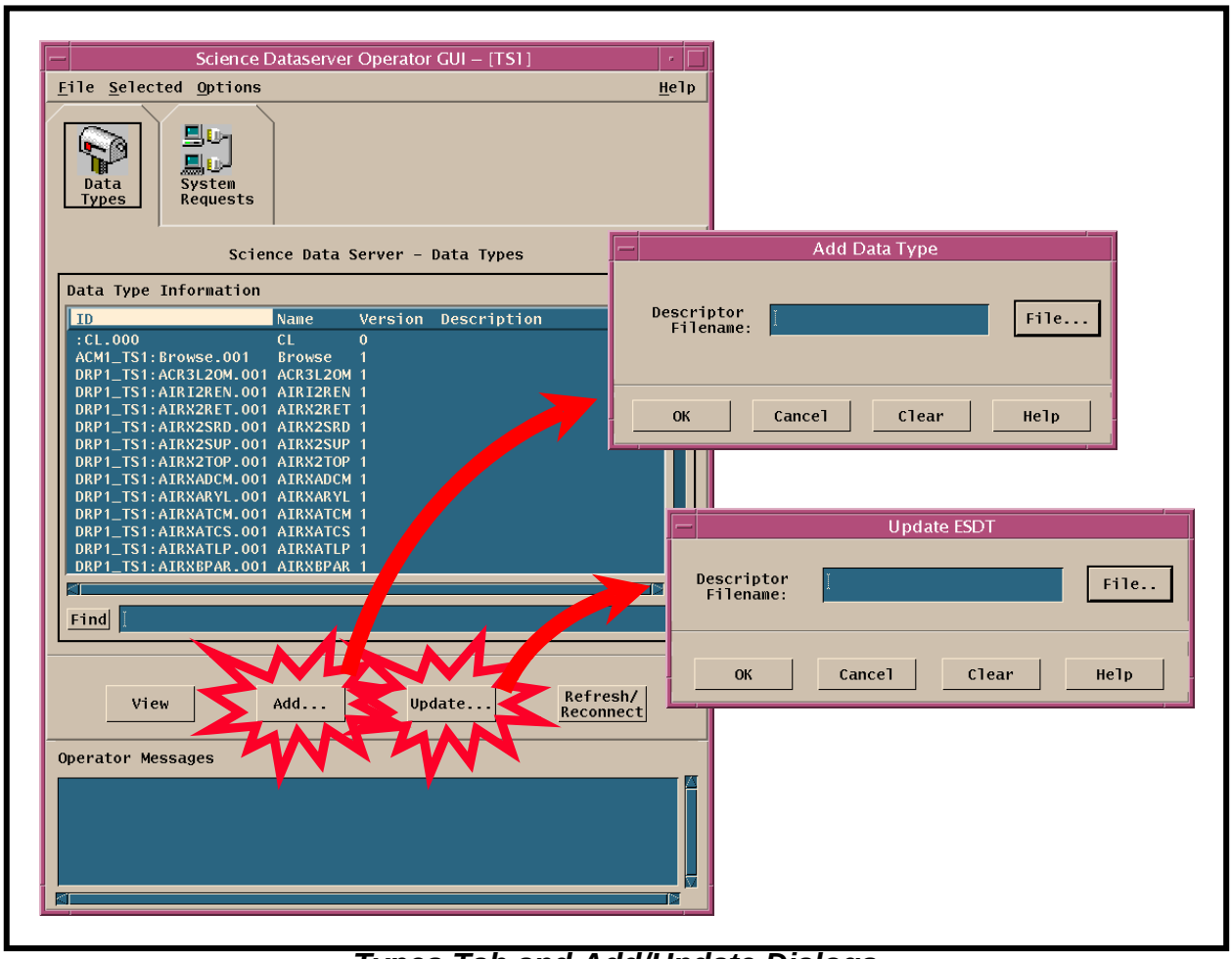
8 To launch the Data Distribution GUI, use a similar procedure and type **EcDsDdistGuiStart <*MODE*>** then press the **Return/Enter** key.

- The Data Distribution GUI is displayed.

9 To launch the **DSS Science Data Server** GUI, use a similar procedure and type **EcDsSdSrvGuiStart** *<MODE>* then press the **Return/Enter** key.

- The Science Data Server Operator GUI is displayed.

Suppose you are an Archive Manager and are asked by someone on the SSI&T team to provide archive path information for a particular Earth Science Data Type (ESDT) stored in the archive. The SSI&T team can tell you the name for the ESDT, because that information is typically in the descriptor file specified when an ESDT is loaded, using the **Data Types** tab of the Science Data Server GUI and the **Add Data Type** dialog illustrated in Figure 13. The figure also illustrates the dialog used to **Update ESDT** information, providing the capability to make updates without reloading ESDTs when changes are necessary.***Figure 13. Science Data Server Data***



Types Tab and Add/Update Dialogs

Use the following procedure to display specific path information for the archive server.

Display Archive Path Information Using Storage Management GUIs

1 Launch the DSS Storage Management GUI using UNIX commands.

- For detailed instructions refer to the **Launch DSS GUIs Using UNIX Commands** procedure (previous section of this lesson).
- The DSS Storage Management GUI is displayed.

2 Click on the **Storage Config.** tab to ensure that the Storage Configuration display is available.

- The **Storage Config.** tab is displayed.

3 In the field listing **Server Type** click on the **ARCHIVE** line to highlight it.

- The selected line is highlighted and the **Server ID** and **Status** of archive servers are displayed in the field listing **Server Name**.

4 Click on the **Vol Grp Config.** tab.

- The **Volume Group Information** is displayed showing volume groups and their current paths.

5 If it is desirable to display the path history for a data type, on the **Vol Grp Config.** tab, click on the **Data Type Name** entry for the specific server for which path history information is desired.

- The selected line is highlighted.

6 Click on the **Display History** button.

- A **Volume Group History** window is displayed showing the path history for the highlighted data type.

Archive Resource Management

The management of archive resources and data is governed by local policy. The software permits sites to establish unique naming and mapping conventions for relating ESDTs to logical volume groups and physical archive volume groups. However, to foster consistency and ease of management of the data in the archive, and to increase the supportability of the system, it is desirable to establish and follow conventions (e.g., naming conventions, assignment of data to volume groups). To facilitate support of local archives by centralized resources (e.g., the System Monitoring and Coordination Center), it is desirable that the DAACs work together (e.g., through Operations Working Groups) to implement consistent and compatible data management practices. Of course, it is necessary for each DAAC to enforce policies and procedures to ensure the long-term viability of archived data.

Logical volume groups are specified using the ESDT short name with the version ID as an extension (e.g., **MOD01.001**). Achieving this convention can be facilitated by adoption of an approach in which specific data products are assigned to the same storage path. In fact, it may be desirable to assign all related products (e.g., **MOD01.001**, **MOD01.002**, **MOD01.003** . . .) to the same storage path. This will have the effect of consolidating related products ultimately in the same physical archive volume group.

The Storage Management GUIs provide tools for managing archive resources. As noted previously, the **Storage Config.** tab (see Figure 11) provides information and control functions for setting and modifying configurations of various Server Types (e.g., 8mm tape, Archive Server) and to manage data location within the archive and on disk. Another Storage Management GUI is the **Resource Mngmnt.** tab, shown in Figure 14. It allows the operator to monitor and adjust the availability of given storage devices. Buttons provide access to pop-up windows permitting management of hardware, including the ability to put specific resources on line or take them off line, and management of media sets.

Another Storage Management GUI is the **Cache Stats.** tab, shown in Figure 15. It permits the operator to view various system caches, along with statistics on their use.

The custom software implements an active "just-enough" cache cleanup strategy. Under this strategy, the Cache Manager identifies and removes just enough to accommodate new files. Once full, STMGT caches and the pull area will generally remain full, but this does not require any action by the operator. Instead, the system assesses how much cache or pull disk space is needed and removes expired files only to the extent needed to provide the needed space.

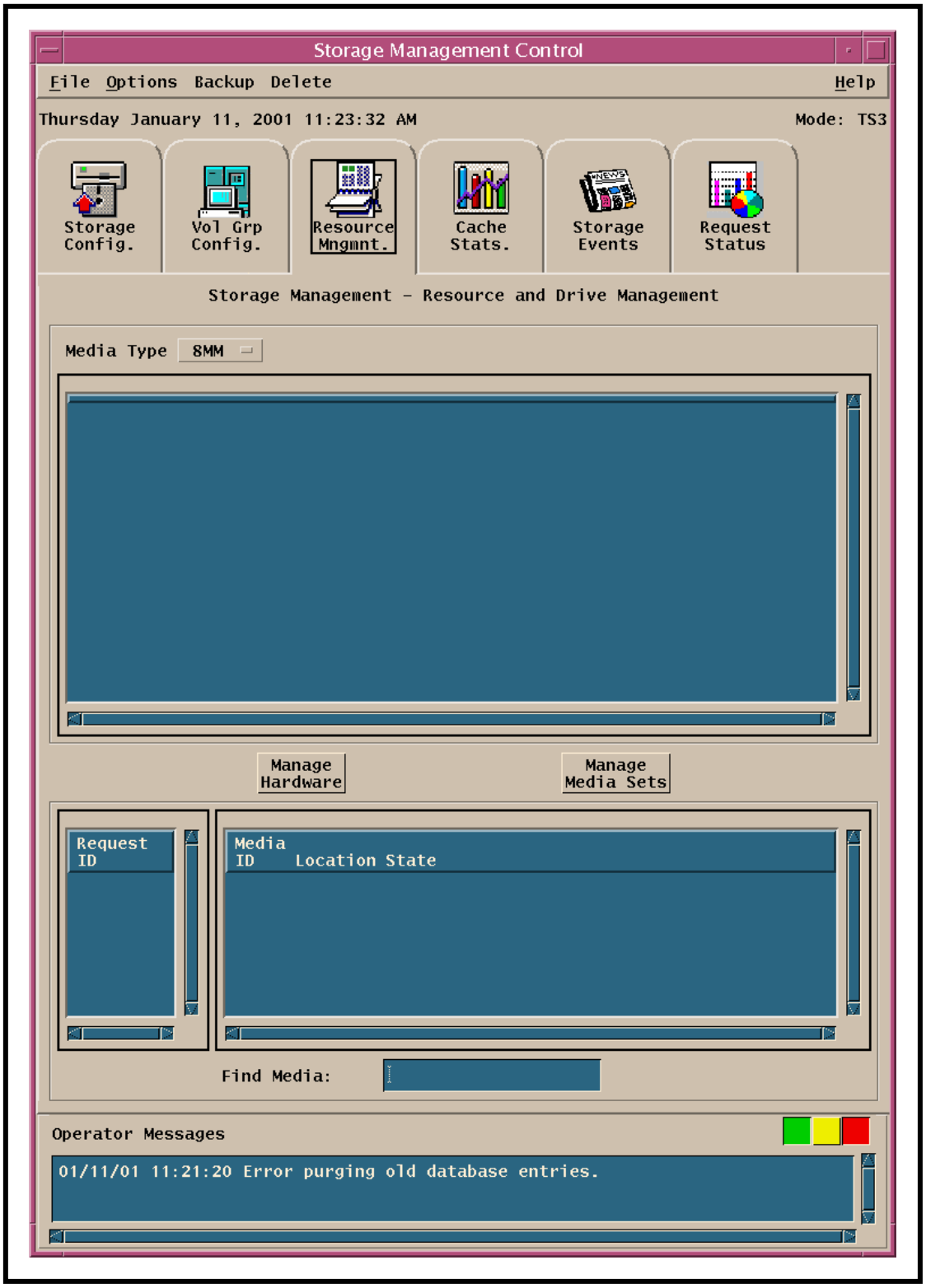


Figure 14. Storage Management Resource Mngmnt Tab

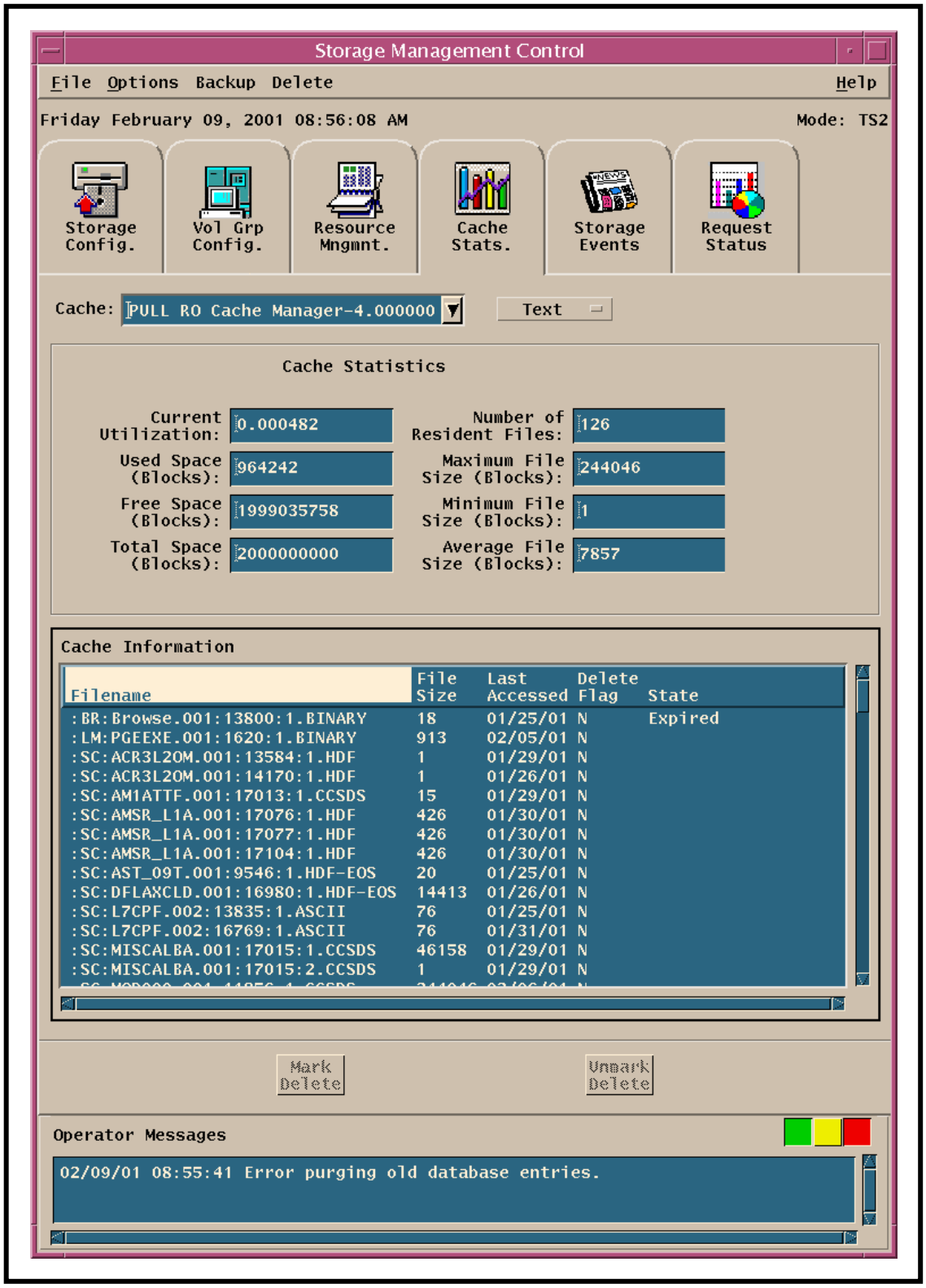


Figure 15. Storage Management Cache Stats Tab

Insert Data Into the Archive

Storing new data in the archive repository is largely an automated process that does not normally require operator interaction and occurs as a result of operations such as ingest and data production. Any operations involvement would be to support archive administration operations, resolve problems, periodically monitor working storage and archival operations, and coordinate with the appropriate external/internal sources to resolve schedule conflicts.

Because of the automated nature of this process, step-by-step procedures are not required.

Archive Insert Scenario

As we have seen, data and associated metadata can be received from numerous sources. This scenario focuses on a routine data insert from the processing subsystem. It assumes that all components are active and not in any degraded modes of operation, that ESDT data collection types have been established, and that the data server's nominal activity rate is 50% of capacity.

Insert Data into the Archive

1 Initiate the session between the Processing Subsystem and the Data Server.

- The Processing Subsystem sends a Data Insert Request to the Science Data Server.
- Receipt of the request is logged (via MSS Logging Services) and a request identifier is associated with the Data Insert Request.
- The content of the request is validated; if successful, it is queued for later processing. If unsuccessful, a rejection message is issued.
- *The operator may examine the progress of a request by reviewing storage events using the Data Server Subsystem **SDSRV and DDIST GUIs** as described in the next section.*

2 Transfer data from Processing Subsystem to Data Server.

- The queued Data Insert Request is reached and processing begins.
- Associated data granules and metadata are transferred from the Processing Subsystem to the Data Server working storage.
- Data transfer status, including recoverable errors, is indicated in the event log via MSS Logging Services.
- *The operator may check the request status using the Data Server Subsystem **SDSRV & DDIST GUIs**.*

3 Validate metadata received from the Processing Subsystem.

- The metadata update file(s) produced by the associated product PGEs are validated for completeness and correctness.
- Validation success or failure is logged via MSS Logging Services with the associated Data Insert Request Identifier and the appropriate status message is returned to the Processing Subsystem.

4 Store data granules in the permanent archive.

- Upon successful validation of the metadata update file, Science Data Server sends a Data Storage Request to Storage Management.
- The data granules in working storage associated with the Data Storage Request are stored.
- The Archive Activity Log (via MSS Logging Services) records each data product being stored and storage status of each storage operation.
- A checksum value is calculated for each data object associated with each granule.
  - This calculation can be turned off, and if it is, it may result in the archiving of a corrupted granule with no ready means of detecting the corruption.)
- The checksum value (if calculated), storage status, and other selected metadata are forwarded to the Science Data Server in a status message upon completion of the Data Storage Request.

5 Store metadata.

- Science Data Server receives and logs the Data Storage Request status message from Storage Management.
- The additional metadata items are validated.
- The PGE produced metadata update file and the storage management provided metadata are loaded into the metadata database.
- The status of the metadata load is entered in the event log.
- *The operator may examine the progress of a request by reviewing storage events using the Data Server Subsystem* ***SDSRV and DDIST GUIs*** *as described in the next section*.

6 Report Data Insert Request status.

- The Science Data Server logs completion of the Data Insert Request in the event log and reports completion of the Data Insert Request to the Data Archive Manager, the operator console and to the insert Requester (the Processing Subsystem in this scenario).

- Each of the above entities would also be notified if the request failed and reason(s) for failure identified.

7 Process subscriptions based on newly inserted data.

- The Science Data Server will then examine the event list for all subscriptions for that event.
- Subscription notifications are sent to the appropriate entities as appropriate and distribution processing is initiated.
- The Science Data Server sends an Advertisement Update Message to the Advertising Server to advertise the new data.

This page intentionally left blank.

Monitor Archive Requests

As previously noted, data that are inserted into the archive can be received from a number of sources including the Ingest Subsystem, Processing Subsystem, other DAACs, and Authorized Users. The Data Ingest Technician (DIT) or other operator can monitor the insertion of data into the archive using the Data Server Subsystem (DSS) GUIs.

Request Status Window

A primary GUI tool for monitoring of archive processing is the **Request Status** window (see Figure 16, accessible from the **Storage Management Control** GUI. Using the **Request Status** tab the Archive Manager or Distribution Technician can detect stalled requests or servers that appear to be idle.

The **Request Status** window displays the following information:

- **Operation** is the type of operation represented by the request.
- **Request ID** is a unique identifier for the request.
- **Progress** is the stage of processing on which the request is currently working (may include a numeric progress indication).
- **Status** provides information about processes attempted and the outcome (e.g., DsEStDRExecuteFailed, DsEStARPathSearchExhausted, OK, . . .WriteFailed, . . .).
- **Priority** is **Xpress**, **Very High**, **High**, **Normal**, or **Low**.
- **When Submitted** is the time and date when the request was received by the Storage Management server that is responsible for the request.
- **Last Updated** is the time and date when the status was last updated for the request.

The operator can reduce the displayed list of requests by clicking on the **Filtering** pull-down menu just above the **Request Status Information** list on the window. This permits filtering on four areas or filter types selectable from the pull-down menu:

- **Server** controls what activity is displayed by limiting the list to the requests being/having been serviced by a specific server. Selecting **All** displays all requests throughout Storage Management. Other selections include the individual archive servers, cache manager servers, ftp servers, request manager server, and staging disk servers.
- **Operation** allows the operator to focus on a specific type of operation. The list of operations is dynamically generated to reflect those operations for which requests are currently in queue (e.g., **All**, **CMLink**, **ArStore**, **ArRetrieve**, **FtpPull**, **FtpPush**).

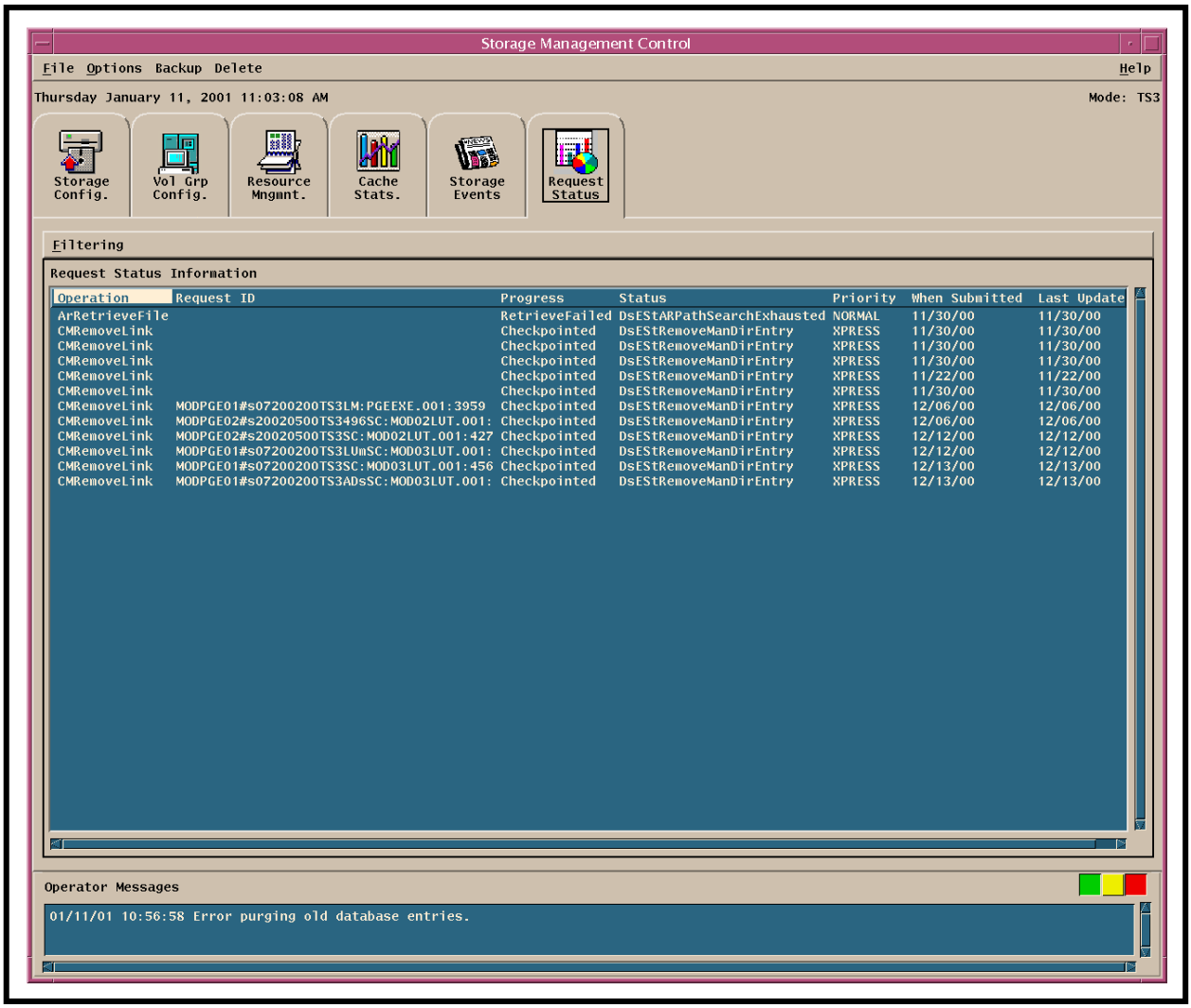


Figure 16. Storage Management Request Status Window

- **Processing State** allows the operator to differentiate among requests that are being actively processed; have been completed, either successfully or to a retryable error state; or have been suspended and are awaiting the outcome of another event. The following selections are available: **All**, **Processing**, **Suspended**, and **Completed**.
- **Submitter** allows the operator to see the status of requests submitted by a specific client process. The list of possible clients is dynamically generated to reflect the list of clients with outstanding requests (e.g., **All**, **DSDD**, **HDFC**, **SDSV**, **this**, **[various servers]**).

Figure 17 illustrates the Request Filtering menus. The following procedure may be used to monitor archive requests using the Storage Management GUI and Request Status window.

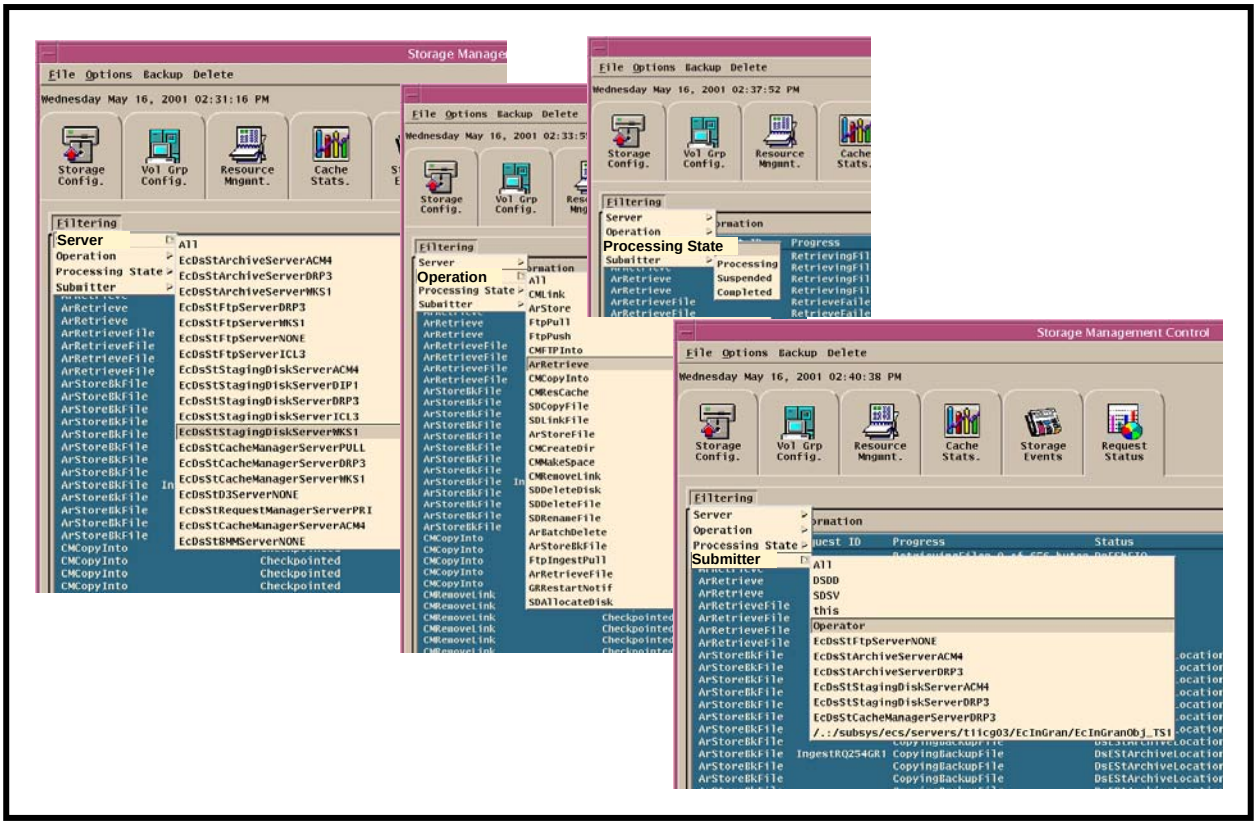


Figure 17. Storage Management Request Filtering Menus

Monitor Archive Requests Using the Storage Management Request Status Window

1 Launch the DSS Storage Management GUI using UNIX commands.

- For detailed instructions refer to the **Launch DSS GUIs Using UNIX Commands** procedure (previous section of this lesson).
- The DSS Storage Management GUI is displayed.

2 Click on the **Request Status** tab.

- The **Request Status** tab is displayed.

3 Observe information displayed on the **Request Status** tab of the **Storage Management Control** GUI.

- The **Request Status Information** table displays the following information:
  - **Operation.**
  - **Request ID.**
  - **Progress.**

- **Status.**
- **Priority.**
- **When Submitted.**
- **Last Updated.**

- By default all storage management server requests for the last 24 hours are shown in the **Request Status Information** table of the **Request Status** tab.
- Clicking on any of the column headers of the **Request Status Information** table causes the listed requests to be sorted in order by the column selected.
  - For example, clicking on the **Last Updated** column header causes requests to be listed in order from the least recently updated to the most recently updated.
- The **Operator Messages** field at the bottom of the GUI displays messages concerning events occurring in storage management operations.
- Note that storage management servers control virtually all data inserted into or retrieved from the archive; the resulting large amount of activity on the **Request Status** tab may make it useful to restrict the number of requests displayed by applying a filter (see next step).

4 To filter the list of requests, use the **Filtering** pull-down menu above the top left corner of the **Request Status Information** table, selecting as desired to display requests associated with a particular **Server**, **Operation**, **Processing State**, or **Submitter**.

- The list of requests displayed in the **Request Status Information** table is restricted by the filtering choice.

5 Observe the Storage Management requests displayed in the **Request Status Information** table.

- The **Progress** and **Status** column entries in the table may provide indication for particular requests of potential problems or conditions requiring attention.

6 Repeat Steps 4 and 5 as necessary to monitor Storage Management requests.

7 To **exit**, follow menu path **File → Exit**.

Distrib'n Requests Window

The Distrib'n Requests window of the Data Distribution Operator GUI, illustrated in Figure 18, displays detailed information on individual data distribution requests and provides the capability to filter requests, change priority of requests, and designate shipping status. The window contains a list of data distribution requests that can be sorted by column. To change the priority

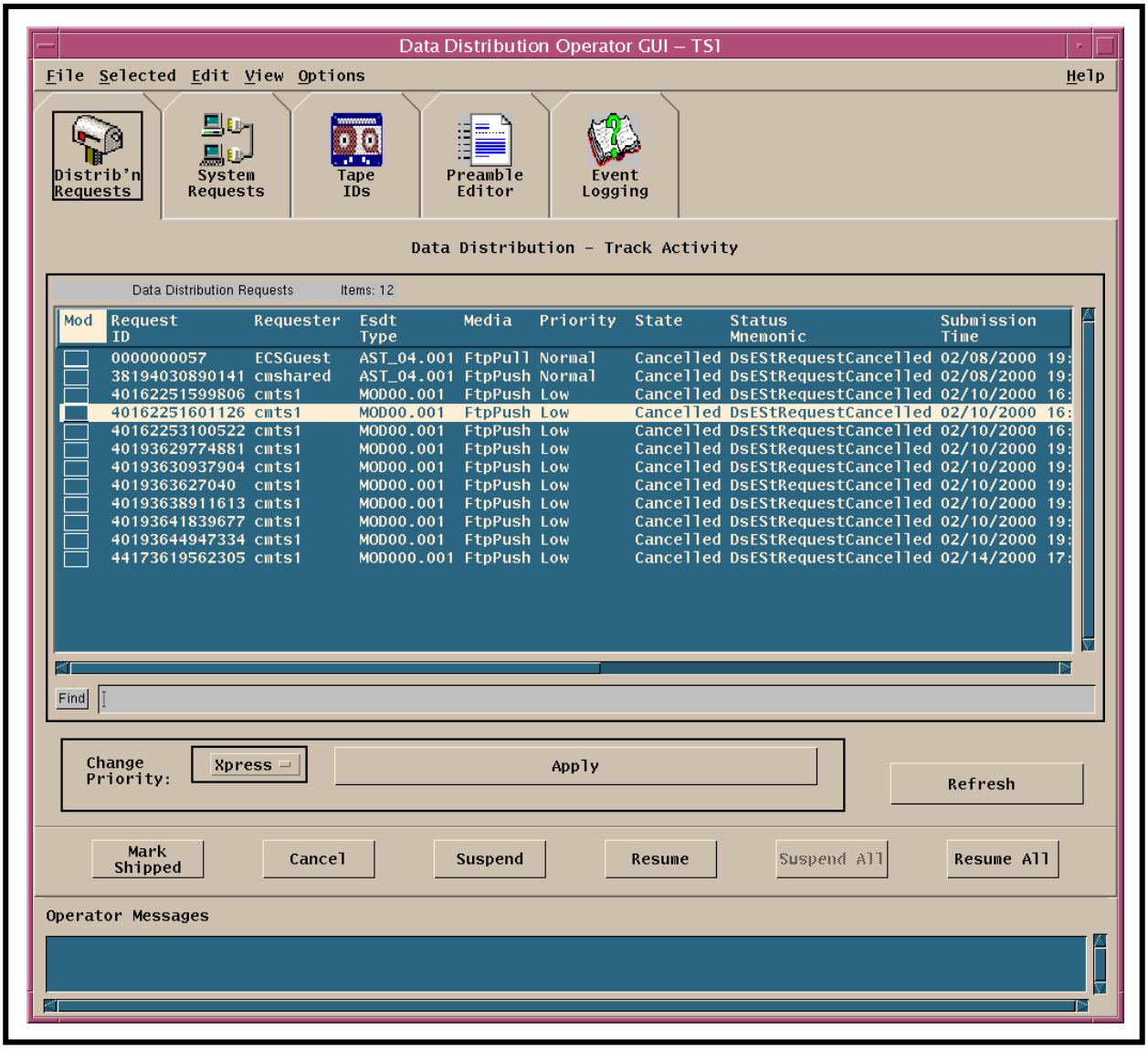


Figure 18. DDIST Distrib'n Requests Window

of a selected request, select the desired priority and click on the **Apply** button in the **Change Priority** area. A selected request can be marked to indicate that it has been shipped. An **Abort** button is used to cancel a selected request. Distribution requests can also be filtered by attributes, using the **Distribution Filter Requests** window shown in Figure 19.The next section addresses retrieval of data from the archive and how to monitor distribution requests using the **Data Distribution GUI** tool and the associated **Filters**.

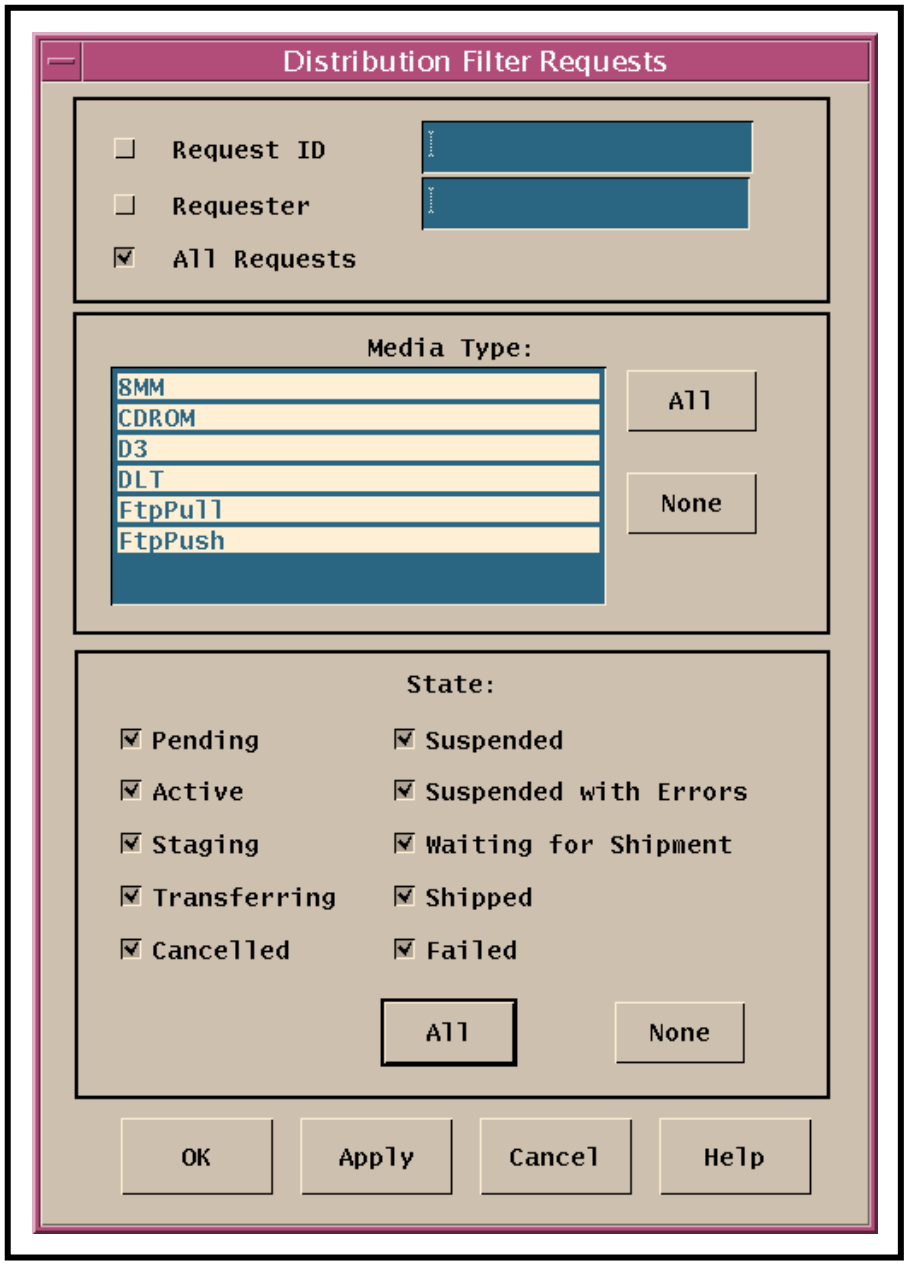


Figure 19. Distribution Filter Requests Window

Retrieve Data from the Archive

Retrieval of data from the archive is a largely automated process that occurs in response to either Data Pool Insert activities or Data Distribution requests.

Data Pool Insert

Data Pool Insert entails inserting ECS or non-ECS data into the Data Pool. ECS data is copied from AMASS into the Data Pool, based on either Data Pool insert actions assigned by the Spatial Subscription Server or Order Manager Server, or on ECS granule IDs submitted by the Batch Insert Utility. Non-ECS data is copied into the Data Pool from a location specified in an XML file created by DAAC operations as input to the Batch Insert Utility.

Data Pool Insert consists of the following six major subcomponents:

- Data Pool Action Driver (EcDlActionDriver) (DPAD) - a C++ executable that schedules Data Pool insert actions based on a queue of Data Pool insert actions populated by the Spatial Subscription Server, the Batch Insert Utility, or the Order Manager Server.
- Data Pool Insert Utility (EcDlInsertUtility) (DPIU) - a java executable that manages the insert of a data granule into the Data Pool, updates the Data Pool database inventory, and writes XML metadata files to the Data Pool disk.
- Data Pool Quick Server (EcDlQuickServer) - a C++ executable that is installed on each AMASS host. It checks AMASS file status and performs actual copy operations from AMASS to the Data Pool for ECS data using the EcUtCopyExec.
- Data Pool XML generator (EcDlM2XTApp) - a java executable that translates ECS granule metadata from the Science Data Server database into XML for storage in the Data Pool directories.
- Band extraction utility (bandtool) - a C executable invoked by the DPIU. The bandtool extracts band information from HDF-EOS granules and stores the extracted information in the Data Pool database. The bandtool is invoked only if the granule being inserted is from a collection eligible for conversion by the HDF-EOS to GeoTiff Conversion Tool (HEG).
- The jpeg extraction utility (hdf2jpeg) - a C executable invoked by the DPIU. The jpeg extraction utility extracts browse images (jpeg or raster) from a browse hdfeos file.

Usually orders for granules (including browse granules as well as science granules) are filled from the Data Pool. The Data Pool Insert Service puts order-only granules into special Data Pool directories that are reserved for this purpose. The order-only directories are hidden from ftp users, and the granules in the directories are not visible to Data Pool web drill down. However,

in support of FTP Pull orders the hidden directories are accessible for anonymous ftp via links in user FTP Pull directories.

When an order-only granule is later inserted via a “normal” Data Pool insert, it becomes a normal Data Pool granule, and the Data Pool Insert Service moves it from the hidden directory to the appropriate place in the public directory structure. Such an event should never occur for billable or restricted granules. DAACs should not and generally do not insert granules from billable collections into the Data Pool, and the Data Pool Insert Service will perform various checks (including examination of the DFA flag) to prevent the insert of granules that are flagged as “Hidden” or “Restricted” in ECS.

To decouple Data Pool and OMS file references, ordered granules are always be represented in the hidden directory structure either by files (if the granule is not in the public Data Pool) or by links pointing to the public files (if the granule is in the public Data Pool). If a granule is converted from an order-only to a public Data Pool granule, its files are moved and links are left behind in the order-only directories. A “metadata only” granule is considered public; however, the corresponding science file remains inaccessible; i.e., it is never put in the public Data Pool.

The same general rules apply to order-only browse granules as apply to order-only science granules, i.e., they are placed into the hidden directory structure.

- For example:
  - An EOS Data Gateway (EDG) user requests a browse product for immediate display on the EDG GUI (a feature called an Integrated Browse Request).
  - EDG sends the Integrated Browse Request to the V0 Gateway.
  - V0 Gateway submits the request for a browse granule to the Data Pool (DPL) CI.
  - DPL checks whether the browse granule is already in the Data Pool.
  - If the browse granule is not already in the Data Pool, DPL (i.e., Data Pool Insert) stages the browse file (in the DPL hidden directory structure) by retrieving it from the appropriate AMASS storage location, as specified in the volume group history information in the STMGT database.
  - DPL reads the staged browse file and returns it to the V0 Gateway.
  - V0 Gateway returns the browse file to the EDG.

Monitoring Distribution Requests

Although it is becoming more common to distribute data from the Data Pool rather than from the Archive via Data Distribution, there are still a number of possible sources for Data Distribution requests:

- data orders [especially secure distribution (scp) requests] from scientists or other end users.
  - one-time orders.
  - standing orders placed as subscriptions for acquiring data.
- data requests from other DAAC sites.
  - cross-DAAC orders for end users.
  - data needed as input for processing at other sites (subscriptions placed for ingest by those sites).
- internal requests for data needed for processing.

Placing orders and subscriptions on behalf of scientists or other end users is typically done by User Services representatives. Procedures for these activities are addressed in the training materials for User Services. As the system responds to these requests, and to requests from other DAACs or internal processes, the Archive Manager or other operators can monitor the progress of the distribution requests. Suppose that a User Services representative at your DAAC asks you to check on the status of a data distribution request from a user named Ivan Ohrdurr. The following procedure is applicable.

Monitor Distribution Requests Using the Data Distribution GUI

---

1 Launch the **DDIST GUI** using UNIX commands.

- For detailed instructions refer to the **Launch DSS GUIs Using UNIX Commands** procedure (previous section of this lesson).
- The **Data Distribution GUI** tool is displayed.

2 Click on the **Distrib'n Requests** tab.

- The Distribution Requests window is opened.
- A list of requests is displayed.

3 **Click** on the **Filter** push button.

- The **Distribution Filter Requests** window opens.
- Three filter types are displayed: **Request ID**, **Requester**, and **All**.

4 **Click** on the **Requester** button, in the radio box.

- The cursor moves to the selection field to the right of the **Requester**.

5 Enter the requester's name (in this case, **Ohrdurr**) in the text entry field opposite the **Requester** button and label.

6 Click on the **All** button in the **Media Type:** area.

- All of the Media Type toggle buttons show as selected (depressed).
- If you are seeking only requests for a particular set of media, you can select just the button(s) for that set instead of clicking on the **All** button.

7 Click on the **All** button in the **State:** area.

- All of the State toggle buttons show as selected (depressed).
- If you are seeking only requests in a particular state or states, you can select just the button(s) for the desired state(s) instead of clicking on the **All** button.

8 **Click** on the **OK** push button, located at the bottom of the window.

- The other push buttons located at the bottom of the window are **Apply**, **Cancel**, and **Help**.
- The **Filter Requests** window is closed.
- The Distribution Requests screen shows any requests that meet the filter criteria in the **Data Distribution Requests** field.

9 If necessary, use the scroll bar at the bottom of the **Data Distribution Requests** field to scroll horizontally to view the state of the Ohrdurr request(s).

Deleting Granules

The granule deletion capability provides the Operations Staff with the ability to delete granules using command line interfaces and the **Storage Management Control** GUI. The granules can be deleted from both the inventory and archive or just the archive.

The deletion process can involve deleting the specified SC (science) granules along with associated granules BR (browse), PH (production history), and QA (quality assurance) granules, as long as any other granules do not reference the associated granules. The deletion process can also involve deleting the specified granules even if they are inputs to other granules.

Phases of Granule Deletion

There are three phases to the granule deletion process:

- Logical deletion [marking granules (in the SDSRV database) for deletion].
- Physical deletion (removing from the SDSRV database data concerning granules marked for deletion and putting in the STMGT database data concerning granules to be deleted from the archive).
- Deletion from the archive (removing from the archive the files identified for deletion in the STMGT database).

Phase 1, Logical Deletion: For the first phase, a command-line Bulk Delete utility (EcDsBulkDelete.pl) responds to operator-specified criteria for the deletion of granules by "logically" deleting from the inventory (SDSRV database) those granules that meet the criteria. The granules are marked as “deleted” and can no longer be accessed, but their inventory entries are not removed yet. The logical “deletion” may specify the flagging of granule files to be deleted from the archive (*Delete From Archive*, or DFA) only, leaving the inventory record intact, or it may specify *Physical Deletion*, which entails removal of the inventory record from the database as well as removal of the files from the archive. For each granule to be physically deleted an entry is made in the DsMdDeletedGranules table of the SDSRV database with a time stamp recording the logical deletion time. If applicable, the DFAFlag is set for the granule’s entry in the DsMdDeletedGranules table. Flagging DFA granules involves changing the value of the DeleteFromArchive entry in the DsMdGranules table from **N** to **Y**.

Phase 2, Physical Deletion: The second phase is actual deletion from the inventory (SDSRV database) of the granules marked for physical deletion (not DFA only). Physical deletion occurs when the operations staff runs the Science Data Server Deletion Cleanup utility (EcDsDeletionCleanup.pl). The Deletion Cleanup utility removes all inventory rows (in the SDSRV database) for granules that were flagged as “deleted,” including rows referencing related information (e.g., BR, PH, and QA), and the script creates entries in the DsStPendingDelete table (in the STMGT database) for granules to be deleted. This includes entries for granules that are to be physically deleted, as well as those designated DFA only.

Phase 3, Deletion from the Archive: In Phase 3 the operator uses the Archive Batch Delete utility to initiate the removal from the archive of the files listed in the DsStPendingDelete table (populated by SDSRV in Phase 2). The operator must run the utility separately on each appropriate Archive host (FSMS Server or APC Server host) because the utility lists and deletes files from archive on the local host only. The Archive Batch Delete utility creates requests to the local archive server host to delete selected files. Files that are successfully deleted have their associated rows removed from the DsStPendingDelete table.

Periodically, as sufficient data removal from the archive makes it appropriate, Operations may elect to reclaim the tape space and recycle archive tapes. AMASS software commands (e.g., *volcomp*, *volclean*, *volformat*, and *volstat*) are used for that purpose.

Phase 1, Logical Deletion

Logical Deletion involves the use of the Bulk Delete utility (EcDsBulkDelete.pl). A prerequisite to Logical Deletion is having a file of GeoIDs (corresponding to granules) for use as input to the Bulk Delete utility. Although it is possible to manually create a file of GeoIDs using a UNIX editor, an easier way is to use the Bulk Search utility to generate a list of GeoIDs based on criteria to be specified when running the Bulk Search utility.

Generating a GeoID File to Be Used as Input to the Bulk Delete Utility or the Bulk Undelete Utility

A GeoID is the granule identification portion of a Universal Reference (UR); it contains the BaseType, SubType (ESDT ShortName and VersionID) and databaseID. For example, the GeoID SC:AST_L1BT.001:5850 indicates BaseType SC (science granule), ShortName AST_L1BT (ASTER Level 1B thermal infrared data) VersionID 001, and databaseID 5850. The GeoID is different for each granule in the system.

GeoID files are input files for scripts used in deleting (or “undeleting”) ECS granules from the inventory, archive, or Data Pool. A GeoID file consists of a list of GeoIDs for granules that are to be deleted (or “undeleted”). One way to generate a file of granule GeoIDs is to use the Bulk Search utility, which allows the operator to specify criteria for selecting granules on the basis of certain characteristics (e.g., ESDT, version, and date inserted in the archive). Subsequently, the GeoID file can be used as input to the Bulk Delete utility, the Bulk Undelete utility, or the Data Pool Cleanup Utility.

Generating a GeoID File to Be Used as Input to the Bulk Delete Utility or the Bulk Undelete Utility

---

NOTE: Finding 10,000 granules may take 15 to 20 minutes.

1 Log in at the Sun Consolidation Internal Server host, where the Bulk Search utility is installed.

- Examples of Sun Consolidation Internal Server host names include **e0acs11**, **g0acs11**, **l0acs03**, and **n0acs04**.

- For detailed instructions refer to the procedure for **Logging in to System Hosts** (previous section of this lesson).

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

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

3 To set up relevant environment variables (if desired) at the UNIX prompt enter:

setenv MODE <*ECS mode*>
setenv SYB_SQL_SERVER <*Sybase server*>
setenv SYB_DB_NAME <*database name*>

- For example, the following commands would allow running the Bulk Search utility using the OPS mode Science Data Server database at the ASDC (LaRC):

  setenv MODE OPS
  setenv SYB_SQL_SERVER l0acg02_srvr
  setenv SYB_DB_NAME EcDsScienceDataServer1

- The <***ECS mode***> value specified for the **MODE** parameter indicates the ECS mode (e.g., OPS, TS1, or TS2) to be searched.
  - If this environment variable is set, the **-mode** command line argument does not need to be given when starting the Bulk Search utility.
- The <***Sybase server***> value specified for the **SYB_SQL_SERVER** parameter indicates the Sybase (database) server (e.g., e0acg11_srvr, g0acg01_srvr, l0acg02_srvr, or n0acg01_srvr) for the Science Data Server database.
  - If this environment variable is set, the **-server** command line argument does not need to be given when starting the Bulk Search utility.
- The <***database name***> value specified for the **SYB_DB_NAME** parameter indicates which database (e.g., EcDsScienceDataServer1, EcDsScienceDataServer1_TS1, or EcDsScienceDataServer1_TS2) is involved in the search.
  - If this environment variable is set, the **-database** command line argument does not need to be given when starting the Bulk Search utility.

4 At the UNIX prompt enter:

EcDsBulkSearch.pl [-name <*shortname*>] [-version <*version ID*>] [-begindate <*mm/dd/yyyy*> [*hh:mm:ss*]] [-enddate <*mm/dd/yyyy*> [*hh:mm:ss*]] [-insertbegin <*mm/dd/yyyy*> [*hh:mm:ss*]] [-insertend <*mm/dd/yyyy*> [*hh:mm:ss*]] [-acquirebegin <*mm/dd/yyyy*> [*hh:mm:ss*]] [-acquireend <*mm/dd/yyyy*> [*hh:mm:ss*]] [-DFA] [-physical] [-localgranulefile <*local granule ID file*>] -geoidfile <*geoid file*> [-limit <*granule limit*>] [-mode <*ECS mode*>] [-server <*Sybase server*>] [-database <*database name*>] -user <*database login ID*> [-password <*database login password*>]

- For example, the following command would generate a file of GeoIDs for all PM1ATTNR.077 granules marked "DFA" in the OPS mode at the ASDC (LaRC):

  EcDsBulkSearch.pl -DFA -name PM1ATTNR -version 077 -mode OPS -server l0acg02_srvr -database EcDsScienceDataServer1 -user sdsrv_role -password *password* -geoidfile PM1ATTNR_Dec23.geoid

- If the **MODE**, **SYB_SQL_SERVER**, and **SYB_DB_NAME** environment variables had been set as shown in the example in Step 3, the following command would have the same effect as the previous example:

  EcDsBulkSearch.pl -DFA -name PM1ATTNR -version 077 -user sdsrv_role -password *password* -geoidfile PM1ATTNR_Dec23.geoid

- The following command would generate a file of GeoIDs for all deleted ("physical" deletion) PM1ATTNR.077 granules in the OPS mode at the ASDC (LaRC):

  EcDsBulkSearch.pl -physical -name PM1ATTNR -version 077 -user sdsrv_role -password *password* -geoidfile PM1ATTNR_Dec23.geoid

- The following command would generate a file of GeoIDs for all granules inserted in the OPS mode at the ASDC (LaRC) between 11:00 P.M. on May 6, 2005 and 6:00 P.M. on May 7, 2005:

  EcDsBulkSearch.pl -insertbegin 05/06/2005 23:00:00 -insertend 05/07/2005 18:00:00 -mode OPS -server l0acg02_srvr -database EcDsScienceDataServer1 -user sdsrv_role -password *password* -geoidfile 05062005GranList.geoid

- The **-name** option indicates that the search is restricted to granules with the specified ESDT ShortName only. Typically the **-name** option is used in conjunction with the **-version** option.

- The **-version** option indicates that the search is restricted to granules with the specified ESDT Version only. The **-version** option is used in conjunction with the **-name** option.

- The **-begindate** option indicates that the search is restricted to granules with a BeginningDateTime greater than or equal to the specified date (and optionally time). Typically the **-begindate** option is used in conjunction with the **-enddate** option.

- The **-enddate** option indicates that the search is restricted to granules with an EndingDateTime less than or equal to the specified date (and optionally time). Typically the **-enddate** option is used in conjunction with the **-begindate** option.
- The **-insertbegin** option indicates that the search is restricted to granules with an insertTime greater than or equal to the specified date (and optionally time). The **-insertbegin** option may be used in conjunction with the **-insertend** option.
- The **-insertend** option indicates that the search is restricted to granules with an insertTime less than or equal to the specified date (and optionally time). The **-insertend** option may be used in conjunction with the **-insertbegin** option.
- The **-acquirebegin** option indicates that the search is restricted to granules with a BeginningDateTime greater than or equal to the specified date (and optionally time). The **-acquirebegin** option is the same as the **–begindate** option except that it can be combined with the **-acquireend** option and used in a **BETWEEN** clause.
- The **-acquireend** option indicates that the search is restricted to granules with a BeginningDateTime less than or equal to the specified date (and optionally time). Typically the **-acquireend** option is used in conjunction with the **–acquirebegin** option.
- The **-physical** option indicates that the search is restricted to deleted granules.
- The **-DFA** option indicates that the search is restricted to granules that are in a DFA (deleted from archive) status.
- The <***local granule ID file***> value specified for the **-localgranulefile** optional parameter identifies an input file containing Local Granule IDs that are to be converted into GeoIDs.
- The <***geoid file***> value specified for the **-geoidfile** parameter identifies the output file that contains the GeoIDs (e.g., **SC:PM1ATTNR.077:2013463060**) of the granules that meet the search criteria. A report file with the same name as the GeoID file but with a **.rpt** extension is generated for the purpose of verification.
- The **-limit** option indicates that the search should be limited to (return no more than) the specified number of granules.
- The <***ECS mode***> value specified for the **-mode** parameter indicates the ECS mode (e.g., OPS, TS1, or TS2) to be searched.
- The <***Sybase server***> value specified for the **-server** parameter indicates the Sybase (database) server (e.g., e0acg11_srvr, g0acg01_srvr, l0acg02_srvr, or n0acg01_srvr) for the Science Data Server database.
- The <***database name***> value specified for the **-database** parameter indicates which database (e.g., EcDsScienceDataServer1, EcDsScienceDataServer1_TS1, or EcDsScienceDataServer1_TS2) is involved in the search.
- The <***database login ID***> value specified for the **-user** parameter is the user ID (e.g., sdsrv_role) for logging in to the Science Data Server database.

- The ***<database login password>*** value specified for the **- password** parameter is the password for logging in to the Science Data Server database. If the password option is not entered on the command line, the script prompts for a password.
- The Bulk Search utility runs and records GeoIDs of granules that meet the specified criteria in the GeoID file.

5 When the Bulk Search utility has completed its run and the GeoID output file is available, type **vi** ***<geoid file>*** and then press the **Return/Enter** key.

- ***<geoid file>*** refers to the GeoID file to be reviewed (e.g., SYN5_GranDel.geoid, Dec_23_AllDatatypes.geoid).
- For example:

```
l0acs03{cmops}[10]->vi PM1ATTNR_Dec23.geoid
SC:PM1ATTNR.077:2013463060
SC:PM1ATTNR.077:2013463061
SC:PM1ATTNR.077:2013463062
SC:PM1ATTNR.077:2013463063
SC:PM1ATTNR.077:2013463064
SC:PM1ATTNR.077:2013463065
SC:PM1ATTNR.077:2013463066
SC:PM1ATTNR.077:2013463067
SC:PM1ATTNR.077:2013463068
SC:PM1ATTNR.077:2013463069
[…]
SC:PM1ATTNR.077:2013496391
SC:PM1ATTNR.077:2013496392
SC:PM1ATTNR.077:2013496393
"PM1ATTNR_Dec23.geoid" 9740 lines, 255893 characters
```

  - Many lines have been deleted from the example.
- Although this procedure has been written for the **vi** editor, any UNIX editor can be used to edit the file.

6 Review the file entries to identify problems that have occurred.

- The GeoID file must contain GeoIDs in the format ***<BaseType>:<ESDT_ShortName.VersionID>:<databaseID>***.
- For example:

  SC:PM1ATTNR.077:2013496393

  - The GeoID in the example indicates BaseType SC (science granule), ShortName PM1ATTNR (AQUA attitude data in native format) VersionID 077, and databaseID 2013496393.
- There may be no spaces or blank lines in the file.

7 Use UNIX editor commands to fix problems detected in the file.

- The following vi editor commands are useful:
  - **h** (move cursor left).
  - **j** (move cursor down).
  - **k** (move cursor up).
  - **l** (move cursor right).
  - **a** (append text).
  - **i** (insert text).
  - **r** (replace single character).
  - **x** (delete a character).
  - **dw** (delete a word).
  - **dd** (delete a line).
  - ***n*dd** (delete *n* lines).
  - **u** (undo previous change).
  - **Esc** (switch to command mode).

8 Press the **Esc** key.

9 At the **vi** prompt type **ZZ** and then press the **Return/Enter** key.

- **vi** exits and the edited file is saved.
  - To exit **vi** without saving the new entries in the file type **:q!** then press **Return/Enter**.
- UNIX command line prompt is displayed.

Deleting Granules, Phase 1: Running the Bulk Delete Utility

Once granules have been identified/selected for deletion (e.g., using the Bulk Search utility), the operator runs the Bulk Delete utility, a Perl script, **EcDsBulkDelete.pl**. There are two types of runs that can be performed with the Bulk Delete utility:

- Physical.
- DFA.

A “physical” deletion run results in marking each specified granule and metadata as deleted from both inventory and archive. A “DFA” run involves marking each specified granule and metadata as deleted from the archive only.

As previously mentioned the Bulk Delete utility responds to operator-specified criteria for the deletion of granules by "logically" deleting from the inventory (SDSRV database) those granules that meet the criteria. The granules are marked as “deleted” and can no longer be accessed, but their inventory entries are not removed yet. The logical “deletion” may specify the flagging of granule files to be deleted from the archive (*Delete From Archive*, or DFA) only, leaving the inventory record intact, or it may specify *Physical Deletion*, which entails removal of the inventory record from the database as well as removal of the files from the archive.

For each granule to be physically deleted an entry is made in the DsMdDeletedGranules table of the SDSRV database with a time stamp recording the logical deletion time. If applicable, the DFAFlag is set for the granule's entry in the DsMdDeletedGranules table. Flagging DFA granules involves changing the value of the DeleteFromArchive entry in the DsMdGranules table from **N** to **Y**.

Deleting Granules, Phase 1: Running the Bulk Delete Utility

NOTE: A prerequisite to deleting granules is having a file of GeoIDs (corresponding to granules) for use as input to the Bulk Delete utility. Although it is possible to manually create a file of GeoIDs using a UNIX editor, an easier way is to use the Bulk Search utility to generate a list of GeoIDs based on criteria specified when running the Bulk Search utility [refer to the procedure for **Generating a GeoID File to Be Used as Input to the Bulk Delete Utility or the Bulk Undelete Utility** (previous section of this lesson).]

NOTE: Deleting 10,000 granules may take 15 to 20 minutes.

1 Log in at the Sun Consolidation Internal Server host, where the Bulk Delete utility is installed.

- Examples of Sun Consolidation Internal Server host names include **e0acs11**, **g0acs11**, **l0acs03**, and **n0acs04**.
- For detailed instructions refer to the procedure for **Logging in to System Hosts** (previous section of this lesson).

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

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

3 To set up relevant environment variables (if desired) at the UNIX prompt enter:

setenv MODE <*ECS mode*>
setenv SYB_SQL_SERVER <*Sybase server*>
setenv SYB_DB_NAME <*database name*>

- For example, the following commands would allow running the Bulk Delete utility using the OPS mode Science Data Server database at the ASDC (LaRC):

  setenv MODE OPS
  setenv SYB_SQL_SERVER l0acg02_srvr
  setenv SYB_DB_NAME EcDsScienceDataServer1

- The <***ECS mode***> value specified for the **MODE** parameter indicates the ECS mode (e.g., OPS, TS1, or TS2) in which to run the Bulk Delete utility.
  - If this environment variable is set, the **-mode** command line argument does not need to be given when starting the Bulk Delete utility.

- The <***Sybase server***> value specified for the **SYB_SQL_SERVER** parameter indicates the Sybase (database) server (e.g., e0acg11_srvr, g0acg01_srvr, l0acg02_srvr, or n0acg01_srvr) for the Science Data Server database.
  - If this environment variable is set, the **-server** command line argument does not need to be given when starting the Bulk Delete utility.
- The <***database name***> value specified for the **SYB_DB_NAME** parameter indicates which database (e.g., EcDsScienceDataServer1, EcDsScienceDataServer1_TS1, or EcDsScienceDataServer1_TS2) is to be accessed by the Bulk Delete utility.
  - If this environment variable is set, the **-database** command line argument does not need to be given when starting the Bulk Delete utility.

4 To perform "physical" deletion (i.e., deletion from both inventory and archive), at the UNIX prompt enter:

EcDsBulkDelete.pl -physical [-delref] [-noassoc] -geoidfile <*geoid file*> [-mode <*ECS mode*>] [-server <*Sybase server*>] [-database <*database name*>] -user <*database login ID*> [-password <*database login password*>] [-log <*log filename*>]

- For example:

  EcDsBulkDelete.pl -physical -geoidfile PM1ATTNR_Dec23.geoid -mode OPS -database EcDsScienceDataServer1 -user sdsrv_role -server l0acg02_srvr -password *password* -log PM1ATTNR_Dec23DelPhys.log

- The **-physical** option (i.e., delete from both inventory and archive) indicates that both granule metadata and archive are to be marked for deletion.
- The **-delref** option (i.e., delete granules that are referenced by other granules) indicates that a non-science/limited (non-SC/LM) granule should be deleted even if it is associated with "undeleted" SC/LM granules.

NOTE: The **-delref** option has no effect on deleting SC/LM granules. They are always deleted regardless of whether or not they are referenced.

- The **-noassoc** option indicates that associated granules (e.g., QA, Browse, or PH granules) are not to be deleted.
- <***geoid file***> is the file that contains the GeoIDs of the granules to be deleted (in the form **SC:AST_L1BT.001:4267**).
- <***ECS mode***> is the ECS mode (e.g., OPS, TS1, or TS2) containing the granules to be deleted.
- <***Sybase server***> is the name of the applicable Sybase (database) server (e.g., e0acg11_srvr, g0acg01_srvr, l0acg02_srvr, or n0acg01_srvr) for the Science Data Server database.
- <***database name***> is the name of the Science Data Server database (e.g., EcDsScienceDataServer1, EcDsScienceDataServer1_TS1, or EcDsScienceDataServer1_TS2) for the applicable mode.

- ***<database login ID>*** is the user ID (e.g., sdsrv_role) for logging in to the Science Data Server database.
- ***<database login password>*** is the database login ID's password for logging in to the Science Data Server database. If the password option is not entered on the command line, the script prompts for a password.
- ***<log filename>*** is the name of the log file to which a deletion report is to be written. If no log file name is provided, the Bulk Delete utility creates a log file (BulkDelete.<*date*>.log) in the /usr/ecs/<*MODE*>/CUSTOM/logs/ directory.
- The Bulk Delete utility runs and deletes the granules specified in the GeoID file from the inventory and the archive.
- The Bulk Delete utility records information about utility events in the log file (e.g., PM1ATTNR_Dec23DelPhys.log).

5 To perform "DFA" deletion only, at the UNIX prompt enter:

EcDsBulkDelete.pl -DFA -geoidfile <*geoid file*> [-mode <*ECS mode*>] [-server <*Sybase server*>] [-database <*database name*>] -user <*database login ID*> [-password <*database login password*>] [-log <*log filename*>]

- For example:

  EcDsBulkDelete.pl -DFA -geoidfile PM1ATTNR_Dec23.geoid -mode OPS -database EcDsScienceDataServer1 -user sdsrv_role -server l0acg02_srvr -password *password* -log PM1ATTNR_Dec23DelDFA.log

- The **-DFA** option indicates that the granules listed in the GeoID file are to be marked as "Delete From Archive" only (does not involve a "physical" deletion).
- The Bulk Delete utility runs and deletes the granules specified in the GeoID file from the archive.
- The Bulk Delete utility records information about utility events in the log file (e.g., PM1ATTNR_Dec23DelDFA.log).

6 When the Bulk Delete utility has completed its run and the log file is available, at the UNIX prompt enter:

pg <log filename>

- ***<log filename>*** refers to the log file to be reviewed (e.g., PM1ATTNR_Dec23DelPhys.log, PM1ATTNR_Dec23DelDFA.log).
- The first page of the log file is displayed.
- For example:

```
l0acs03{cmops}[10]->pg PM1ATTNR_Dec23DelPhys.log
*****************************************************
*            Bulk Granule Deletion Tool
*
*
```

```
*  Report created 23 Dec 2005 10:46:22
*
*  GEOID file:    /usr/ecs/OPS/CUSTOM/utilities/PM1ATTNR_Dec23.geoid
*  Deletion mode: OPS
*  Deletion type: physical
*  Deleting associated granules: yes
*
*  Total granules processed: 9740
*****************************************************
*
* TOTAL UPDATED GRANULES
*
* SC Type = 9705
*
* Associated granules deleted = 0
*
*****************************************************
*
* TOTAL FAILED GRANULES
*
* SC Type = 35
*
*****************************************************

SC:PM1ATTNR.077:2013463060 -- Granule successfully marked
SC:PM1ATTNR.077:2013463061 -- Granule successfully marked
SC:PM1ATTNR.077:2013463062 -- Granule successfully marked
SC:PM1ATTNR.077:2013463063 -- Granule successfully marked
SC:PM1ATTNR.077:2013463064 -- Granule successfully marked
SC:PM1ATTNR.077:2013463065 -- No matching granule found
SC:PM1ATTNR.077:2013463066 -- No matching granule found
SC:PM1ATTNR.077:2013463067 -- No matching granule found
SC:PM1ATTNR.077:2013463068 -- No matching granule found
SC:PM1ATTNR.077:2013463069 -- No matching granule found
[…]
SC:PM1ATTNR.077:2013496391 -- Granule successfully marked
SC:PM1ATTNR.077:2013496392 -- Granule successfully marked
SC:PM1ATTNR.077:2013496393 -- Granule successfully marked

(EOF):
```

  – Many lines have been deleted from the example.

- Although this procedure has been written for the **pg** command, any UNIX editor or visualizing command (e.g., **vi**, **view, more, page**) can be used to review the file.

7 Review the log file to determine whether all deletions occurred as expected or experienced errors.

- The following **pg** commands (at the **:** prompt) are useful:
  - **n** then **Return/Enter** (go to Page n).
  - **Return/Enter** or **+1** then **Return/Enter** (go down to the next page).
  - **-1** then **Return/Enter** (go back to the preceding page).
  - **+n** then **Return/Enter** (go down n number of pages).
  - **-n** then **Return/Enter** (go back n number of pages).
  - **+nl** then **Return/Enter** (go down n number of lines).
  - **-nl** then **Return/Enter** (go back n number of lines).
  - **$** then **Return/Enter** [go to the last page (end of file)].
  - **q** then **Return/Enter** (exit from pg).
- The following types of messages related to granule states may be seen in the log file (as shown in the example in Step 6):
  - Granule successfully marked.
  - No matching granule found.
  - Cannot DFA this granule type.
  - Granule was logically deleted already.
  - Golden granule cannot be deleted.
  - Unknown error deleting granule.
  - Granule NOT Deleted/DFAed Because Either It Is an Input To Another Granule OR It Is Associated With A Granule Which Is an Input To Another Granule.
  - Granule was NOT locatable and therefore was NOT Deleted/DFAed.
  - Granule was DFAed already.

8 If the GeoID and log files were written to the utilities directory (e.g., /usr/ecs/OPS/CUSTOM/utilities) rather than a home directory, log directory, or data directory, move or remove the files from the utilities directory.

- For example:

  mv PM1ATTNR_Dec23DelPhys.log ../logs
  mv PM1ATTNR_Dec23.geoid /home/cmops/geoid

- Leaving GeoID files and log files in the utilities directory can eventually make it more difficult to find needed scripts in the directory.
- The utilities directory is intended for scripts, not log files or data files. Leaving GeoID files and log files in the utilities directory causes it to load up with extraneous files.

"Undeleting" Granules from the Archive and Inventory

In the event that it is desirable to restore granules that have been marked for deletion in the Science Data Server database (e.g., using the Bulk Delete utility), the Bulk Undelete utility provides an "undelete" capability.

"Undeleting" Granules from the Archive and Inventory

NOTE: A prerequisite to "undeleting" is having a file of GeoIDs (corresponding to granules) for use as input to the Bulk Undelete utility. Although it is possible to manually create a file of GeoIDs using a UNIX editor, an easier way is to use the Bulk Search utility to generate a list of "deleted" GeoIDs based on criteria specified when running the Bulk Search utility [refer to the procedure for **Generating a GeoID File to Be Used as Input to the Bulk Delete Utility or the Bulk Undelete Utility** (previous section of this lesson).]

NOTE: "Undeleting" 10,000 granules may take 15 to 20 minutes.

1 Log in at the Sun Consolidation Internal Server host, where the Bulk Undelete utility is installed.

- Examples of Sun Consolidation Internal Server host names include **e0acs11**, **g0acs11**, **l0acs03**, and **n0acs04**.
- For detailed instructions refer to the procedure for **Logging in to System Hosts** (previous section of this lesson).

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

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

3 To set up relevant environment variables (if desired) at the UNIX prompt enter:

setenv MODE <*ECS mode*>
setenv SYB_SQL_SERVER <*Sybase server*>
setenv SYB_DB_NAME <*database name*>

- For example, the following commands would allow running the Bulk Undelete utility using the OPS mode Science Data Server database at the ASDC (LaRC):

  setenv MODE OPS
  setenv SYB_SQL_SERVER l0acg02_srvr
  setenv SYB_DB_NAME EcDsScienceDataServer1

- The <***ECS mode***> value specified for the **MODE** parameter indicates the ECS mode (e.g., OPS, TS1, or TS2) in which to run the Bulk Undelete utility.
  - If this environment variable is set, the **-mode** command line argument does not need to be given when starting the Bulk Undelete utility.

- The <***Sybase server***> value specified for the **SYB_SQL_SERVER** parameter indicates the Sybase (database) server (e.g., e0acg11_srvr, g0acg01_srvr, l0acg02_srvr, or n0acg01_srvr) for the Science Data Server database.
  - If this environment variable is set, the **-server** command line argument does not need to be given when starting the Bulk Undelete utility.
- The <***database name***> value specified for the **SYB_DB_NAME** parameter indicates which database (e.g., EcDsScienceDataServer1, EcDsScienceDataServer1_TS1, or EcDsScienceDataServer1_TS2) is to be accessed by the Bulk Undelete utility.
  - If this environment variable is set, the **-database** command line argument does not need to be given when starting the Bulk Undelete utility.

4 To perform "physical undeletion," at the UNIX prompt enter:

EcDsBulkUndelete.pl -physical [-noassoc] -geoidfile <*geoid file*> [-mode <*ECS mode*>] [-server <*Sybase server*>] [-database <*database name*>] -user <*database login ID*> [-password <*database login password*>] [-log <*log filename*>]

- For example:

  EcDsBulkUndelete.pl -physical -geoidfile PM1ATTNR_Dec23.geoid -mode OPS -database EcDsScienceDataServer1 -user sdsrv_role -server l0acg02_srvr -password *password* -log PM1ATTNR_Dec23UndelPhys.log
- The **-physical** option (i.e., "undelete" from both inventory and archive) indicates that both granule metadata and archive are to be marked for "undeletion."
- The **-noassoc** option indicates that associated granules (e.g., QA, Browse, or PH granules) are not to be "undeleted."
- **<*geoid file*>** The file that contains the GeoIDs of the granules to be "undeleted" (in the form **SC:AST_L1BT.001:4267**).
- **<*ECS mode*>** is the ECS mode (e.g., OPS, TS1, or TS2) containing the granules to be "undeleted."
- **<*Sybase server*>** is the name of the applicable Sybase (database) server (e.g., e0acg11_srvr, g0acg01_srvr, l0acg02_srvr, or n0acg01_srvr) for the Science Data Server database.
- **<*database name*>** is the name of the Science Data Server database (e.g., EcDsScienceDataServer1, EcDsScienceDataServer1_TS1, or EcDsScienceDataServer1_TS2) for the applicable mode.
- **<*database login ID*>** is the user ID (e.g., sdsrv_role) for logging in to the Science Data Server database.
- **<*database login password*>** is the database login ID's password for logging in to the Science Data Server database. If the password option is not entered on the command line, the utility prompts for a password.

- ***<log filename>*** is the name of the log file to which a deletion report is to be written. If no log file name is provided, the Bulk Undelete utility creates a log file (BulkUndelete.<*date*>.log) in the /usr/ecs/<*MODE*>/CUSTOM/logs/ directory.
- The Bulk Undelete utility runs and removes "deleted" markings for the granules/metadata specified in the GeoID file in both inventory and archive.
- The Bulk Undelete utility records information about utility events in the log file (e.g., PM1ATTNR_Dec23UndelPhys.log).

5 To perform a "DFA undeletion," at the UNIX prompt enter:

EcDsBulkUndelete.pl -DFA -geoidfile <*geoid file*> [-mode <*ECS mode*>] [-server <*Sybase server*>] [-database <*database name*>] -user <*database login ID*> [-password <*database login password*>] [-log <*log filename*>]

- For example:

  EcDsBulkUndelete.pl -DFA -geoidfile PM1ATTNR_Dec23.geoid -mode OPS -database EcDsScienceDataServer1 -user sdsrv_role -server l0acg02_srvr -password *password* -log PM1ATTNR_Dec23UndelPhys.log

- The **-DFA** option indicates that "deleted" markings are to be removed for the granules in the archive only.
- The Bulk Undelete utility runs and removes "deleted" markings for the granules specified in the GeoID file in the archive.
- The Bulk Undelete utility records information about utility events in the log file (e.g., PM1ATTNR_Dec23UndelPhys.log).

6 When the Bulk Undelete utility has completed its run and the log file is available, at the UNIX prompt enter:

pg *<log filename>*

- ***<log filename>*** refers to the log file to be reviewed (e.g., PM1ATTNR_Dec23UndelPhys.log, PM1ATTNR_Dec23UndelDFA.log).
- The first page of the log file is displayed.
- For example:

```
l0acs03{cmops}[10]->pg PM1ATTNR_Dec23UndelPhys.log
*****************************************************
*            Bulk Granule Undeletion Tool
*
*
*  Report created 23 Dec 2005 10:46:22
*
*  GEOID file:    /usr/ecs/OPS/CUSTOM/utilities/PM1ATTNR_Dec23.geoid
*  Undeletion mode: OPS
*  Undeletion type: physical
*  Undeleting associated granules: yes
```

```
*
*  Total granules processed: 9740
*****************************************************
* TOTAL UPDATED GRANULES
*
* SC Type = 9740
*
* Associated granules undeleted = 4163
*
*****************************************************
*
* TOTAL FAILED GRANULES
*
*
*****************************************************

SC:PM1ATTNR.077:2013463060 -- Granule successfully marked
SC:PM1ATTNR.077:2013463061 -- Granule successfully marked
SC:PM1ATTNR.077:2013463062 -- Granule successfully marked
SC:PM1ATTNR.077:2013463063 -- Granule successfully marked
SC:PM1ATTNR.077:2013463064 -- Granule successfully marked
SC:PM1ATTNR.077:2013463065 -- Granule successfully marked
SC:PM1ATTNR.077:2013463071 -- Granule successfully marked
SC:PM1ATTNR.077:2013463072 -- Granule successfully marked
SC:PM1ATTNR.077:2013463073 -- Granule successfully marked
SC:PM1ATTNR.077:2013463074 -- Granule successfully marked
[…]
SC:PM1ATTNR.077:2013496391 -- Granule successfully marked
SC:PM1ATTNR.077:2013496392 -- Granule successfully marked
SC:PM1ATTNR.077:2013496393 -- Granule successfully marked

(EOF):
```

  - Many lines have been deleted from the example.
- Although this procedure has been written for the **pg** command, any UNIX editor or visualizing command (e.g., **vi**, **view, more, page**) can be used to review the file.

7 Review the log file to determine whether all "undeletions" occurred as expected or experienced errors.

- The following **pg** commands (at the **:** prompt) are useful:
  - **n** then **Return/Enter** (go to Page n).
  - **Return/Enter** or **+1** then **Return/Enter** (go down to the next page).
  - **-1** then **Return/Enter** (go back to the preceding page).
  - **+n** then **Return/Enter** (go down n number of pages).
  - **-n** then **Return/Enter** (go back n number of pages).
  - **+nl** then **Return/Enter** (go down n number of lines).

  - **-nl** then **Return/Enter** (go back n number of lines).
  - **$** then **Return/Enter** [go to the last page (end of file)].
  - **q** then **Return/Enter** (exit from pg).
- The following types of messages related to granule states may be seen in the log file (as shown in the example in Step 6):
  - Granule successfully marked.
  - No matching granule found.
  - Cannot UnDFA this granule type.
  - Granule is not DFA'd.
  - Granule is not Logically Deleted.
  - Unknown error undeleting granule.
  - Granule no longer in DsMdDeletedGranules table.
  - Invalid UR format.

8 If the GeoID and log files were written to the utilities directory (e.g., /usr/ecs/OPS/CUSTOM/utilities) rather than a home directory, log directory, or data directory, move or remove the files from the utilities directory.

- For example:

  mv PM1ATTNR_Dec23UndelPhys.log ../logs
  mv PM1ATTNR_Dec23.geoid /home/cmops/geoid

- Leaving GeoID files and log files in the utilities directory can eventually make it more difficult to find needed scripts in the directory.
- The utilities directory is intended for scripts, not log files or data files. Leaving GeoID files and log files in the utilities directory causes it to load up with extraneous files.

Phase 2, Physical Deletion

Physical Deletion involves the use of the Deletion Cleanup utility (EcDsDeletionCleanup.pl). A prerequisite to Physical Deletion is the completion of Phase 1, Logical Deletion. [Refer to the procedure for **Deleting Granules, Phase 1: Running the Bulk Delete Utility** (previous section of this lesson).]

Deleting Granules, Phase 2: Running the Deletion Cleanup Utility

Once granules have been marked/flagged for deletion, the operator runs the Deletion Cleanup utility, **EcDsDeletionCleanup.pl**. As previously mentioned the Deletion Cleanup utility removes all inventory rows (in the SDSRV database) for granules that were flagged as “deleted,” including rows referencing related information (e.g., BR, PH, and QA), and the script creates entries in the DsStPendingDelete table (in the STMGT database) for granules to be deleted.

The operations staff can control the lag time between logical deletion and physical deletion. The lag time is entered into the Deletion Cleanup utility, which deletes inventory entries only for granules that have been logically deleted prior to that time period.

Deleting Granules, Phase 2: Running the Deletion Cleanup Utility

---

1 Log in at the Sun Consolidation Internal Server host, where the Deletion Cleanup utility is installed.

- Examples of Sun Consolidation Internal Server host names include **e0acs11**, **g0acs11**, **l0acs03**, and **n0acs04**.
- For detailed instructions refer to the procedure for **Logging in to System Hosts** (previous section of this lesson).

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

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

3 To set up relevant environment variables (if desired) at the UNIX prompt enter:

setenv MY_MODE <*ECS mode*>
setenv SYB_USER sdsrv_role
setenv SYB_SQL_SERVER <*Sybase server*>
setenv SDSRV_DB_NAME <*database name*>
setenv STMGT_DB_NAME <*STMGT database name*>
setenv BATCH_SIZE <*batch size*>
setenv BATCH_SIZE_GRANULE <*batch size granule*>
setenv BATCH_SIZE_PH <*batch size history*>

- For example, the following commands would allow running the Deletion Cleanup utility using the OPS mode Science Data Server and Storage Management databases at the ASDC (LaRC):

  setenv MY_MODE OPS
  setenv SYB_USER sdsrv_role
  setenv SYB_SQL_SERVER l0acg02_srvr
  setenv SDSRV_DB_NAME EcDsScienceDataServer1
  setenv STMGT_DB_NAME stmgtdb1
  setenv BATCH_SIZE 5000
  setenv BATCH_SIZE_GRANULE 30
  setenv BATCH_SIZE_PH 5

- The **<*ECS mode*>** value specified for the **MY_MODE** parameter indicates the ECS mode (e.g., OPS, TS1, or TS2) in which to run the Deletion Cleanup utility.
- The **<*Sybase server*>** value specified for the **SYB_SQL_SERVER** parameter indicates the Sybase (database) server (e.g., e0acg11_srvr, g0acg01_srvr, l0acg02_srvr, or n0acg01_srvr) for the Science Data Server database.

- The <***database name***> value specified for the **SDSRV_DB_NAME** parameter indicates which Science Data Server database (e.g., EcDsScienceDataServer1, EcDsScienceDataServer1_TS1, or EcDsScienceDataServer1_TS2) is to be accessed by the Deletion Cleanup utility.
- The <***STMGT database name***> value specified for the **STMGT_DB_NAME** parameter indicates which Storage Management database (e.g., stmgtdb1, stmgtdb1_TS1, or stmgtdb1_TS2) is to be accessed by the Deletion Cleanup utility.
- The <***batch size***> value specified for the **BATCH_SIZE** parameter indicates the size of the batch used to transfer granules from SDSRV to STMGT.
  - The default value is 10,000, which is accepted by pressing **Return/Enter** at the prompt without entering a value first.
- The <***batch size granule***> value specified for the **BATCH_SIZE_GRANULE** parameter indicates the number of granules that will be deleted simultaneously from granule tables during granule cleanup.
  - The default value is 50, which is accepted by pressing **Return/Enter** at the prompt without entering a value first.
- The <***batch size history***> value specified for the **BATCH_SIZE_PH** parameter indicates the number of granules that will be deleted simultaneously from processing history tables during granule cleanup.
  - The default value is 5, which is accepted by pressing **Return/Enter** at the prompt without entering a value first.
- If the environment variables are set, the corresponding values do not have to be entered when starting the Deletion Cleanup utility.

4 To execute the Deletion Cleanup utility at the UNIX prompt type **EcDsDeletionCleanup.pl** and then press the **Return/Enter** key.

- If the environment variable MY_MODE is **not** set, the utility prompts **Enter Mode of Operation :.**

5 If prompted, at the **Enter Mode of Operation :** prompt type <***MODE***> and then press the **Return/Enter** key.

- If the environment variable SYB_USER is **not** set , the utility prompts **Enter Sybase User Name :**.

6 If prompted, at the **Enter Sybase User Name :** prompt type **sdsrv_role** and then press the **Return/Enter** key.

- The utility prompts **Enter Sybase password:**.

7 At the **Enter Sybase password:** prompt type <***password***> and then press the **Return/Enter** key.

- If the environment variable SYB_SQL_SERVER is **not** set, the utility prompts **Enter Sybase SQL Server Name :**.

8 If prompted, at the **Enter Sybase SQL Server Name :** prompt type ***<Sybase server>*** and then press the **Return/Enter** key.

- ***<Sybase server>*** is the Sybase (database) server (e.g., e0acg11_srvr, g0acg01_srvr, l0acg02_srvr, or n0acg01_srvr) for the Science Data Server database.
- If the environment variable SDSRV_DB_NAME is **not** set, the utility prompts **Enter SDSRV's database name :**.

9 If prompted, at the **Enter SDSRV's database name :** prompt type ***<SDSRV database name>*** and then press the **Return/Enter** key.

- The ***<SDSRV database name>*** is the relevant Science Data Server database (e.g., EcDsScienceDataServer1, EcDsScienceDataServer1_TS1, or EcDsScienceDataServer1_TS2).
- If the environment variable STMGT_DB_NAME is **not** set, the utility prompts **Enter STMGT's database name :**.

10 If prompted, at the **Enter STMGT's database name :** prompt type ***<STMGT database name>*** and then press the **Return/Enter** key.

- The ***<STMGT database name>*** is the relevant Storage Management database (e.g., stmgtdb1, stmgtdb1_TS1, or stmgtdb1_TS2).
- If the environment variable BATCH_SIZE is **not** set, the utility prompts **Enter Batch Size for stmgt migration(10000) :**.

11 If prompted, at the **Enter Batch Size for stmgt migration(10000) :** prompt type ***<batch size>*** and then press the **Return/Enter** key.

- ***<batch size>*** is the size of the batch used to transfer granules from SDSRV to STMGT.
- The default value is 10,000, which is accepted by pressing **Return/Enter** at the prompt without entering a value first.
- If the environment variable BATCH_SIZE_GRANULE is **not** set, the utility prompts **Enter Batch Size for Granule Deletion sdsrv(50) :**.

NOTE: Take care when increasing the BATCH_SIZE_GRANULE and BATCH_SIZE_PH values beyond the recommended default values. If the values are set too high, the database tables will be locked and all available Sybase locks will be used up.

12 If prompted, at the **Enter Batch Size for Granule Deletion sdsrv(50) :** prompt type ***<batch size granule>*** and then press the **Return/Enter** key.

- ***<batch size granule>*** represents the number of granules that will be deleted simultaneously from granule tables during granule cleanup.
- The default value is 50, which is accepted by pressing **Return/Enter** at the prompt without entering a value first.
- If the environment variable BATCH_SIZE_PH is **not** set, the utility prompts **Enter Batch Size for Processing History Deletion sdsrv(5) :**.

13 If prompted, at the **Enter Batch Size for Processing History Deletion sdsrv(5) :** prompt type ***<batch size history>*** and then press the **Return/Enter** key.

- ***<batch size history>*** represents the number of granules that will be deleted simultaneously from processing history tables during granule cleanup.
- The default value is 5, which is accepted by pressing **Return/Enter** at the prompt without entering a value first.
- The utility prompts **Enter Log File name:**.

14 To use a log file name other than the default name at the **Enter Log File name:** prompt type ***<log file name>*** and then press the **Return/Enter** key.

- ***<log file name>*** (e.g., **DeletionCleanup20051130.LOG**) is the name of the file where EcDsDeletionCleanup.pl records process events.
- To accept the default log file name (**DeletionCleanup.LOG**) press **Return/Enter** at the prompt without entering a value.
- If there is already a log file with the name that was entered, new entries will be appended to that log; otherwise a new log file will be created.
- If ***<log file name>*** does not include a **.LOG** file extension, Deletion Cleanup appends the extension to the name.
- Deletion Cleanup prepares to connect to the database and displays a **Ready to get into sdsrv database...** message.
- Then Deletion Cleanup connects to the database and checks for leftover granules that need to be processed. The following messages are displayed:

  First check if there are leftover granules need to be processed*************

  You may skip the following part if the log file for the previous run indicated a success or if it gave suggestion that it could be done automatially.

  Do you want to proceed with this part? [y/n] :

15 To process the granules in the DsMdStagingTable only (to skip checking for leftover granules from a previous run of the Deletion Cleanup utility) at the **Do you want to proceed with this part? [y/n] :** prompt type **n** and then press the **Return/Enter** key.

- The following message is displayed:

  Operator chose to only process the granules left over in the DsMdStagingTable

- Deletion Cleanup either processes leftover granules or determines that there are no leftover granules from previous runs.
- Eventually the following message is displayed:

  Do you wish to continue deleting some more granules? [y/n] :

- Go to Step 20.

16 To check for leftover granules from a previous run of the Deletion Cleanup utility at the **Do you want to proceed with this part? [y/n] :** prompt type **y** and then press the **Return/Enter** key.

- The following message is displayed:

  To ensure that all the granules are deleted for the previous run, StartTime, EndTime and subType file from the previous log file in the user input section need to be passed in.
  If these info are not available in the log file, the previous run can not be recovered because script failed before the user input was gathered

  Please enter the StartTime from the log file for the last run or hit return if not available:

17 To continue checking for leftover granules at the **Please enter the StartTime from the log file for the last run or hit return if not available:** prompt type ***<start time>*** and then press the **Return/Enter** key.

- ***<start time>*** represents the start time (e.g., 10/15/2005) for the last run (from the log file).
- If the start time for the last run is not available, press **Return/Enter** at the prompt without entering a value.
- The utility prompts **Please enter the EndTime from the log file for the last run or hit return if not available:**

18 To continue checking for leftover granules at the **Please enter the EndTime from the log file for the last run or hit return if not available:** prompt type ***<end time>*** and then press the **Return/Enter** key.

- ***<end time>*** represents the end time (e.g., 10/18/2005) for the last run (from the log file).
- If the end time for the last run is not available, press **Return/Enter** at the prompt without entering a value.
- The utility prompts **Please enter the subType file name from the log file for the last run or hit return if not available:**

19 To continue checking for leftover granules at the **Please enter the subType file name from the log file for the last run or hit return if not available:** prompt type ***<subtype>*** and then press the **Return/Enter** key.

- ***<subtype>*** represents the subType (e.g., PM1ATTNR.077) for the last run (from the log file).
- If the subType for the last run is not available, press **Return/Enter** at the prompt without entering a value.
- Deletion Cleanup either processes leftover granules or determines that there are no leftover granules from previous runs.
- If there are leftover files, the following series of messages is displayed (there may be some time delay between messages):

  Start processing leftover granules*********************

executing…

No. of files to be migrated:: …

Successfully Committed the migration of …files

Successful Migration.

Finish processing leftover granules from previous run(s)***************

Do you wish to continue deleting some more granules? [y/n] :

- If there are no leftover files, the following series of messages is displayed:

  No leftover granules from previous run(s)************************

  Do you wish to continue deleting some more granules? [y/n] :

20 To exit Deletion Cleanup without doing any new cleanup at the **Do you wish to continue deleting some more granules? [y/n] :** prompt type **n** and then press the **Return/Enter** key.

- An **Operator chose to Terminate granule deletion process after processing the leftover granules** message is displayed.
- Go to Step 30.

21 To proceed with deletion cleanup at the **Do you wish to continue deleting some more granules? [y/n] :** prompt type **y** and then press the **Return/Enter** key.

- The following menu is displayed:

  The user must select the menu option:
  1. Select granules for a specific day (lag number or date <mm/dd/yyyy> format)
  2. Select all granules older than (include) a specific day (lag number or date <mm/dd/yyyy> format)
  3. QUIT
  Choose 1, 2 or 3 ==>

22 At the menu prompt type **<*number*>** and then press the **Return/Enter** key.

- **<*number*>** is **1**, **2**, or **3**.
- Enter **1** to clean up granules for a specific day.
  - The message **Enter Lag Time in Days or date <mm/dd/yyyy>:** is displayed.
- Enter **2** to clean up all granules older than (and including) a specific day.
  - The message **Enter Lag Time in Days or date <mm/dd/yyyy>:** is displayed.
- Enter **3** to exit from Deletion Cleanup.
  - The message **bye** is displayed.
- Go to Step 30.

23 At the **Enter Lag Time in Days or date <mm/dd/yyyy>:** prompt type ***<lag time>*** and then press the **Return/Enter** key.

- ***<lag time>*** is expressed as either a number of days (e.g., 17) or a date (in ***mm/dd/yyyy*** format, e.g., 10/18/2005).
  - An entry of **0** (zero) is equivalent to the current date.
- The lag time is the time from the current date at which the operator wishes to delete granules and is in units of days. For example, if the current date were November 30, 2005 and if it were necessary to delete all granules marked for deletion starting on November 25, 2005 (i.e. to delete all granules with a date of November 25, 2005 or earlier), the operator would enter either a lag time of **5** or the date **11/25/2005**.
- A lag time of **0** (zero) results in the deletion of **all** granules that have been flagged for deletion.
- One of the following types of messages is displayed:

  A Lag Time of '5' Days has been entered.
  Is this correct? [y/n]

  OR

  A specific day of '11/25/2005' has been entered.
  Is this correct? [y/n]

24 If the displayed lag time is **not** correct, at the **Is this correct? [y/n]** prompt type **n** and then press the **Return/Enter** key.

- An **Enter Lag Time in Days or date <mm/dd/yyyy>:** prompt message is displayed.
- Return to Step 23.

25 If the displayed lag time is correct, at the **Is this correct? [y/n]** prompt type **y** and then press the **Return/Enter** key.

- The following menu is displayed:

  The user must select the menue option:
  1. Select above for specific datatypes(wildcard as an option) and versions
  In the list of subtype format: AST_L1BT.001 or AST_L1B*.001...
  2. Select the above for all datatypes and versions
  3. QUIT
  Choose 1, 2 or 3 ==>

26 At the menu prompt type ***<number>*** and then press the **Return/Enter** key.

- ***<number>*** is **1**, **2**, or **3**.
- Enter **1** to specify a file of dataypes for cleanup.
  - The following message is displayed:

    This will select granules for specific datatypes in the User Specified Input File
    Please specify the full path/filename ==>

- Enter **2** to select all datatypes for cleanup.
  - The following type of message is displayed:

    Search against DsMdDeletedGranules table return the following results:

    ShortName: PM1ATTNR, VersionID: 77, Num of Granules: 9740, DFA granules: YES

    The number of Inventory Granules to be deleted = 9740
    The number of DFA Granules to be deleted = 9740
    Do you wish to continue deleting these granules?
- Enter **3** to exit from Deletion Cleanup.
  - The message **bye** is displayed.
  - Go to Step 30.

27 If a **Please specify the full path/filename ==>** prompt is displayed, type **<*path*>** and then press the **Return/Enter** key.

- **<*path*>** is the full path and file name for the input file of data types for cleanup.
- For example: **/usr/ecs/OPS/CUSTOM/data/DSS/datatypes20051130.dat**
- The following type of message is displayed:

  Search against DsMdDeletedGranules table return the following results:

  ShortName: PM1ATTNR, VersionID: 77, Num of Granules: 9740, DFA granules: YES

  The number of Inventory Granules to be deleted = 9740
  The number of DFA Granules to be deleted = 9740
  Do you wish to continue deleting these granules?

28 To **not** continue with the cleanup at the **Do you wish to continue deleting these granules?** prompt type **n** and then press the **Return/Enter** key.

- The message **Operator chose to Terminate granule deletion process after getting the counts from DsMdDeletedGranules** is displayed.
- Go to Step 30.

29 To continue with the cleanup at the **Do you wish to continue deleting these granules?** prompt type **y** and then press the **Return/Enter** key.

- The following series of messages is displayed (there may be some time delay between messages):

  executing exec ProcPhysicalDeleteFmStaging 50,5
  First delete all the science related granules

  Begin Insert to #t1 for SC granules that will be deleted
  ....
  .....
  .....
  finshed Physical Deleting. Continue migrating deleted granules to Stmgt
  No. of files to be migrated:: 2

Successful Migration.
Wed Nov 30 13:03:16 EST 2005

Successfully finished running Deletion CleanUp Utility.

30 To rerun the utility (e.g., in response to a message recommending that the utility be rerun) return to Step 4.

Phase 3, Deletion from the Archive

Deletion from the Archive involves the use of the Archive Batch Delete utility (EcDsStArchiveBatchDelete.pl). A prerequisite to Deletion from the Archive is the completion of Phase 2: Physical Deletion. [Refer to the procedure for **Deleting Granules, Phase 2: Running the Deletion Cleanup Utility** (previous section of this lesson).]

Deleting Granules, Phase 3: Running the Archive Batch Delete Utility

The operator uses the Archive Batch Delete utility to initiate the removal from the archive of the files listed in the DsStPendingDelete table (populated by SDSRV in Phase 2). The operator must run the utility separately on each appropriate Archive host (FSMS Server or APC Server host) because the utility lists and deletes files from archive on the local host only. The Archive Batch Delete utility creates requests to the local archive server host to delete selected files. Files that are successfully deleted have their associated rows removed from the DsStPendingDelete table.

Deleting Granules, Phase 3: Running the Archive Batch Delete Utility

---

1 Log in at the Sun Consolidation Internal Server host, where the Archive Batch Delete utility is installed.

- Examples of FSMS Server host names include **e0drg11**, **g0drg01**, **l0drg01**, and **n0drg01**.
- Examples of APC Server host names include **e0acg11**, **g0acg01**, **l0acg02**, and **n0acg01**.
- For detailed instructions refer to the procedure for **Logging in to System Hosts** (previous section of this lesson).

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

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

3 To set up relevant environment variables (if desired) at the UNIX prompt enter:

setenv MODE <*ECS mode*>
setenv SYB_SQL_SERVER <*Sybase server*>
setenv SYB_DB_NAME <*database name*>

- For example, the following commands would allow running the Archive Batch Delete utility using the OPS mode Storage Management database at the ASDC (LaRC):

  setenv MODE OPS
  setenv SYB_SQL_SERVER l0acg02_srvr
  setenv SYB_DB_NAME stmgtdb1

- The <***ECS mode***> value specified for the **MODE** parameter indicates the ECS mode (e.g., OPS, TS1, or TS2) in which to run the Archive Batch Delete utility.
- The <***Sybase server***> value specified for the **SYB_SQL_SERVER** parameter indicates the Sybase (database) server (e.g., e0acg11_srvr, g0acg01_srvr, l0acg02_srvr, or n0acg01_srvr) for the Science Data Server database.
- The <***database name***> value specified for the **SYB_DB_NAME** parameter indicates which Storage Management database (e.g., stmgtdb1, stmgtdb1_TS1, or stmgtdb1_TS2) is to be accessed by the Archive Batch Delete utility.
- If the environment variables are set, the corresponding values do not have to be entered when starting the Archive Batch Delete utility.

4 To execute the Deletion Cleanup utility at the UNIX prompt type **EcDsStArchiveBatchDelete.pl** and then press the **Return/Enter** key.

- If the environment variable MODE is **not** set, the utility prompts **Enter Mode of Operation :**.

NOTE: If **Ctrl-c** is pressed during the deletion process, the process is terminated, the utility prints out the reports and displays the message **Process was aborted.**

5 If prompted, at the **Enter Mode of Operation :** prompt type <***MODE***> and then press the **Return/Enter** key.

- The utility prompts **Enter log file name (or return if use default log file):**.

6 To use a log file name other than the default name at the **Enter log file name (or return if use default log file):** prompt type <***log file name***> and then press the **Return/Enter** key.

- <***log file name***> (e.g., **EcDsStArchiveBatchDelete.ALOG**) is the name of the file where EcDsStArchiveBatchDelete.pl records process events.
- To accept the default log file name (**ArchiveBatchDelete.** <***date***>**.**<***time***>**.log**) press **Return/Enter** at the prompt without entering a value.
- The utility displays a message similar to the following message:

  log information will be put to
  /usr/ecs/OPS/CUSTOM/logs/ArchiveBatchDelete.20051201.161030.log

- The utility prompts **Enter db login :**.

7 At the **Enter db login :** prompt type **stmgt_role** and then press the **Return/Enter** key.

- The utility prompts **Enter db password :**.

8 At the **Enter db password :** prompt type ***<password>*** and then press the **Return/Enter** key.

- If the environment variable SYB_SQL_SERVER is **not** set, the utility prompts **Enter Sybase SQL Server Name :**.

9 If prompted, at the **Enter Sybase SQL Server Name :** prompt type ***<Sybase server>*** and then press the **Return/Enter** key.

- ***<Sybase server>*** is the Sybase (database) server (e.g., e0acg11_srvr, g0acg01_srvr, l0acg02_srvr, or n0acg01_srvr) for the Storage Management database.
- If the environment variable SYB_DB_NAME is **not** set, the utility prompts **Enter STMGT's database name :**.

10 If prompted, at the **Enter STMGT's database name :** prompt type ***<database name>*** and then press the **Return/Enter** key.

- The ***<database name>*** is the relevant Storage Management database (e.g., stmgtdb1, stmgtdb1_TS1, or stmgtdb1_TS2).
- Archive Batch Delete connects to the database, lists MissingVolGroup information from the Stage column of the DsStPendingDelete table, and allows the operator to reset the Stage field to NULL. It displays the following kinds of messages:

  Ready to get into stmgt database... stmgtdb1.
  Listing MissingVolGroup information in DsStPendingDelete...

  The following is the list of MissingVolGroup information in DsStPendingDelete table:

  VersionedDataType filecount

  total file count: 0

  No files in DsStPendingDelete table having stage MissingVolGroup.

  Do you want to continue? [y/n] :

  – In the example no file was found with missing volgroup.

11 To exit from the utility at the **Do you want to continue? [y/n] :** prompt type **n** and then press the **Return/Enter** key.

- The following message is displayed:

  Terminating batch delete process.

12 To continue with the archive batch deletion process at the **Do you want to continue? [y/n] :** prompt type **y** and then press the **Return/Enter** key.

- The following types of messages are displayed:

  Trying to get HardWareCI from stmgtdb1...
  HardWareCI is DRP1
  server id is 3
  exec DsStPDInsertComplete to fill on DsStPendingDelete, this may take a while...

- Then there is a prompt for the batch size for log deletion:

  Please enter batch size for log deletion progress (or return if use default value 100) :

13 To continue with the archive batch deletion process at the **Please enter batch size for log deletion progress (or return if use default value 100) :** prompt type ***<batch size>*** and then press the **Return/Enter** key.

- ***<batch size>*** represents the batch size for log deletion progress, which means that every time the specified number of files have been processed Archive Batch Deletion reports/logs the total number of files successfully deleted and total number of files which failed to be deleted.
- To accept the default value (100) press **Return/Enter** at the prompt without entering a value.
  - If the default value (100) is accepted, every time 100 files have been processed Archive Batch Deletion reports/logs the total number of files successfully deleted and total number of files which failed to be deleted.
- The utility shows deletion information from the DsStPendingDelete table:

  Listing information in the DsStPendingDelete...

  The following is the list of information in DsStPendingDelete table for DRP1 :

VersionedDataType	VolumeGroupId	VolumeGroupPath	filecount
0: AST_L1BT.001	**182**	**/archive/OPS/drp/astl1**	**4**
1: AST_L1BT.001B	**167**	**/archive/OPS/acm/browse**	**4**
2: Browse.001	**104**	**/archive/OPS/acm/browse**	**2**

  total file count: 10

  Do you want to continue? [y/n] :

14 To exit from the utility at the **Do you want to continue? [y/n] :** prompt type **n** and then press the **Return/Enter** key.

- The following message is displayed:

  Terminating batch delete process.

15 To continue with the archive batch deletion process at the **Do you want to continue? [y/n] :** prompt type **y** and then press the **Return/Enter** key.

- Then there is a prompt for selecting the index number(s) of the VersionedDataType/VolumeGroupPath set(s) to be deleted:

  Please enter the indexes of the desired VersionedDataType and VolumeGroupPath pair with space between indexes or just -1 for all)=>

16 To continue with the archive batch deletion process at the **Please enter the indexes of the desired VersionedDataType and VolumeGroupPath pair with space between indexes or just -1 for all)=>** prompt type ***<index1>* [*<index2>* … *<index*n*>*]** and then press the **Return/Enter** key.

- ***<index1>* [*<index2>* … *<index*n*>*]** represent the index numbers (from the previously displayed list) of the VersionedDataType/VolumeGroupPath sets to be deleted. (Refer to the example in Step 13 for a list of VersionedDataType/VolumeGroupPath sets with index numbers.)
- To select all VersionedDataType/VolumeGroupPath sets for deletion either enter all index numbers separated by spaces or just enter **–1**.
- For example:

  Please enter the indexes of the desired VersionedDataType and VolumeGroupPath pair with space between indexes or just -1 for all)=> 0

  You selected the following VersionedDataType/VolumeGroupPath sets:

  0: AST_L1BT.001 182 /archive/OPS/drp/astl1 4

- In the example index number **0** was entered and Archive Batch Deletion displayed relevant information associated with that entry in the DsStPendingDelete table (as shown in Step 13)
- Then there is a prompt for deleting a subset of the selected granules:

  Do you want to delete a subset of the granules above? [y/n] :

17 To delete the entire set of selected granules (to skip selecting a subset) at the **Do you want to delete a subset of the granules above? [y/n] :** prompt type **n** and then press the **Return/Enter** key.

- The following types of messages are displayed:

  Select files in DsStPendingDelete for deletion...

  Start of Batch Delete...
  get Directory /archive/OPS/drp/astl1
  Total No. of files to be deleted in this session: 3
  No. of files to be deleted: 0, deleted: 3, failed to be deleted: 1
  Files deleted in this session: 3
  Files failed to be deleted in this session: 1
  Total files deleted so far: 3
  Total files failed to be deleted so far: 1

  Listing information in the DsStPendingDelete...

  The following is the list of information in DsStPendingDelete table for DRP1 :

VersionedDataType	VolumeGroupId	VolumeGroupPath	filecount
0: AST_L1BT.001	182	/archive/OPS/drp/astl1	1
1: AST_L1BT.001B	167	/archive/OPS/acm/browse	4
2: Browse.001	104	/archive/OPS/acm/browse	2

total file count: 7

Do you want to continue? [y/n] :

- The example indicates that three files were deleted from Index 0 (AST_L1BT.001) and one failed to be deleted. (Compare with the example in Step 13.)

- Return to Step 14.

18 To delete a subset of the selected granules at the **Do you want to delete a subset of the granules above? [y/n] :** prompt type **y** and then press the **Return/Enter** key.

- Then there is a prompt for entering the subset size:

  Please enter subset size :

- If a subset is entered, the utility selects the file(s) to be deleted based on the selected VersionedDataType/VolumeGroupPath and the granule FileName. Older files are deleted first.

19 To delete a subset of the selected granules at the **Please enter subset size :** prompt type <***subset size***> and then press the **Return/Enter** key.

- <***subset size***> represents the subset size.
- For example:

  Please enter subset size : 1

  Subset size: 1
  Select files in DsStPendingDelete for deletion...

  Start of Batch Delete...
  get Directory /archive/OPS/drp/astl1
  Total No. of files to be deleted in this session: 1
  No. of files to be deleted: 0, deleted: 1, failed to be deleted: 0
  Files deleted in this session: 1
  Files failed to be deleted in this session: 0
  Total files deleted so far: 1
  Total files failed to be deleted so far: 0

  Listing information in the DsStPendingDelete...

  The following is the list of information in DsStPendingDelete table for DRP1 :

VersionedDataType	VolumeGroupId	VolumeGroupPath	filecount
0: AST_L1BT.001	182	/archive/OPS/drp/astl1	3
1: AST_L1BT.001B	167	/archive/OPS/acm/browse	4
2: Browse.001	104	/archive/OPS/acm/browse	2

  total file count: 9

  Do you want to continue? [y/n] :

- The example indicates that one file was deleted from Index 0 (AST_L1BT.001). (Compare with the example in Step 13.)

20 Return to Step 14.

Loading Archive Media

Automatically Loading Archive Media

For the STK storage facility, each Powderhorn is equipped with a 21-tape Cartridge Access Port (CAP). In automatic mode, tapes may be placed in the CAP for automatic loading. Tapes are also ejected through the CAP when identified for ejection using a command at the host for the STK Automated Cartridge System Library Software (ACSLS).

The following procedure addresses automatic media loading.

Automatically Load STK Powderhorn Archive Media

1 Log in as **amass** or **root** at the FSMS SGI host (workstation ***x*0drg##**).

- The ***x*** in the workstation name will be a letter designating your site: **g** = GSFC, **m** = SMC, **l** = LaRC, **e** = LP DAAC, **n** = NSIDC, **o** = ORNL, **a** = ASF, **j** = JPL; the **##** will be an identifying two-digit number (e.g., **n0drg01** indicates an FSMS SGI server at NSIDC).

2 At the FSMS host, type **/usr/amass/bin/bulkinlet SP** and then press the **Return/Enter** key.

- The Cartridge Access Port (CAP) door unlocks (audible unlatching sound).
- If you have removed an existing volume and are re-inserting it, do not use the **SP** option, which puts the volume in the general space pool. Instead type **/usr/amass/bin/bulkinlet <*volgrp*>**, where **<*volgrp*>** is the volume group from which the volume was removed. This will put the volume back where it was before removal.

3 Write down or note the bar code number(s) on the label(s) of the cartridge(s), open the recessed latch on the CAP door and insert the tape(s), solid black side up, with the bar code label facing you, and close the door.

- The robot scans all the volumes.
- Data for the newly inserted media are displayed, including bar codes, associated volume numbers, and, in the **flag** column, the letters **IUO**, indicating that the volumes are inactive (**I**), unformatted (**U**), and offline (**O**).

4 For any newly inserted media, it is necessary to issue a formatting command. For the new 9940 tapes, type **/usr/amass/bin/volformat -b 256k ###** and then press the **Return/Enter** key.

- ### is the volume number. You can enter more than one volume number, separating each number from the preceding one with a space.
- A message requests confirmation that you wish to continue.

5 Type **y** and then press the **Return/Enter** key.

- A message is displayed requesting further confirmation, stating that **The following volumes will be formatted:** and listing volume numbers, followed by **(Y-N)**.

6 Type **y** and then press the **Return/Enter** key.

- After a few minutes, a message **Completed formatting all volumes** is displayed.

7 To verify that the volume(s) are inserted, type **/usr/amass/bin/vollist** and then press the **Return/Enter** key.

- Data for the media are displayed; the **flag** column shows that the newly formatted volumes are inactive (**I**).

8 To activate the media for use, type **/usr/amass/bin/volstat** and then press the **Return/Enter** key.

- Data for the media are displayed; the **flag** column shows that the volumes are now active (**A**).

Manually Loading Archive Media

With the **bulkload** command, you bypass the CAP and manually load media directly into the library bins. Typically, this will only be done at the initial load of the system with large numbers of media volumes. The **bulkload** command enables AMASS to determine what type of media have been placed in the library and to convey this information to the AMASS database. The following procedures are applicable.

Manually Load STK Powderhorn Archive Media

---

1 To manually insert a tape into the Powderhorn, login to the control software (ACSLS) using the acssa account.

2 Type **enter 0,0,0** and then press the **Return/Enter** key.

- The Cartridge Access Port (CAP) door unlocks (audible unlatching sound).

3 Write down or note the bar code number(s) on the label(s) of the cartridge(s), open the recessed latch on the Cartridge Access Port (CAP) door and insert the tape(s), solid black side up, with the bar code label facing you, and close the door.

- The robot scans all the volumes.

4 At the AMASS host, type **bulkload -s SP** and then press the **Return/Enter** key.

- The AMASS database is populated with data for the volumes in the library, including bar codes, associated volume numbers, and status -- inactive (**I**), unformatted (**U**), and offline (**O**). The data may be reviewed using the **vollist** command.

NOTE: If you are loading a very large number of volumes, such as at initial load, and choose to bypass the CAP and place the volumes directly in the LSM slots, data about the volumes will not be immediately available to ACSLS for communication to AMASS. You will first have to use the ACSLS **audit** command to initiate an audit of the LSM, a process that may take several hours.

Caution

Inactivate AMASS before using the following command.

5 To view a list of media in the library, type **medialist -3** and then press the **Return/Enter** key.

- The **-3** option indicates the STK Powderhorn.
- The utility reads the library element status stored in the library, and information about the library contents, including the status (**FULL** or **EMPTY**) of the elements.

Formatting a Volume

Volume formatting was part of the procedure for automatically loading media (previous section of this lesson). However, it may be necessary or desirable at some time to format a volume that is already present in the archive. To format a volume, it must be online. A volume is placed online using the **volloc** command. Formatting a volume destroys any files on that volume. Before formatting a volume, check to make sure it does not have any files that should be saved. The following procedure is applicable.

Format a Tape Volume

1 To put the volume online, at the FSMS host, type **/usr/amass/bin/volloc -n** ***###*** and then press the **Return/Enter** key.

- ***###*** is the number of the volume.

2 To verify there are no files on volume, type **/usr/amass/bin/volfilelist** ***<Vol. No.>*** and then press the **Return/Enter** key.

- No files are displayed.
- If a list of files is returned, indicating that the volume is not empty, before proceeding verify that you have the correct volume and that it is to be formatted.

3 To format the volume, type **/usr/amass/bin/volformat -b 256k** ***###*** and then press the **Return/Enter** key.

- ***###*** is the number of the volume.

4 To verify status of the volume, type **/usr/amass/bin/volprint -a** ***###*** and then press the **Return/Enter** key.

- ***###*** is the number of the volume.

Remove Media

To remove media from the archive, use the following procedures.

Remove Media from the STK Powderhorn

1 Determine which volumes you want to remove by utilizing the volume number. If necessary to review volume numbers and other information, log into the AMASS host, type **/usr/amass/bin/vollist** and then press the **Return/Enter** key.

2 If there are only a few volumes to remove, from AMASS, for each volume to be removed type **/usr/amass/bin/voloutlet** ***###*** and then press the **Return/Enter** key.

- ***###*** is the volume number.
- AMASS marks the volume off-line and the volume is transferred to the CAP.

3 For the STK Powderhorn, open the recessed latch on the Cartridge Access Port (CAP) door and remove the tape(s)

Backup Archived Data

The archive design incorporates programmed backups of archived data. System requirements specify that a percentage of archived data be duplicated for local and off-site storage to provide for data safety. However, the large volume of archived data merits finding alternatives to complete backup of all volumes in the libraries. Selection of data for backup is based on assessment of the feasibility of recovery in the event of data loss.

It is imperative to backup data that would be irretrievable if lost. Such data are saved to the archive, saved to local backup, and saved to offsite backup. Many data elements that will be archived, however, could be retrieved in the event of loss. For example, in the event of loss of a higher-level product that is an output of processing a lower level product, it would be possible to restore the higher-level product by reprocessing the lower level product. As another example, the system will often archive a lower level product from a data provider, but that product may also be retained in the archives of the data provider. If the product were lost from the ECS archive, it would be possible to ingest it again from the data provider, using appropriate Ingest procedures.

Thus, when data are inserted into the archive (e.g., through Ingest, from Processing), up to three copies of the data may be created, reflecting different types of data use:

- the active archive copy, available for distribution or other use (volume group is specified in the *Archive ID*).
- a copy to be retained for local backup (volume group is specified in the *Backup ID*, typically created by appending "B" to the ShortName and VersionID; e.g., AST_L1B.001B).
- a copy to be sent to offsite backup storage (volume group is specified in the *Offsite ID*, typically created by appending "O" to the ShortName and VersionID).

Creating Offsite Backups

The paths for creation of the data copies are specified for each ESDT when it is loaded, using the **Add Volume Group** dialog accessible from the **Vol Grp. Config.** tab on the Storage Management GUI as illustrated in Figure 20. The Archive ID (for the archive copy) and the Backup ID (for the local backup copy) should reflect different archives if possible (i.e., different LSMs), to spread the risk of loss. The Offsite ID will not be a remote site path, but rather a local path for making copies to be sent for offsite storage. The requirements to implement creation of offsite backups (see Document 611-EMD-001, *Release 7.11 Mission Operation Procedures for the EMD Project*, Section 17.6.1.1) include:

- creating a subdirectory and volume group for offsite backups.

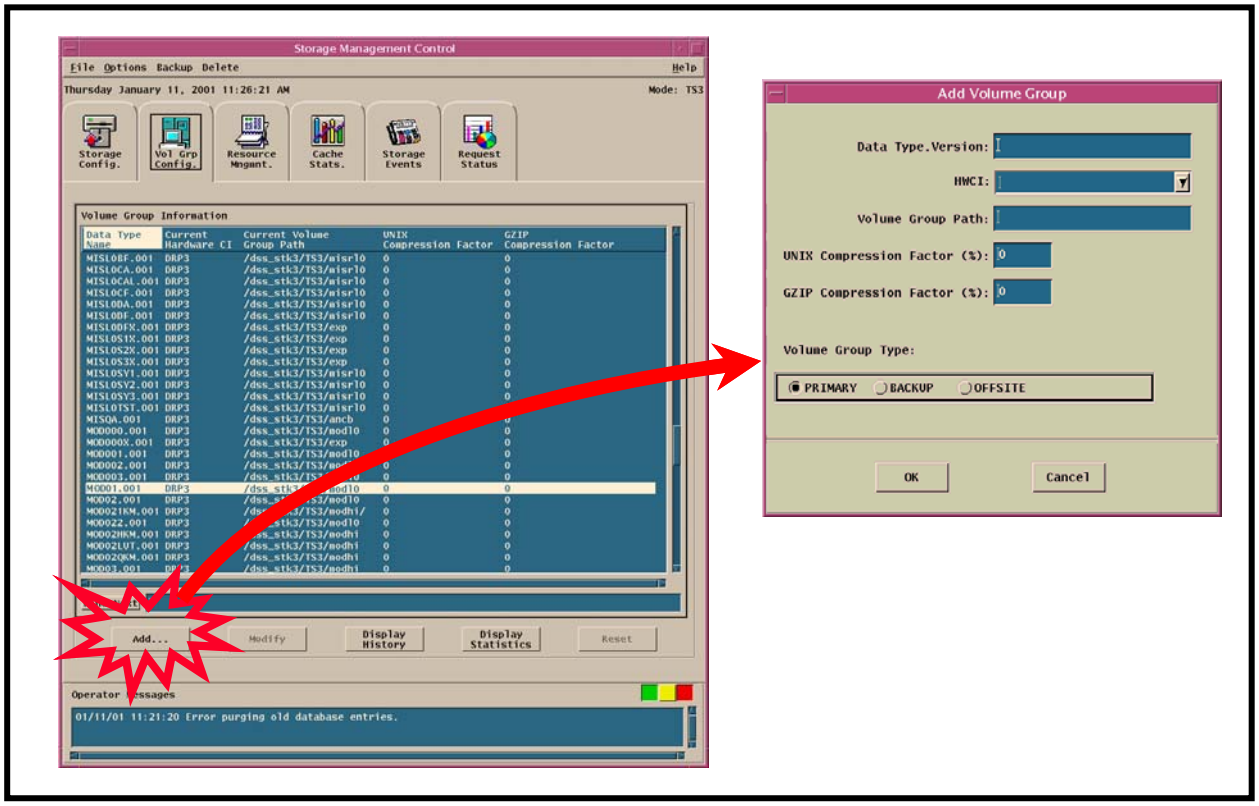


Figure 20. Storage Management, Add Volume Group Dialog

- using the **Vol Grp Config.** tab of the Storage Management GUIs to add the volume group to the appropriate archive server and set the offsite ID to be the three-character specification for the local site (e.g., EDC, GSF, LAR, NSC).
- updating the STMGT database offsite table with the volume group name created for offsite storage (may require assistance of Database Administrator).
- adding volumes to the volume group as needed.

Each site is responsible for arranging its own secure offsite storage. The offsite backup cartridges are removed from the archive storage facility using procedures already described. For local and/or offsite storage of specific archive data, the DAAC Archive Manager generates or directs the generation of a list of selected data. At the time the files are archived, they are written to specific volume groups that correspond to the three data usage types identified in the preceding paragraph. Only files belonging to the data usage type are written to the tapes in a specific volume group. Hence, the Archive Manager can determine the tapes that should be stored for local backup and those for offsite storage. This can be accomplished using the AMASS administration **vollist** command.

vollist - this command lists all volumes and their current status information. It will optionally accept a specific volume number. Information displayed includes: volume number, volume group, jukebox number, slot position, volume label/bar code, current

volume status, amount of used space on the volume, amount of available space on volume, percentage of space no longer referenced, and number of read or write errors on volume.

If there are other files designated for local and offsite backup that have not been written to a specific volume group, the Archive Manager can use the appropriate AMASS administration commands, **dirfilelist** and/or **volfilelist**, to locate the appropriate archive volume that contains the designated archived files.

dirfilelist: this command lists the files under a directory. One directory or file is displayed on a line. Information displayed includes: file name, volume number, starting block number, file size, file permissions, number of hard links, numeric user id, numeric group id, last time file was accessed, and last time file was modified.

volfilelist: this command lists all of the files on a volume and accepts as input a volume number. One file is displayed on a line. Information displayed includes: file name, volume number, starting block number, file size, file permissions, number of hard links, numeric user id, numeric group id, last time file was accessed, and last time file was modified.

For more information about these AMASS commands, refer to the *AMASS System Administrator's Guide*. The following procedure is applicable for creating offsite backups for science data.

Create Offsite Backups for Science Data

---

1 Launch the DSS Storage Management GUI using UNIX commands.

- For detailed instructions refer to the **Launch DSS GUIs Using UNIX Commands** procedure (previous section of this lesson).
- The DSS Storage Management GUI is displayed.

2 Click on the **Vol Grp Config.** tab to display the Volume Group information.

- The **Vol Grp Config.** tab information is displayed.

3 Click on the **Add . . .** button below the Volume Group Information field.

- The **Add Volume Group** window is displayed.

4 In the **Add Volume Group** window click in the **Data Type.Version:** field.

- The cursor moves to the **Data Type.Version:** field.

5 Type the ESDT ***ShortName*** and ***Version*** (e.g., MOD01.001) of the data type for which the volume group is to be created.

- The typed entry appears in the **Data Type.Version:** field.

6 In the **Add Volume Group** window click on the pull-down arrow at the end of the **HWCI:** field.

- A pull-down menu displays designators of the hardware configuration items available for storing data.

7 Click on the designator for the hardware configuration item where the archive copies of data for the ESDT are to be stored.

- The selected designator is displayed in the **HWCI:** field.

8 In the **Add Volume Group** window click in the **Volume Group Path:** field.

- The cursor moves to the **Volume Group Path:** field.

9 Type the full path identification for the storage of active archive data for the ESDT.

- Typically, the path will be of the form **dss_stk*n*/<*MODE*>/*xxxxx***, where ***n*** is a number designating a StorageTek Library Storage Module, ***MODE*** is **OPS**, **TS1**, or **TS2**, and ***xxxxx*** is a short identifier for what is being stored; e.g., **dss_stk1/OPS/modl0**.
- The typed entry appears in the **Volume Group Path:** field.

10 In the **Volume Group Type:** radio box, click on the **PRIMARY** button.

- The button's depressed appearance indicates selection of **PRIMARY**, signifying that the volume group being created is for primary storage for active archive use.

11 Click on the **Save and Add Next VG** button at the bottom of the **Add Volume Group** window.

- The volume group is created for display in the **Volume Group Information** field on the **Vol Grp Config.** tab of the Storage Management GUI.

12 In the **Add Volume Group** window click in the **Volume Group Path:** field.

- The cursor moves to the **Volume Group Path:** field.

13 Change the data entered at Step 9 to identify the full path for the storage of local backup data for the ESDT.

- This step is only for those ESDTs that require local backup.
- The typed entry appears in the **Volume Group Path:** field.

14 In the **Volume Group Type:** radio box, click on the **BACKUP** button.

- The button's depressed appearance indicates selection of **BACKUP**, signifying that the volume group being created is for storage for local backup use.

15 Click on the **Save and Add Next VG** button at the bottom of the **Add Volume Group** window.

- The volume group is created for display in the **Volume Group Information** field on the **Vol Grp Config.** tab of the Storage Management GUI.

16 In the **Add Volume Group** window click in the **Volume Group Path:** field.

- The cursor moves to the **Volume Group Path:** field.

17 Change the data entered at Step 13 to identify the full path for the creation and initial storage of offsite backup data for the ESDT.

- This step is only for those ESDTs that require offsite backup.
- The typed entry appears in the **Volume Group Path:** field.

18 In the **Volume Group Type:** radio box, click on the **OFFSITE** button.

- The button's depressed appearance indicates selection of **OFFSITE**, signifying that the volume group being created is for creation and initial storage for offsite backup use.

19 Click on the **Save and Exit** button at the bottom of the **Add Volume Group** window.

- The volume group is created for display in the **Volume Group Information** field on the **Vol Grp Config.** tab of the Storage Management GUI.
- Data stored in the volume group for OFFSITE backup can be safeguarded by removing tapes that have data stored on them and transporting the tapes to a secure offsite storage location.

Creating a Backup for AMASS

The Archive Manager should periodically create a backup to guard against loss of the AMASS database and functioning. The archive storage format used by AMASS is a proprietary format designed to optimize storage and retrieval speed. The command **vgexport -q** can be used to create a text file, storable on magnetic media, which can be used with the AMASS format archive tapes and the command **vgimport** to recover from the loss. This command exports the AMASS database for a specified volume group to standard out (**stdout**), a file containing the directory structure and media attributes (e.g., media type, ownership, timestamp) for the volume group. The file is located in **/usr/amass/filesysdb** and is exported as standard ASCII text. Use the following procedure.

Create a Backup for AMASS

1 Log in to the FSMS host (e.g., e0drg11, g0drg01, l0drg01, or n0drg01) as **amass** or **root**.

2 Type **/usr/amass/bin/vgexport –q** then press the **Return/Enter** key.

- A file named **stdout** is created in **/usr/amass/filesysdb**.
  - The **stdout** file is useful only with the archive volumes represented in the AMASS database.

Replacing the AMASS Database Backup Volume (Volume 1)

The AMASS database backup is stored in the archive on Volume 1. "Volume 1," hard coded to be the backup volume, actually designates one of the last volumes in the StorageTek LSM, to prevent its inadvertent use as a data volume. Whenever **amassbackup** is run, AMASS issues an e-mail message with information on volume capacity and usage. It is also possible to issue the command **vollist 1** to display how much space is left on the volume, or **volprint 1** for still more detail. If the volume becomes full *during* a backup attempt, the backup will fail and it is necessary to initialize a new backup volume and perform a full backup as described in the following procedure.

Replace a Full Backup Volume (Volume 1)

---

1 Log in to the FSMS host (e.g., e0drg11, g0drg01, l0drg01, or n0drg01) as **amass** or **root**.

2 Type **/usr/amass/bin/voloutlet 1** and then press the **Return/Enter** key.

- The LSM robot places the Backup Volume in the CAP.

3 Open the recessed latch on the CAP door; remove the Backup Volume tape and store it in a safe place.

4 Physically designate the new Backup Volume tape so that it can be easily discriminated from other volumes (e.g., write “Backup Volume” on the tape, color code the tape, or make and display a note of its home storage slot or preprinted barcode).

5 Note the pre-printed number on the volume label (e.g., 112102), insert the new Backup Volume in the CAP, and close the door.

- The robot scans the volume.

6 At the AMASS host, type **/usr/amass/bin/bulkinlet -u** and then press the **Return/Enter** key.

- AMASS assigns the Backup Volume a unique volume number.
- AMASS marks the volume **ONLINE** in the AMASS database.
- AMASS assigns the Backup Volume to the last barcode position in the library.
- AMASS gives the volume a **BACKUP VOLUME** label.

7 Type **/usr/amass/bin/vollist 1** and then press the **Return/Enter** key.

- AMASS displays the following:

VOL NUM	VOL GRP	JUKE NUM	POS	VOL LABEL	FLAGS	USED (MB)	AVAIL (MB)	DEAD (%)	ERRS
1	**0**	**1**		**BACKUP-VOLUME**	**I**	**0**	**20000**	**0**	**0**

8 To change the Volume Label field from **BACKUP-VOLUME** to the preprinted media number (e.g., 112102), type **/usr/amass/bin/vollabel 1 112102** and then press the **Return/Enter** key.

9 Type **/usr/amass/bin/vollist 1** and then press the **Return/Enter** key.

- AMASS displays the following:

VOL NUM	VOL GRP	JUKE NUM	POS	VOL LABEL	FLAGS	USED (MB)	AVAIL (MB)	DEAD (%)	ERRS
1	**0**	**1**		**112102**	**I**	**0**	**20000**	**0**	**0**

10 Type **/usr/amass/bin/volformat -u** and then press the **Return/Enter** key.

- A message requests confirmation that you wish to continue.

11 Type **y** and then press the **Return/Enter** key.

- A message is displayed requesting further confirmation, stating that **The following volumes will be formatted: 1 (Y-N)**.

12 Type **y** and then press the **Return/Enter** key.

- After a few minutes, a message **Completed formatting all volumes** is displayed.

13 To verify that the volume is inserted, type **/usr/amass/bin/vollist 1** and then press the **Return/Enter** key.

- Data for the media are displayed; the **flag** column shows that the newly formatted volume is inactive (**I**).

14 Type **/usr/amass/bin/amassbackup -fv** and then press the **Return/Enter** key.

- AMASS performs a full backup with the verbose option of the AMASS database and transaction logs.

Create Replacement Backups Manually from Existing Archives

If loss of data necessitates obtaining and inserting backup data from local or off-site storage, it is necessary to create replacement data to be returned to backup storage. Use the following procedure.

Create Replacement Backups

---

1 Log in to the FSMS host (e.g., e0drg11, g0drg01, l0drg01, or n0drg01) as **amass** or **root**.

2 Type **/usr/amass/bin/volcopy -c** <***source***> < ***destination*** > then press the **Return/Enter** key.

- <***destination***> is the volume number of the destination volume.
- <***source***> is the volume number of the source volume.
- The **-c** option specifies copy of the source to the destination.
- A bit for bit copy of the source (the cartridge to be copied) is made at the destination (an available, unused cartridge). Because the copy procedure depends on the amount of data on the source cartridge, the process can take as long as an hour to complete.

NOTE: After starting a **volcopy** procedure, do not attempt to kill the process with the **kill -9** command.

3 A hardcopy/softcopy list of the files backed up should be created and kept for future file restoration operations.

4 Remove the backup volume(s) and send to offsite storage area, as appropriate.

Restore Archive Data

Although Archive hardware has been selected for high reliability, there may be an occasional tape failure or drive error. In the event that such errors cause loss or corruption of the primary copy of Archive data, it will be necessary to attempt data recovery or to re-archive equivalent data. There are a number of potential means for restoring lost or corrupted data.

Use of Backup Data for Recovery

Depending upon the circumstances and nature of the loss, restoration of archive data may take the form of:

- copying from a local backup to the original or a new primary copy.
- reprocessing.
- obtaining replacement data from an external data provider.
- restoring the AMASS database and/or the ACSLS database.

In the event of catastrophic loss of the archive and local backup cartridges, it will be necessary to retrieve backups from off-site storage. A different process required to recover from a failed attempt at archive storage is addressed in a subsequent section on archive monitoring and fault notification.

Manual Data Recovery from Local or Offsite Backup Tapes

If a backup volume is available and contains the data that were lost or corrupted on the primary copy, the data can be copied using standard UNIX commands. Detailed procedures are available in Section 17.6.2 of Document 611-EMD-001, *Release 7.11 Mission Operation Procedures for the EMD Project*. If the backup volume must be obtained from offsite storage, it must then be inserted into the archive and brought on line. The procedures for loading archive media were addressed under a preceding topic. The requirements then entail:

- using the **Storage Config.** tab of the Storage Management GUIs to view the volume groups of the appropriate archive server and to find the files in the primary and backup volume groups.
- using the UNIX copy command (**cp** or **dd**) to copy the lost or corrupted file from the backup version to the primary version.
- as appropriate (i.e., if the recovery is one of a set of files to be restored, for example, because they were lost from a damaged tape), removing the names of the files recovered from the list of files to be recovered by other means.

If an entire volume is to be copied, perform the procedure to create replacement backups as addressed under a previous topic; if recovery is from offsite, send the backup back to secure offsite storage.

Reprocessing

If it is possible to produce a lost data product by running a Product Generation Executive (PGE) on other data available in the archive, recovery of the lost product may be achieved by this reprocessing. The reprocessing will be a case of on-demand processing, for which procedures are addressed in a separate lesson on Production Planning and Processing. In this case, the resultant recovered file will have a new Universal Reference (UR) and a new Production Date and Time. As part of the input to this process, the operator needs the following information for each file to be recovered:

- the Archive unique filename for the file.
- the Archive IDs of the primary archive and backup archive.
- the file checksum if available.

Requesting Replacement Data from Provider

Where the archived data that are lost are not available in local or off-site backups, but were originally acquired from an external data provider who retains a copy of the data in the archives of the data provider, recovery may be achieved by re-ingesting the data. Most Level 0 (L0) instrument data in the ECS archive was originally supplied by the EOS Data and Operations System (EDOS). Missing or corrupt ECS L0 files can usually be restored from the EDOS archive. If ECS operations at the Goddard Space Flight Center Earth Sciences DAAC requests a specific L0 product or products, EDOS can fulfill the request by supplying DTF-2 archive tape(s) containing the EDOS Production Data Sets (PDSs). For other DAACs, without DTF-2 tape drives available, EDOS supplies replacement PDSs using an FTP interface.

Requests for replacement L0 data from EDOS are made using an ESDIS-sponsored Data Reorder web tool. The tool is accessed from an ESDIS web page listing several links for reports and tools in the End-to-End Tracking System, as illustrated in Figure 21. The specific information to be entered using the tool in a reorder may vary somewhat depending on the platform and instrument involved in the initial capture of the data. Accordingly, as illustrated in the figure, there are separate links in the **Data Reorder** row of the Related Links table on the page.

A click on the appropriate Data Reorder link results in display of a page specific to the platform specified in the link. For example, a click on the **Terra** link in the Data Reorder row of the page illustrated in Figure 21 results in display of the page shown on the left side of Figure 22. A subsequent click on the **Add New Request** link displays the page shown on the right side of Figure 22, permitting specification and submission of the reorder request.

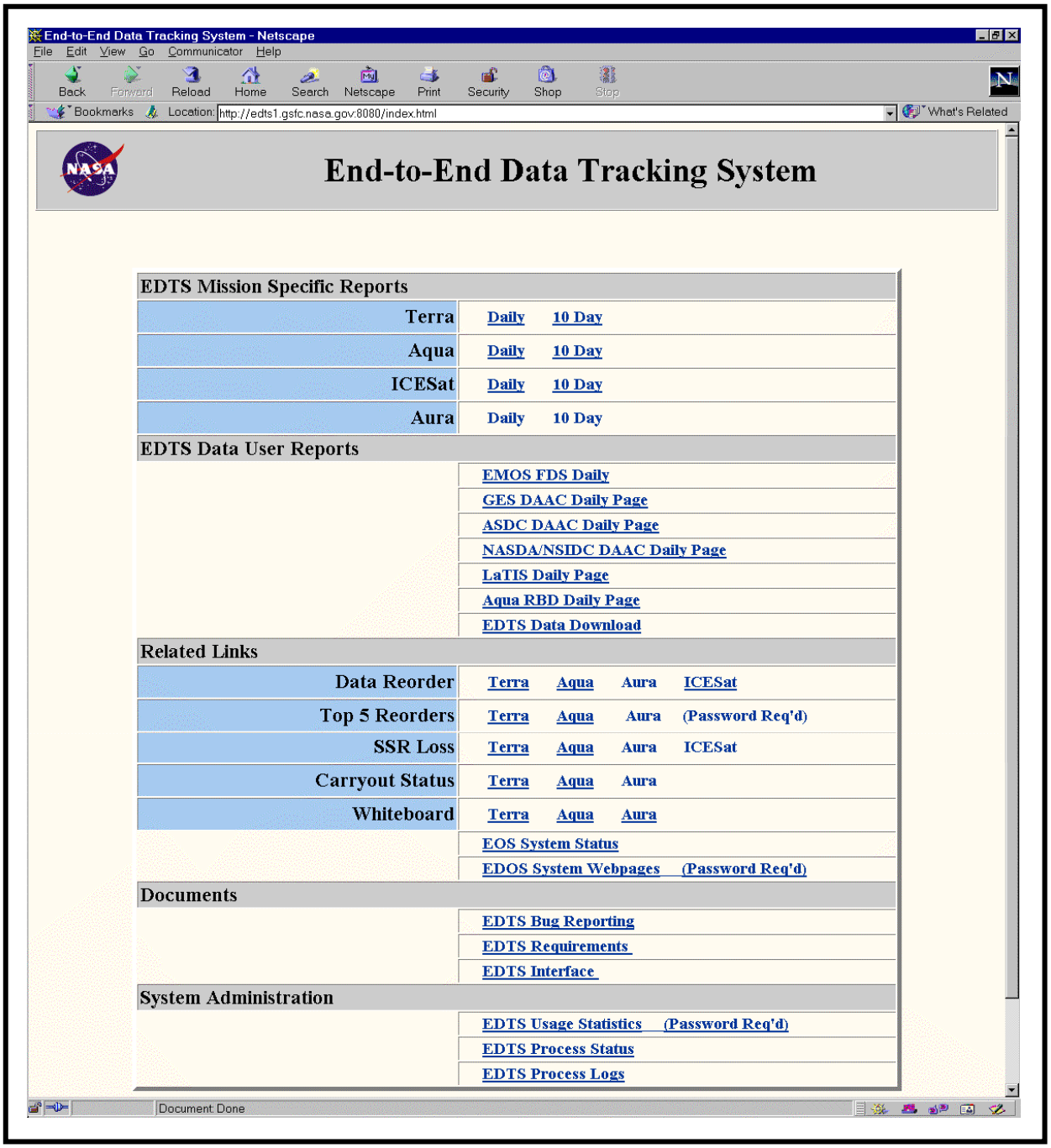


Figure 21. End-to-End Tracking System Web Page Providing Access to the Data Reorder Tool

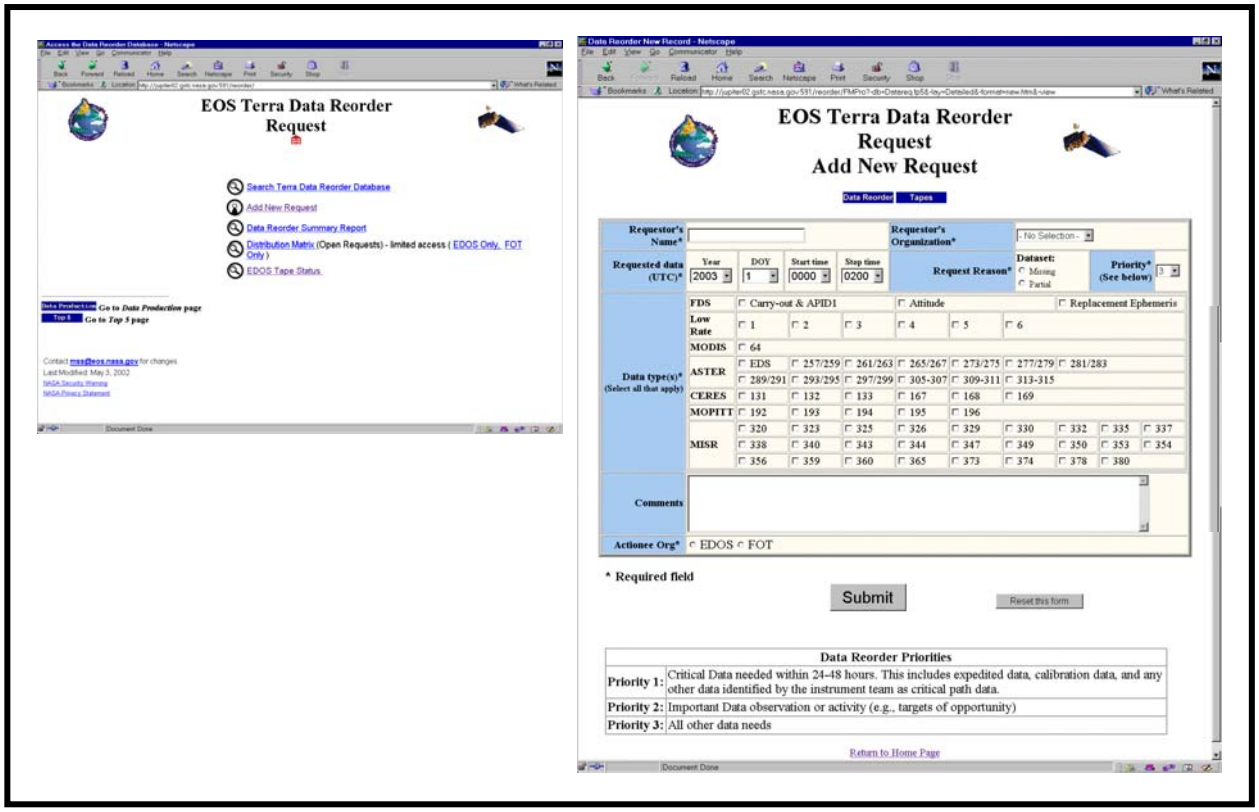


Figure 22. Add New Request Page for EOS Terra Data Reorder

Some data providers have decided not to support data re-supply; consult appropriate Interface Control Documents to determine suppliers able to re-supply data. The re-ingest may entail ingest procedures that are addressed in a separate lesson on Ingest, or procedures unique to ingest of replacement data (e.g., when replacement data from EDOS are provided on a tape with multiple files, only some of which are needed to replace a lost granule). As with re-processed data, the resultant recovered file will have a new UR.

The following procedure addresses re-ingest of lost data.

Re-Ingest Lost Data

1 Identify the source for each of the lost granules that were ingested.

2 If you have not already done so, retrieve the file location metadata for each file (see Document 611-EMD-001, *Release 7.11 Mission Operation Procedures for the EMD Project*, Procedure 17.6.2.5.1, **SDSRV Retrieval of File Location Metadata**).

NOTE: Except for Step 48, the remaining steps may be performed by the Ingest Technician.

NOTE: Some data suppliers do not support re-supply of data.

3 With reference to the applicable Interface Control Document (ICD) and using the granule metadata retrieved in Step 2, initiate the required data re-supply requests as defined in the ICD for those data suppliers able to re-supply data.

- To re-order Level 0 production data sets (PDSs) from EDOS, the DAACS use the ESDIS-sponsored EOS Data Reorder Web Tool. Steps 4 through 26 address use of the tool to submit a re-order request.
- EDOS furnishes L0 replacement data to the GSFC Earth Sciences (GES) DAAC on DTF-2 tapes. A tape may contain multiple granules and files, a subset of which is needed to replace the lost granule(s). Steps 27 through 54 address recovery of the lost data.
- EDOS furnishes L0 replacement data to other DAACs that do not have DTF-2 tape drives. In this case, EDOS transfers the necessary PDS(s) to ECS in the automated Ingest process for polling with delivery record, monitored by the Ingest Technician (see Document 611-EMD-001, *Release 7.11 Mission Operation Procedures for the EMD Project*, Procedure 16.2.5, **Monitor/Control Ingest Requests**). The re-order is accomplished as specified in Steps 4 through 26; the only other necessary steps in this procedure are 45, 48, and 49.

NOTE: Steps 4 through 8 are to access and launch browser software.

4 At the UNIX command shell prompt, type **setenv DISPLAY *clientname*:0.0** and then press the **Return/Enter** key.

- For ***clientname***, use either the local terminal/workstation IP address or its machine name.

5 Start the log-in to a Netscape host by typing **/tools/bin/ssh *hostname*** (e.g., e0ins01, g0ins01, l0ins01, or n0ins02) at the UNIX command shell prompt, and press the **Return/Enter** key.

- 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)?** type **yes** ("**y**" alone does 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 6.
- If you have not previously set up a secure shell passphrase, go to Step 7.

6 If a prompt to **Enter passphrase for RSA key '<*user@localhost*>'** appears, type your ***Passphrase*** and then press the **Return/Enter** key. Go to Step 8.

7 At the ***<user@remotehost>*'s password:** prompt, type your ***Password*** and then press the **Return/Enter** key.

- You are logged in and a UNIX command shell prompt is displayed.

8 Type **netscape &** and then press the **Return/Enter** key.

- **The Netscape web browser is displayed.**

9 Click in the **Netsite:** or **Location:** field.

- The field is highlighted.

10 Type the Universal Resource Locator (URL) for the End-to-End Data Tracking System (e.g., http://edts1.gsfc.nasa.gov:8080/index.html) and then press the **Return/Enter** key.

- The **End-to-End Data Tracking System** index page is displayed, offering access to **various reports and related links.**

11 Under **Related Links** click on the appropriate Data Reorder link.

- There are links for various satellite platforms (e.g., Terra, Aqua, or Aura).
- The appropriate **Data Reorder Request** page for the selected Data Reorder link is displayed.

12 On the Data Reorder Request page click on the **Add New Request** link.

- The **Add New Request** page is displayed.

13 On the **Add New Request** page click in the **Requestor's Name** field.

- The cursor is displayed in the field.

14 Type the name of the person making the request.

- The typed entry is displayed in the field.

15 Single-click on the pull-down arrow at the end of the **Requestor's Organization** field and then single-click on the name of the requesting organization to select it, or, if the requesting organization is not displayed in the pull-down menu, select **Other**.

- The choices are **LaRC DAAC**, **LaTIS**, **GSFC DAAC**, **EDC**, **EDOS**, **FOT**, **ESDIS**, **ASTER GDS**, and **Other**.
- The selected choice is displayed in the field.

16 In the **Requested Data (UTC)** block of the page single-click on the pull-down arrow at the end of the **Year** field and then single-click on the year for the missing data to select it.

- The selected choice is displayed in the field.

17 In the **Requested Data (UTC)** block of the page, single-click on the pull-down arrow at the end of the **DOY** (Day of Year) field and then single-click on the day of the year for the missing data, first scrolling with the scroll bar if necessary to display the desired day.

- The selected choice is displayed in the field.

18 In the **Requested Data (UTC)** block of the page, single-click on the pull-down arrow at the end of the **Start time** field and then single-click on the hour representing the start of a two-hour time window for the missing data.

- The choices are in even two-hour time intervals beginning with **0000** and proceeding to **2200**.
- The selected choice is displayed in the field.

19 In the **Requested Data (UTC)** block of the page, single-click on the pull-down arrow at the end of the **Stop time** field and then single-click on the hour representing the end of a two-hour time window for the missing data.

- The choices are in even two-hour time intervals beginning with **0200** and proceeding to **2400**.
- The selected choice is displayed in the field.

20 In the **Request Reason** block of the page click on the appropriate radio button to indicate the reason for the request, specifying that the dataset is **Missing** or **Partial**.

- The selected button is filled to indicate its selection.

21 For any data priority other than **3**, click on the pull-down arrow at the end of the **Priority** field and then click on the appropriate priority for the request.

- The choices are **1** (critical data needed within 24-48 hours), **2** (important data observation or activity, such as a target of opportunity), and **3** (all other data needs). The default is **3**, and this step may be skipped if that is the priority for the request.
- The selected choice is displayed in the field.

22 Click in the applicable check boxes in the **Data type(s)** block of the page to specify the desired **FDS** or Flight Dynamics System information (Carry-out and APID1 or satellite housekeeping data, Attitude, and Replacement Ephemeris), **Low Rate** information, and instrument (e.g., **MODIS**, **ASTER**, **CERES**, **MOPITT**, and **MISR**) application process identifiers (APIDs).

- The selected check boxes each display a checkmark to indicate selection.

23 If it is desirable to enter any comments concerning the request, click in the **Comments** field; otherwise, go to Step 25.

- The cursor is displayed in the field.

24 Type any comments to be submitted with the request.

- The typed entry is displayed in the field.

25 In the **Actionee Org** block of the page click on the appropriate radio button to identify the actionee for the request, specifying **EDOS** or **FOT**.

- The selected button is filled to indicate the selection (in this case, **EDOS**).

26 Click on the **Submit** button at the bottom of the form.

- Following confirmation, the request submittal is acknowledged with a request ID.

27 If the replacement data are on a DTF-2 tape from EDOS, load the tape into a DTF-2 drive, using Document 611-EMD-001, *Release 7.11 Mission Operation Procedures for the EMD Project*, Procedure 16.3.3.4, **Perform DTF-2 Drive Loading**.

- The tape is loaded.

28 Access a terminal window logged in to the appropriate host (e.g., Sun internal server).

- Examples of Sun internal server host names include **e0acs11**, **g0acs11**, **l0acs03**, and **n0acs04**.

29 Type **cd *path*** and then press the **Return/Enter** key.

- ***path*** represents the directory path to the location where the data from the EDOS archive tape should be copied.
- Using an empty directory would help identify the data from the tape.

30 Type **tar xvf *device*** and then press the **Return/Enter** key.

- ***device*** is the DTF-2 drive device name (e.g., /dev/rmt/2n) as it is known to the shell.
- For example:

  tar xvf /dev/rmt/2n

- As files are read from the tape the file names, file sizes (in bytes), and number of blocks are listed on the screen.
  - For example:

    x DZ9ZA49.MDR, 17393 bytes, 34 tape blocks

31 Type **pg *PPMUDR_name*** and then press the **Return/Enter** key.

- ***PPMUDR_name*** represents the file name of the PDS Physical Media Unit Delivery Record (PPMUDR).
  - The PPMUDR file name has a **.MDR** extension.
  - The PPMUDR is the first item on the EDOS archive tape.

- For example:

  pg DZ9ZA49.MDR

- Although this procedure has been written for the **pg** command, any UNIX editor or visualizing command (e.g., **vi**, **view, more**) can be used to review the log file.

32 Observe the contents of the PPMUDR to identify the PDS(s) to be archived.

- Packet date/time ranges in the PPMUDR can be used to determine which PDS(s) is (are) to be archived.
  - In the PPMUDR the PDSs on the tape are listed in file groups, which represent data sets [i.e., science data file(s) and corresponding metadata file].
  - Each file group (data set) includes the date/time range of the data specified as FIRST_PACKET_TIME and LAST_PACKET_TIME.
- For example (extract from a PPMUDR):

```
OBJECT = FILE_GROUP
        DATA_SET_ID = P0420064AAAAAAAAAAAAAA03101231459600
        DATA_TYPE = MOD000
        FIRST_PACKET_TIME = 2003-04-10T00:00:00.000000Z
        LAST_PACKET_TIME = 2003-04-10T01:59:59.999999Z
        PACKET_COUNT = NOT USED
        OCTET_COUNT = NOT USED
        TEST_FLAG = F
        APID_COUNT = 1
        OBJECT = APID_SPEC
            APID_IN_PDS = 64
        END_OBJECT = APID_SPEC
        FILE_COUNT = 2
        OBJECT  = FILE_SPEC
                DIRECTORY_ID = NOT USED
                FILE_ID = P0420064AAAAAAAAAAAAAA03101231459600.PDS
                FILE_TYPE = METADATA
                FILE_SIZE = 384
        END_OBJECT = FILE_SPEC
```

OBJECT = FILE_SPEC

DIRECTORY_ID = NOT USED

FILE_ID = P0420064AAAAAAAAAAAAAA03101231459601.PDS

FILE_TYPE = DATA

FILE_SIZE = 108000

END_OBJECT = FILE_SPEC

END_OBJECT = FILE_GROUP

- In the preceding example one data set is defined (as a "FILE_GROUP").
  - The data type for the set is MOD000.
  - The data were collected on April 10, 2003 between midnight GMT (00:00:00.000000Z) (FIRST_PACKET_TIME) and just before 2:00 A.M. GMT (01:59:59.999999Z) (LAST_PACKET_TIME).
  - There are two files in the data set (FILE_COUNT = 2).
  - One file (P0420064AAAAAAAAAAAAAA03101231459600.PDS) is a metadata file (in EDOS terminology, a "construction record").
  - The other file (P0420064AAAAAAAAAAAAAA03101231459601.PDS) is a data file.
  - Based on information embedded in the file names, the data set was created on April 11, 2003 at 11:14:59 P.M. (as described under the next bullet).

- The EDOS archive tape may contain both nominal and reprocessed PDSs but creation times in file names differentiate between the versions.
  - Ingest the latest (most recent) version if there is more than one version.
  - PDS file names consist of 40 bytes (characters) and Bytes 23 through 33 specify the creation time of the file.
  - For example, **03101231459** is the creation time in the following file name:

    P0420064AAAAAAAAAAAAAA03101231459601.PDS

    - **03** indicates the year (2003).
    - **101** specifies the Julian day (April 11, the 101st day of the year).
    - **231459** is the time of file creation (11:14:59 P.M.).

- It is the Archive Manager's responsibility to resolve any questions concerning which PDSs should be archived (see Step 2).

33 Type **cp** ***filename1 filename2 [… filenameN] path*** and then press the **Return/Enter** key.

- ***filename1 filename2 [… filenameN*** represent the file names of the PDS files to be ingested.
  - Copy both the data and metadata files (as identified in the PPMUDR) for each data set.
- ***path*** is the directory path to the Ingest pollEDOS directory; i.e., the directory in which the ECS software for EDOS ingest routinely looks for EDOS delivery records and data.
  - The EDOS polling directory is specified as a parameter in the Registry database or in the configuration file for EDOS polling (e.g., EcInPolling.EDOS.CFG).
- For example:

  cp P0420064AAAAAAAAAAAAAA03101231459600.PDS P0420064AAAAAAAAAAAAAA03101231459601.PDS /usr/ecs/OPS/CUSTOM/icl/x0icg01/data/pollEDOS

NOTE: If a DAAC-unique script is available for creating delivery records and signal files and placing the files in the polling directory, use the script and skip Steps 34 through 44 (go to Step 45 after running the script). Otherwise, manually generate delivery records and signal files as described in Steps 34 through 44.

34 Type **cd** ***path*** and then press the **Return/Enter** key.

- ***path*** is the directory path to the Ingest pollEDOS directory.
- For example:

  cd /usr/ecs/OPS/CUSTOM/icl/x0icg01/data/pollEDOS

NOTE: Steps 35 through 39 describe how to use an old delivery record (PDR) as a template for generating a new PDR.

35 Type **cp** ***old_PDR_filename new PDR filename*** and then press the **Return/Enter** key.

- ***old_PDR_filename*** represents the file name of an old PDR that is being used as a template for creating a PDR for PDS files to be ingested.
- ***new_PDR_filename*** represents the file name of the new PDR that is being created for PDS files to be ingested.
  - Use the EDOS file-naming convention for PDRs (refer to the EDOS ICD, 423-ICD-EDOS/EGS):
    - PDR file names consist of 38 bytes (characters).
    - Byte 1 identifies the file as either a PDS Delivery Record (“X”) or EDS Delivery Record (“Y”).

- Bytes 2 through 8 identify the spacecraft ID (SCID) (three bytes) and first Applications Process Identifier (APID) (four bytes) in the data set (right-justified and, if necessary, zero-filled on the left).
- Bytes 9 through 15 identify the SCID and second APID in the data set (right-justified and, if necessary, zero-filled on the left), if applicable. If no second APID is present in the data set, this item has a value of "AAAAAAA".
- Bytes 16 through 22 identify the SCID and third APID in the data set (right-justified and, if necessary, zero-filled on the left), if applicable. If no second APID is present in the data set, this item has a value of "AAAAAAA".
- Bytes 23 through 33 identify the GMT/ZULU time when the data set was created.
- Byte 34 is a numeric identification in the range of "0" to "9" to aid in distinguishing the order of data set creation during the day and to provide uniqueness to the file name.
- Bytes 35 through 38 are the file name extension (i.e., ".PDR" or ".EDR")
- For example:

  X0420064AAAAAAAAAAAAAA031012314596.PDR

  X identifies the file as a PDS Delivery Record.

  0420064 identifies the SCID (**042** = Terra) and first APID (**0064** = MOD000 data type) in the data set.

  AAAAAAA indicates that there is no second APID in the data set.

  AAAAAAA indicates that there is no third APID in the data set.

  03101231459 is the GMT/ZULU time when the data set was created [**03** indicates the year (2003); **101** specifies the Julian day (April 11, the 101st day of the year); **231459** is the time of data set creation (11:14:59 P.M.)].

  6 is a numeric identifier (sixth data set of the day).

  .PDR is the file-name extension for a PDS Delivery Record.

36 Type **vi** ***new_PDR_filename*** and then press the **Return/Enter** key.

- The PDR template file is opened (displayed by the vi text editor).
- Although this procedure has been written for the **vi** editor, any UNIX editor can be used to create the PDR.

37 Using vi editor commands modify the PDR file to specify ingest of one of the data sets to be ingested.

- Create a separate PDR for each data set [science data file(s) and corresponding metadata file – refer to the PPUDR "file group" example in Step 32].
- The following vi editor commands are useful:
  - **h** (move cursor left).
  - **j** (move cursor down).
  - **k** (move cursor up).
  - **l** (move cursor right).
  - **a** (append text).
  - **i** (insert text).
  - **r** (replace single character).
  - **x** (delete a character).
  - **dw** (delete a word).
  - **dd** (delete a line).
  - ***n*dd** (delete *n* lines).
  - **u** (undo previous change).
  - **Esc** (switch to command mode).

38 Press the **Esc** key.

39 Type **ZZ**.

- New PDR file is saved.
- UNIX prompt is displayed.

40 Type **vi *XFR_filename*** and then press the **Return/Enter** key.

- A new file with the specified ***XFR_filename*** is opened.
  - Use the EDOS file-naming convention for signal files (refer to the EDOS ICD, 423-ICD-EDOS/EGS):
    - Signal file name is the corresponding PDR file name plus the signal file name extension (i.e., ".XFR").
    - For example:

      X0420064AAAAAAAAAAAAAA031012314596.PDR.XFR

- The signal file indicates that the relevant data files and PDR have been put in the polling directory and are ready to be ingested.
- Although this procedure has been written for the **vi** editor, any UNIX editor can be used to create the signal file.

41 Using vi editor commands create a file that contains the name of the relevant PDR.

- A signal file contains the name of the relevant PDR only.
- For example:

  X0420064AAAAAAAAAAAAAA031012314596.PDR.

42 Press the **Esc** key.

43 Type **ZZ**.

- New signal file is saved.
- UNIX prompt is displayed.
- At the next polling occasion, the EDOS polling client should detect the signal file and initiate ingest of the data specified in the corresponding PDR.

44 Repeat Steps 34 through 43 as required to create delivery records and signal files for all remaining data sets (from the EDOS archive tape) to be ingested.

45 To monitor Ingest request processing (polling with delivery record), perform the procedure for **Monitoring/Controlling Ingest Requests** (see Document 611-EMD-001, *Release 7.11 Mission Operation Procedures for the EMD Project*, Procedure 16.2.5, **Monitor/Control Ingest Requests**).

46 Remove the EDOS-provided tape from the DTF-2 drive, using Document 611-EMD-001, *Release 7.11 Mission Operation Procedures for the EMD Project*, Procedure 16.3.3.5, **Perform DTF-2 Drive Unloading**.

47 Verify that the data have been inserted into the archive as described in the procedure for **Verifying the Archiving of Ingested Data** (see Document 611-EMD-001, *Release 7.11 Mission Operation Procedures for the EMD Project*, Procedure 16.2.10**, Verify the Archiving of Ingested Data**).

48 When insertion into the archive has been verified, the Archive Manager specifies "set delete" for the replaced data/metadata by using procedures for granule deletion to mark the data/metadata for deletion from the archive (see Document 611-EMD-001, *Release 7.11 Mission Operation Procedures for the EMD Project*, Section 17.4.2, **Selecting Granules for Deletion** and Procedure 17.4.2.3, **Selection Using a Separate Input File**).

49 When insertion into the archive has been verified, ensure that the EDOS archive tape is returned to the EDOS Level 0 Processing Facility (LZPF).

NOTE: Clean up (as described in Steps 50 through 54) the directory into which data were originally copied from the EDOS archive tape. If preferred, skip Steps 50 through 54 and use the script described in the procedure **Clean the Polling Directories** (see Document 611-EMD-001, *Release 7.11 Mission Operation Procedures for the EMD Project*, Procedure 16.2.11, **Clean the Polling Directories**).

50 Type **cd *path*** and then press the **Return/Enter** key.

- ***path*** represents the directory path to the location where the data from the EDOS archive tape were first copied.

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

- A listing of the files in the current directory is displayed.

52 Type **rm *filename1 filename2 [… filenameN]*** and then press the **Return/Enter** key.

- ***filename1 filename2 [… filenameN]*** represent the names of the files to be removed from the directory.
- A wildcard may be used if some of the files have common characteristics.
  - For example:

    rm *.PDS
- A prompt is displayed requesting whether or not a particular file should be removed.
  - For example:

    rm: remove DZ9ZA49.MDR (yes/no)?

53 Type **y** and then press the **Return/Enter** key.

- The specified file is deleted and (if applicable) a prompt is displayed requesting whether or not another particular file should be removed.

54 Repeat Step 53 as necessary.

Restoring the AMASS Database

The AMASS database is restored manually by the System Administrator or the Archive Manager using the AMASS command **amassrestore**. This command restores the last full backup, the last partial backup, and all journal transactions that have occurred since the last backup. It creates a sub-directory under filesysdb called **journal**. All restored files are copied to the **journal** directory. The following restore procedure uses a backup volume or tape device.

Restore the AMASS Database

---

1 Log in to the FSMS host (e.g., e0drg11, g0drg01, l0drg01, or n0drg01) as **amass** or **root**.

Caution

Do not use the **amassrestore** command when AMASS is running. To shutdown AMASS, refer to the Special Shutdown Procedures in the AMASS technical documentation *Installing AMASS*.

2 To inactivate the AMASS file system, type **/usr/amass/bin/amassstat -i** and then press the **Return/Enter** key.

- The AMASS file system is inactivated.

3 Make sure the backup drive is available.

- If there is another volume in the drive, return it to its home slot by entering **/usr/amass/daemons/amassrecovery -s** (the option **-s** prevents system startup and performs file recovery).

4 Type **/usr/amass/bin/amassrestore -v -L <*barcodelabel*>** and then press the **Return/Enter** key.

- If you do not know the barcode label number for the backup volume, it can be obtained by entering **/usr/amass/bin/vollist 1**.
- The AMASS database is restored from the backup volume.

Restoring the ACSLS Database

ACSLS provides the **rdb.acsss** utility to restore the database in case of severe disk or data problems. If you have made regular backups, it should be possible to restore the database with little or no loss of data. Restoring the database is likely to be necessary if there has been a system crash, or if the database cannot be started or has a physical or logical error. The following procedure is applicable.

Restore the ACSLS Database

---

1 At the host for ACSLS (e.g., **e0drs03**, **g0drs04**, **l0drs02**, **n0drs03**), log in using the **acsss** user ID and password.

- The **acsss** command-process window is displayed with the **ACSSS>** prompt.

2 Load the restore tape into the backup drive.

3 Type **rdb.acsss**, and then press the **Return/Enter** key.

- If you enter **bdb.acsss** with no options, the backup utility defaults to the default tape device attached and configured to the ACSLS server.

- The system displays the following message.

  Check tape device (/dev/rmt/0) to make sure you have a tape in the tape drive.
  [Hit RETURN to continue or Ctrl-C to exit]

4 Press the **Return/Enter** key.

- The rdb.acsss utility restores the ACSLS database and miscellaneous library resource files.

This page intentionally left blank.

AMASS Graphical User Interface

Using the AMASS GUI

AMASS offers a Graphical User Interface (GUI) called the AMASS Administration Window (AAWin) through which operators can administer volumes and volume groups that are managed by AMASS. AAWin provides a point-and-click interface for identifying volumes their groups, and their configurable parameters. Figure 23 shows the AAWin main window, which is composed of a menu bar, a large middle section called the *workroom*, a utility bar at the right with icons for a trash can, a volume group, and a volume, and a status bar at the bottom with indicator "lights" that represent the current status of AMASS. The figure shows how the window looks when the volume icon on the utility bar has been selected to populate the workroom with icons for volumes, and illustrates the type of volume-related information that appears in a pop-up display as the cursor is moved over one of the icons.

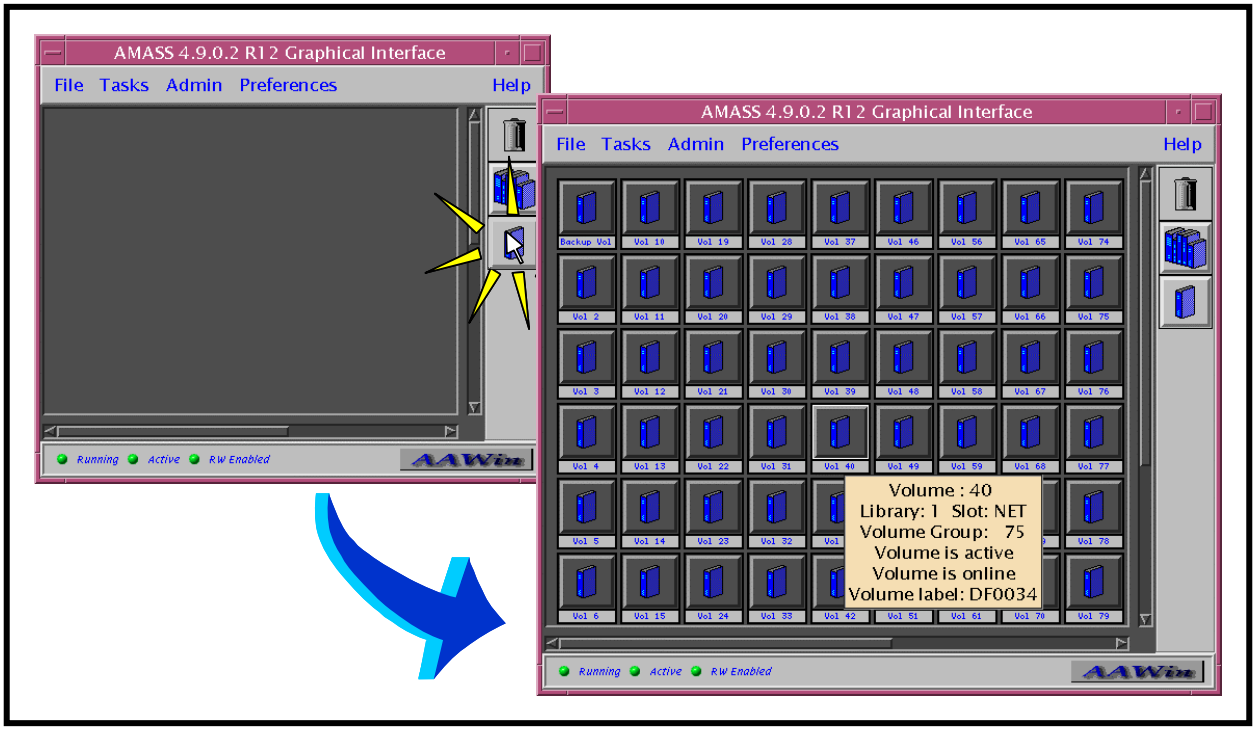


Figure 23. AMASS AAWin Main Window

Use the following procedure to launch the AMASS GUI and view information about volume groups and volumes in the archive.

View Volume Group and Volume Information Using the AMASS GUI

---

1 Log in as **amass** at the FSMS host (e.g., e0drg11, g0drg01, l0drg01, or n0drg01).

2 Type **/usr/amass/bin/aawin** and then press the **Return/Enter** key.

- The AMASS GUI main window is displayed.

3 Click on the **View by Volume Groups** button (middle button at the right of the *workroom*).

- The *workroom* is populated by icons for volume groups.
- The **Block List** window is displayed; it is a vertically scrolled list of blocks of items (in this case, volume groups). The *workroom* can display up to 256 icons; the **Block List** window provides access to additional items in blocks of 256.

4 Move the cursor over one of the icons for a volume group.

- A pop-up display shows data for the volume group (**Volume Group**, **Volumes in Group**, **Free Space**, **Dead Space**, **Error Count**).

5 Click on the **View by Volumes** button (at the bottom right side of the *workroom*).

- The *workroom* is populated by icons for volumes.
- The **Block List** window is also displayed; it is a vertically scrolled list of blocks of items (in this case, volumes).

6 Move the cursor over one of the icons for a volume.

- A pop-up display shows data for the volume group (**Volume**, **Library**, **Slot**, **Volume Group**, **Volume Status**, **Volume Label**).

Modify a Volume Group

Figure 24 shows the *Modify a VG* window. This window is opened by selecting *Modify a Volume Group* from the *Tasks* menu. The window is used to modify the characteristics of a volume group. The top portion of the window (not modifiable) lists root directories already configured for a volume group. The middle portion of the window permits adding directories to the list of root directories for the specified volume group. The third major portion of the window, near the bottom, contains indicators of the status of the volume group and buttons for selecting a volume group, as well as buttons across the very bottom of the window for accepting or canceling the modifications.

NOTE: The *Modify a VG* window also is opened if you have the *workroom* populated with volume group icons and you click on one of them. However, in this case you may only modify the volume group on which you clicked; the bottom of the window will not display buttons for selecting a volume group.)

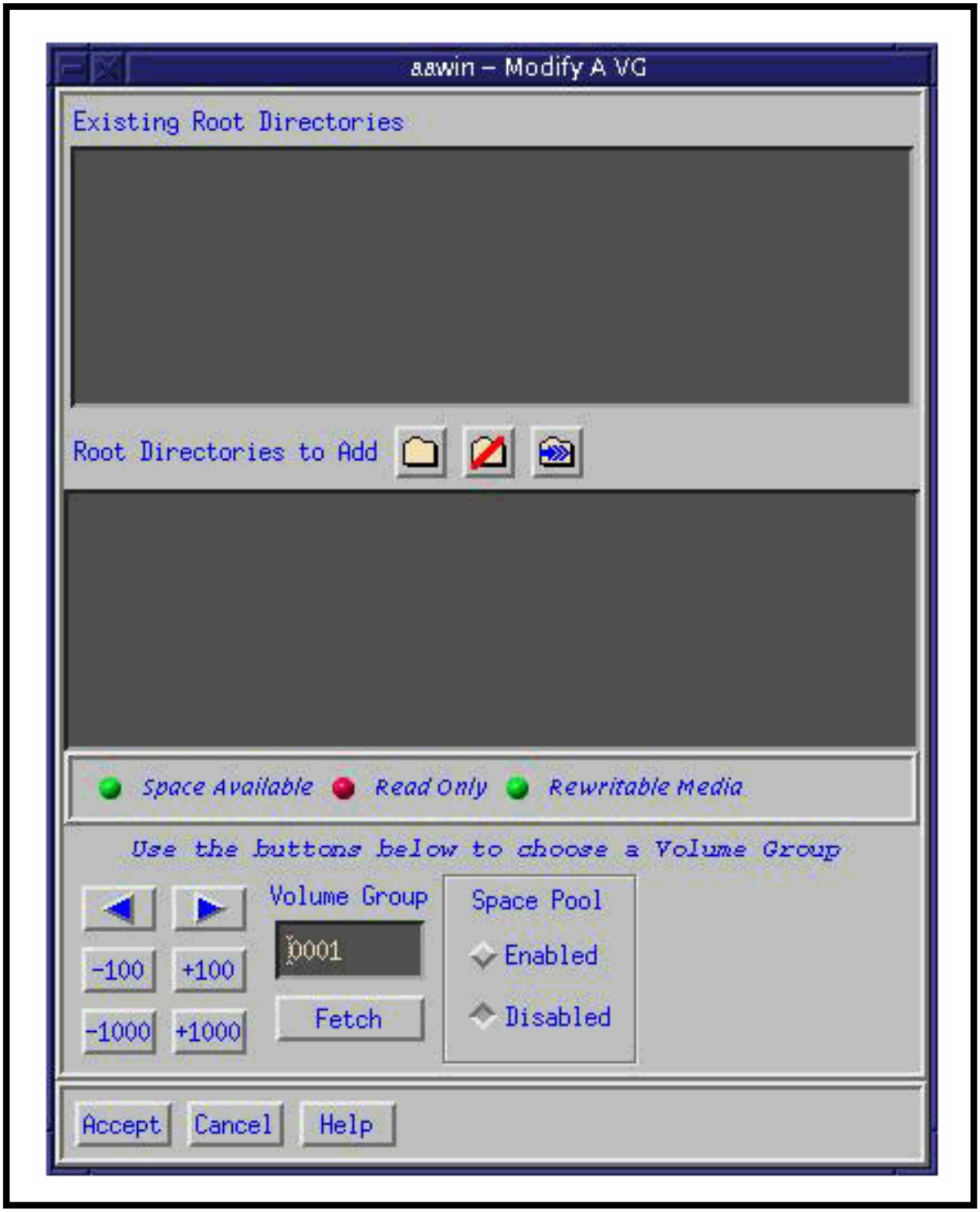


Figure 24. Modify a VG Screen of the AMASS GUI

To see how a volume group can be modified, lets examine how you might assign a new root directory in the AMASS file system to a volume group. Use the following procedure to create the directory and then modify a volume group.

Modify a Volume Group

---

1 Open a second terminal window (other than the one used to launch the AMASS GUI).

NOTE: This procedure assumes that the AMASS GUI is open after previously being launched from a terminal window.

2 In the second terminal window, log in as **amass** at the FSMS host (e.g., **e0drg11**, **g0drg01**, **l0drg01**, or **n0drg01).**

3 To change to the **dss_amass** directory, type **cd /dss_amass** and then press the **Return/Enter** key.

- The working directory is changed to **/dss_amass**.

4 To create an empty directory with path **/dss_amass/training/** to assign to the volume group, type **mkdir training** and then press the **Return/Enter** key.

5 On the AMASS GUI main window, click on the **View by Volume Groups** button (middle button at the right of the *workroom*).

- The *workroom* is populated by icons for volume groups.
- The **Block List** window is also displayed; it is a vertically scrolled list of blocks of items (in this case, volume groups).

6 Follow menu path **Tasks→Modify a Volume Group**.

- The **Modify a VG** window is displayed, showing data for Volume Group 0001.

7 In the area for choosing a volume group, near the bottom of the window, use the buttons to set the number displayed in the **Volume Group** field to the desired volume group.

- A click on the right-pointing arrow button or the left-pointing arrow button respectively increases or decreases the number by one. Buttons below the arrow buttons may be used to increase or decrease the number in multiples of 100 or 1000, as indicated on the buttons.

8 When the **Volume Group** field displays the number of the desired volume group, click on the **Fetch** button.

- The list of root directories already configured for the selected volume group is displayed in the **Existing Root Directories** field.
- The status indicators show the status of the selected volume group.

9 Click on the **File/Directory Selection** button (leftmost button after the label **Root Directories to Add**, with folder icon).

- A **File Selection** filter window is displayed.

10 In the **File Selection** filter window, click on the **Filter** button.

- The **Filter** field displays **/usr/amass/***, and directories and files are displayed in the **Directories** and **Files** windows, respectively.

11 Use the **Filter** button and selection of directories in the **Directories** window to display **/dss_amass/training/** in the **Selection** field.

- The **Selection** field displays **/dss_amass/training/**.

12 In the **File Selection** filter window, click the **OK** button.

- The **Root Directories to Add** field of the **Modify a VG** window displays **/dss_amass/training/**.

13 To examine the capability to edit the list of Directories to Add, click on the entry **/dss_amass/training/** to highlight it in the **Root Directories to Add**, then click on the **Remove a File/Directory from List** button (middle button after the label **Root Directories to Add**, with folder icon crossed out with a red line).

- The entry **/dss_amass/training/** is removed from the **Root Directories to Add** field.

14 Repeat steps 10 - 12 to restore the entry **/dss_amass/training/** to the **Root Directories to Add** field.

- The **Root Directories to Add** field of the **Modify a VG** window displays **/dss_amass/training/**.

15 In the **Modify a VG** window click on the **Accept** button at the bottom of the window.

- The entry **/dss_amass/training/** is removed from the **Root Directories to Add** field and appears in the **Existing Root Directories** field.
- The **Modify a VG** window is closed.

Modify a Volume

Figure 25 shows the *Modify a Volume* window. This window is opened by selecting *Modify a Volume* from the *Tasks* menu. The window is used to modify the characteristics of a volume. The right side of the window shows the current set of statistics and configuration information (not modifiable) for the volume listed in the **Volume** field on the left side of the window (the **Volume** field looks like a button, but if you click on it, a “spinbox” is displayed, with arrow buttons permitting increases or decreases to the volume number, and buttons at the bottom to **Accept** or **Cancel** the change; accepting the change closes the spinbox, displays the new number in the **Volume** field, and displays data for that volume). The left side of the *Modify a Volume*

window provides access to modifiable characteristics of the volume. Changes made to the buttons and fields in the window do not take effect until the **Accept** button at the bottom of the window is clicked.

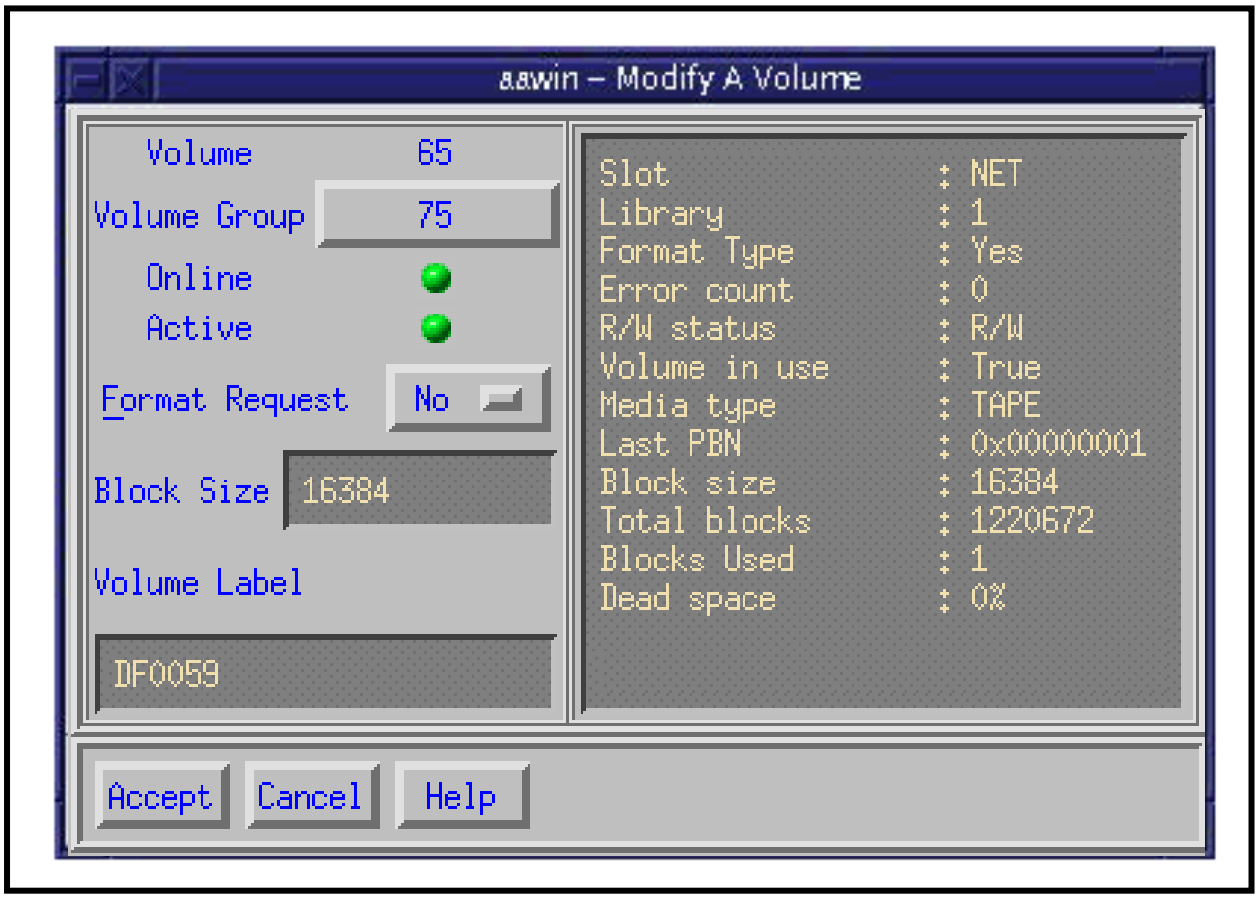


Figure 25. Modify a Volume Window of the AMASS GUI

NOTE: The *Modify a Volume* window also is opened if you have the *workroom* populated with volume icons and you click on one of them. However, in this case you may only modify the volume on which you clicked; the **Volume** field does not look like a button and may not be changed.)

There are six fields that can be edited for a given volume:

- The first is a button for setting the **Volume Group**. Clicking the button opens a spinbox for selecting the volume group to which the volume is to be assigned.
- Below the Volume Group button is an **Online/Offline** indicator light with label. Clicking on the indicator toggles its state and updates the text field (label) next to it.
- Below the Online/Offline indicator is an **Active/Inactive** indicator light with label. Clicking on the indicator toggles its state and updates the text field (label) next to it.

- Next is a **Format Request** option button permitting selection of a formatting option for the volume.
- Next is the **Block Size** field, applicable only to tape libraries when a format is requested to be done on the volume. This field requires a numeric value, which should be a multiple of 16384.
- The last modifiable field is a text field for specifying the volume label.

To examine the functioning of the *Modify a Volume* window, use the following procedure.

Modify a Volume

---

1 On the AMASS GUI main window, click on the **View by Volume Group** button (middle button at the right of the *workroom*).

- The *workroom* is populated by icons for volume groups.
- The **Block List** window is also displayed; it is a vertically scrolled list of blocks of items (in this case, volume groups).

2 Click on the icon for a high-numbered volume group with several volumes in it.

- The *workroom* is populated with icons for the volumes in the selected volume group, and the **Modify a VG** window is displayed, showing data for the selected volume group.

3 Click on the icon for one of the volumes.

- The **Modify a VG** window is closed and the **Modify a Volume** window is displayed, showing data for the selected volume.

4 To change the volume group to which the volume is to be assigned, note the **Volume Group** number indicated on the **Volume Group** button, and then click on the button.

- A spinbox is displayed showing the **Volume Group** number, with right-pointing and left-pointing arrow buttons respectively to increase or decrease the number.

5 Use the arrow buttons to change the **Volume Group** number, and then click on the **Accept** button in the spinbox.

- The spinbox is closed and the new number appears in the **Modify a Volume** window as the **Volume Group** number.

6 Return the **Volume Group** number to its original value by repeating steps 4 and 5, using the spinbox to set the number to that which you noted originally.

- The spinbox is closed and the original number appears in the **Modify a Volume** window as the **Volume Group** number.

7 Experiment with the **Online/Offline** and **Active/Inactive** indicators.

- When the volume is indicated to be **Online**, clicks on the **Active/Inactive** indicator toggle the color and label for the indicator.
- When the volume is indicated to be **Inactive**, clicks on the **Online/Offline** indicator toggle the color and label for the indicator.

8 Click on the **Format Request** option button.

- A pop-up option menu is displayed for selection of **Yes** or **No**, and when one of those options is clicked, the indicated choice is displayed on the option button.

9 Use the mouse to move the cursor to the **Block Size** field.

- A blinking cursor appears in the **Block Size** field.

10 Use the keyboard to enter or change the value in the **Block Size** field.

- The entered data appear in the **Block Size** field.

11 Use the mouse to move the cursor to the **Volume Label** field.

- A blinking cursor appears in the **Volume Label** field.

12 Use the keyboard to enter or change the value in the **Volume Label** field.

- The entered data appear in the **Volume Label** field.

13 If you wish to cancel any request for changes to the volume, click on the **Cancel** button at the bottom of the window. If you wish to accept the changes, click on the **Accept** button at the bottom of the window.

- When you click the **Accept** button, *AAWin* attempts to make the requested changes. For most changes, specifically changes to **Online/Offline** and **Active/Inactive** status, the requested **Volume Group** for the volume, and the **Volume Label**, the changes can be made immediately. But if a format has been requested, then the **Online/Offline** and **Active/Inactive** status changes are not applied immediately. Instead, the requests for these status changes and the format changes are passed to the *AAWin* **Scheduler** daemon for processing. Changes made by the **Scheduler** occur when the job is processed, which depends on how many other jobs are currently scheduled.

Archive Monitoring and Troubleshooting

Trouble Symptoms

Troubleshooting is a process of identifying the source of problems on the basis of observed trouble symptoms. Most Archive problems can be traced to some part of the Data Server Subsystem, especially components of the Science Data Server or Storage Management.

However, a common source of problems involves the reliance on messages or data from other subsystems. Like many other operational areas in the system, Archive has interfaces with other subsystems. Consequently, it is possible to trace some problems to another subsystem, including (but not necessarily limited to) those in the following list:

- Ingest Subsystem (INS).
- Planning Subsystem (PLS).
- Data Processing Subsystem (DPS).
- Communications Subsystem (CSS).
- Data Management Subsystem (DMS).

The general process of troubleshooting involves the following activities:

- Review the trouble symptoms.
- Check the status of relevant hosts/servers (as necessary).
- Check log files (as necessary).
- Take action to correct the problem(s).

Fault Recovery

Each request that crosses a client/server boundary is assigned a system-unique identifier referred to as an RPC ID. (RPC refers to Remote Procedure Call, the mechanism by which requests are submitted from client to server.) The RPC ID facilitates the automatic fault recovery events that occur whenever there is a client or server failure.

- As a request propagates through the system, each associated client/server exchange is assigned a unique RPC ID.
  - The RPC ID for each interaction is derived from the previous RPC ID received by the client for the request.
    - Consequently, all RPC IDs associated with a given request have a common portion that relates the various client/server calls to one another.

  - Given the previous RPC ID, clients consistently reproduce the same RPC ID that was submitted to the server on the subsequent event.

- The concept of reproducible RPC IDs is central to the system fault recovery capability.
  - When requests are retried from client to server, they are always submitted with the same RPC ID that was used in the original submission of the request, even if either the client or server has crashed between retries.

- The RPC ID is also central to the check-pointing aspect of fault recovery.
  - As requests arrive at fault recovery-enabled servers, they are recorded in a persistent store (typically a database), tagged with the RPC ID, which identifies the request.
  - As the request is serviced, check-pointing state information may be updated in the persistent store, up to and including the completion status of the request.
  - This allows the servers to resume servicing from the last check-pointed state, particularly upon resubmission from a client.

Data Server Subsystem components checkpoint the following types of information:

- **EcDsScienceDataServer** - Asynchronous “acquire” requests that have been accepted for processing and subscription server event notifications.
- **EcDsHdfEosServer** - None.
- **EcDsDistributionServer** - Requests (which have been accepted for processing).
- **EcDsStArchiveServer** - “Store” and “retrieve” request state information.
- **EcDsStStagingDiskServer** - Resource allocation and ownership for staging disks.
- **EcDsStFtpServer** - Request state information.
- **EcDsStCacheManagerServer** - None.
- **EcDsStDTFServer** - None.
- **EcDsStRequestManagerServer** - None.

Fault Handling

Failure events are classified according to the following three severity levels:

- Fatal error.
  - Returned when a request cannot be serviced, even with operator intervention.

    - For example, if a request is made to distribute data via ftp to a non-existent host, the request is failed with a fatal error.

- Retry error.
    - Potentially recoverable error.
    - Normally, a retry error would be returned to the client only when the server cannot recover from the error automatically.
    - A retry error may require operator assistance during recovery.
        - For example, the Distribution Technician would use the Data Distribution Operator GUI to manually request resumption of a distribution request that had been "suspended with errors."

- Warning.
    - Provided when operations can proceed without interruption, but an unexpected circumstance was detected.
    - For example, if a client requests a file to be removed but the file does not exist, there is no error per se; however, a warning is generated to caution the client that the file to be removed did not exist in the first place.

Transient errors (such as network errors) are always retry errors.

- In general, clients and servers that experience transient retry errors first attempt to recover by retrying the operation automatically.
- One special case of this is "rebinding," which refers to the process by which a client automatically attempts to re-establish communication with a server in the event communication is disrupted.
    - The disruption may be caused by transient network failure, or by the server crashing or being brought down.
    - In any case, the client automatically attempts to reconnect to the server for a configurable period of time on a client-by-client basis.

System processes encountering an error or receiving an error from a server request can either pass the error back to a higher-level client or present it to the operator for operator intervention. The specific fault handling policies for Data Server Subsystem client processes are shown in Table 3.

Table 3. Data Server Subsystem and Order Manager Subsystem Fault Handling Policies

Client Process	Fault Handling Policy
EcDsScienceDataServer **EcDsHdfEosServer**	**Retry errors:** Errors are retried a configurable number of times, then passed back to the calling client process unchanged. The default retry policy for Science Data Servers is "retry forever." For asynchronous "acquire" requests involving subsetting, retry errors encountered with the HDF servers are not returned to the client. Instead, the request is queued for future execution. **Fatal errors:** Errors are passed back to the calling client process. **NOTE:** Errors associated with asynchronous requests are logged but do not appear on any GUI. The Operator restarts HDF servers manually.
EcDsDistributionServer	Errors are presented to the operator via the Data Distribution Operator GUI. **Retry errors:** Errors are presented as "Suspended with Errors" and can be resumed by the operator. **Fatal errors:** Errors are presented as "Failed." For synchronous requests, fatal errors are also passed back to the calling client process. For asynchronous requests, fatal errors are sent as part of the e-mail notification.
EcDsStRequestManagerServer **EcDsStDTFServer**	**Retry errors:** Errors are passed back to the calling client process. **Fatal errors:** Errors are passed back to the calling client process.

Client Crash and Restart

When a client of a SDSRV or DDIST server crashes, the server (i.e., EcDsScienceDataServer, EcDsHdfEosServer, or EcDsDistributionServer) continues to service the requests that were in process at the time of the client's crash. When a client of a STMGT server (i.e., EcDsStArchiveServer, EcDsStRequestManagerServer, EcDsStCacheManagerServer, EcDsStPullMonitorServer, EcDsStFtpServer, EcDsStDTFServer, or EcDsStStagingDiskServer) crashes, the requests that were in process are cancelled by another client process and there is no impact to the outside requester server.

When a client restarts in the system, it sends a restart notification to each server with which it interacts.

- Clients notify servers that they have come up either "cold" or "warm."
- Generally, the notification temperature sent to the server matches the temperature at which the client process is restarted.
- However, there are some exceptions; for example:
  - EcDsScienceDataServer always notifies EcDsDistributionServer that it has performed a warm restart.

  - The default behavior for both EcDsHdfEosServer and EcDsStDTFServer is to send EcDsStRequestManagerServer cold restart notification.

- When a client sends restart notification to the EcDsStRequestManagerServer, the server calls a stored procedure to clean up the old request and staging disk (if any) created by the client, based on whether it was a cold or warm start.
  - The Storage Management Servers are not directly notified when a restart has occurred.
  - The Storage Management Servers respond to the event according to the fact that a previous request has been marked as failed and any staging disk resources they have allocated have been released.

The default server behavior in response to startup notification from a client is as follows:

- Warm Notification.
  - Outstanding requests for the restarted clients remain available in the persistent store.
  - The outstanding requests may be resubmitted by the client, and are serviced to completion upon resubmission.
  - Associated resources are left allocated until the requests are completed.
- Cold Notification.
  - All outstanding requests for the restarted client are cancelled.
  - If the client resubmits any cancelled request using the same RPC ID (e.g., by pressing the Retry button from an operator GUI), it is failed with a fatal error due to the client cold startup notification.
  - Any resources associated with the cancelled requests are released and reclaimed by the system.

The following servers have some non-standard responses to startup notification:

- **EcDsStArchiveServer**.
  - **Warm Notification:**
    - Default server behavior (as previously described).
  - **Cold Notification:**
    - For partially completed Ingest operations, all files stored are removed. (Partial granules are never permitted in the archive.)
- **EcDsStStagingDiskServer**.
  - **Warm Notification:**

        - All staging disks owned by the restarted client are retained, including temporary staging disks.
    - **Cold Notification:**
        - All staging disks owned by the restarted client are released.

Server Crash and Restart

When a server crashes, clients cannot continue to submit requests for processing.

- Synchronous requests in progress result in a Distributed Computing Environment (DCE) exception being thrown back to the client process, which enters a rebinding failure recovery mode (as previously mentioned).
- Attempts to submit requests while the server is down result in the client blocking until a communication timeout has been reached.
- Although DCE has been replaced by socket-based library calls (i.e., CCS Middleware), the DCE exception code is handled by the CCS Middleware.

When a server restarts, it may perform various resynchronization activities in order to recover from an unexpected termination.

- In the event of a server cold start or cold restart, the server typically cancels all outstanding requests and reclaims all associated resources.
- In general, existing request queues are retained for warm restarts and cleared for cold starts or cold restarts.
- **EcDsScienceDataServer**- and **EcDsHdfEosServer**-specific activities upon start/restart:
    - **Warm Restart:**
        - Restart asynchronous “acquire” requests that were in progress before the crash.
        - Retain the queue of asynchronous “acquire” requests.
        - It is expected that synchronous requests would be resubmitted by the respective senior client applications (i.e., PRONG or INGST).
        - Send event notifications to the Subscription Server for any services completed before the crash for which a subscribed event is registered but has not been sent to the Subscription Server.
    - **Cold Start or Cold Restart:**
        - Purge the queue of asynchronous “acquire” requests.
        - Purge the queue of Subscription Server Event Notifications.

- **EcDsDistributionServer**-specific activities upon start/restart:
  - **Warm Restart:**
    - Request Processing is restarted from the last check-pointed state.
  - **Cold Start or Cold Restart:**
    - STMGT CI is informed of a cold start.
    - EcDsDistributionServer deletes all (prior) request information from its database.
- **EcDsStArchiveServer**-specific activities upon start/restart:
  - **Warm Restart:**
    - Retains existing request queues.
  - **Cold Start or Cold Restart:**
    - For partially completed “store” requests, the files copied into the archive are removed.
    - For partially completed “retrieve” requests, the access count is decremented in the read-only cache.
- **EcDsStCacheManagerServer**-specific activities upon start/restart:
  - **Warm Restart:**
    - The contents of the read-only cache are synchronized with the database.
    - Discrepancies are logged and removed.
  - **Cold Start or Cold Restart:**
    - All files are removed from the read-only cache.
- **EcDsStStagingDiskServer**-specific activities upon start/restart:
  - **Warm Restart:**
    - The set of staging disks in the staging area is synchronized with the database.
    - Discrepancies are logged and removed.
    - Existing request queues are cleared.
  - **Cold Start or Cold Restart:**
    - All staging disks are removed.
- **EcDsStPullMonitorServer**-specific activities upon start/restart:

- **Warm Restart:**
    - The contents of the Pull Area and user request areas are synchronized with the database.
    - Discrepancies are logged and removed.
- **Cold Start or Cold Restart:**
    - All files in the Pull Area and all user request areas are removed.

- **EcDsStFtpServer**-specific activities upon start/restart:
    - **Warm Restart:**
        - Existing request queues are retained.
    - **Cold Start or Cold Restart:**
        - Existing request queues are cleared.

Request Resubmission

Upon restarting a crashed client or server, requests are typically resubmitted. If the restarted process was started warm, the fault-recovery capabilities permit the server to resume processing of the request from its last check-pointed state. This prevents needless repetition of potentially time-consuming activities.

- **EcDsScienceDataServer**- and **EcDsHdfEosServer**-specific activities upon resubmission of a request:
    - All requests are serviced as if they are new requests.
    - RPC IDs are generated automatically and reproducibly; consequently, the Science Data Server typically recreates the same allocation requests on a resubmission.
        - This can trigger special logic to handle requests for which an allocated staging disk has been transferred to the Data Distribution Server.
- **EcDsDistributionServer**-specific activities upon resubmission of a request:
    - If previously submitted and completed with the same RPCId, the request status is returned based on the check-pointed request status.
    - If previously submitted and completed with different RPCIds, the request is re-executed.
    - Otherwise, the client request thread is synchronized with the worker thread actually servicing the request.

- **EcDsStArchiveServer**-specific activities upon resubmission of a request:
  - The request is restored from the last check-pointed state.
  - For “store” requests, copies into the archive are resumed from the last file copied.
  - For “retrieve” requests, the entire “retrieve” request is reprocessed.
    - However, files previously retrieved for the request are, in all likelihood, still in the read-only cache.
- **EcDsStCacheManagerServer**- and **EcDsStFtpServer**-specific activities upon resubmission of a request:
  - If previously submitted and completed, the request status is returned based on the check-pointed request status.
  - Otherwise, the request is processed anew.
- **EcDsStStagingDiskServer**-specific activities upon resubmission of a request:
  - For staging disk allocation, the results are returned to the client if the client resubmits the allocation request under which the disk was created.
- **EcDsStPullMonitorServer**- and **EcDsStDTFServer**-specific activities upon resubmission of a request:
  - The resubmitted request is processed as if it were a new request.

Troubleshooting Archive Problems

Previous sections of this lesson have addressed the use of tools that can assist you in monitoring the Archive. Specifically, the System Requests tab of the Science Data Server Operator GUI provides a view into archive request processes, the Data Distribution GUI provides information on distribution requests, and the AMASS GUI offers ready access to information about the status of archive volumes and volume groups. If archive problems arise, there are additional resources that can provide more detailed monitoring and assistance in troubleshooting. Troubleshooting is a process of identifying the source of problems on the basis of observed symptoms. Because the Archive is at the heart of the system and its Data Server interacts with so many subsystems, problems with Archive functions may be traced to the Data Server subsystem (DSS) or one of many other subsystems, including (but not limited to) the following:

- Ingest Subsystem (INS).
- Planning Subsystem (PLS).
- Data Processing Subsystem (DPS).
- Communications Subsystem (CSS).
- Data Management Subsystem (DMS).

Table 4 summarizes actions to be taken in response to some common Archive problems. If the problem cannot be identified and fixed without help within a reasonable period of time, the appropriate response is to call the Help Desk and submit a trouble ticket in accordance with site Problem Management policy.

Table 4. Troubleshooting Archive Problems

Symptom/Problem	Response
Unable to log in to the FSMS host (e.g., *x*0drg01)	Check with the Operations Controller/System Administrator to ensure that the host is "up."
AMASS is not running	Have the System Administrator restart AMASS.
A volume is inactivated by AMASS	Check for AMASS errors and, unless there are many errors, use the command **/usr/amass/bin/volstat -a <*vol_number*>** to re-activate the volume. [For detailed instructions, refer to the procedure **Display AMASS Errors Using the *amass_log* Script** (subsequent section of this lesson).]
A storage system robot gets out of synchrony with AMASS concerning the location of media.	Re-establish synchrony. [For detailed instructions, follow the procedure **Establish Synchrony Between *quedisplay* and *medialist* Using *mediamove*** (subsequent section of this lesson).]
An Ingest or Data Processing action cannot complete because of failure to store data (reflected as failure on Ingest or Processing GUIs).	Check to ensure AMASS is on line; check for file copy errors, network problems, mount point problems. [For detailed instructions, refer to the procedures in the section on **Recovery from Failure to Store or Retrieve Data** (subsequent section of this lesson).]

Although there are no custom reports of archive events, the Storage Management GUIs provide a **Storage Events** tab, as shown in Figure 26. This tab allows an operator to review events in the storage management Event Log. The Event Log Search Parameters box permits specification of constraints to limit the event log search that is executed when the **Search** button is clicked. This box enables an operator to select a date interval for the search, and it provides option buttons to specify the type and level of events to be displayed. Options on the **Event Type** option button are: **Any**, **Device**, **Cache**, **Software**, **COTS**, **Sybase**, **Pulldisk**, and **Unknown**. Options on the **Event Level** option button are: **Any**, **Information**, **Warning**, **Error**, **Severe**, **Fatal**, and **Unknown**. The resulting events are then displayed in the **Event Log** field. A **Purge Selected** button at the bottom of the window permits deletion of entries in the Event Log that have been selected, or highlighted, by clicking on them in the **Event Log** field.

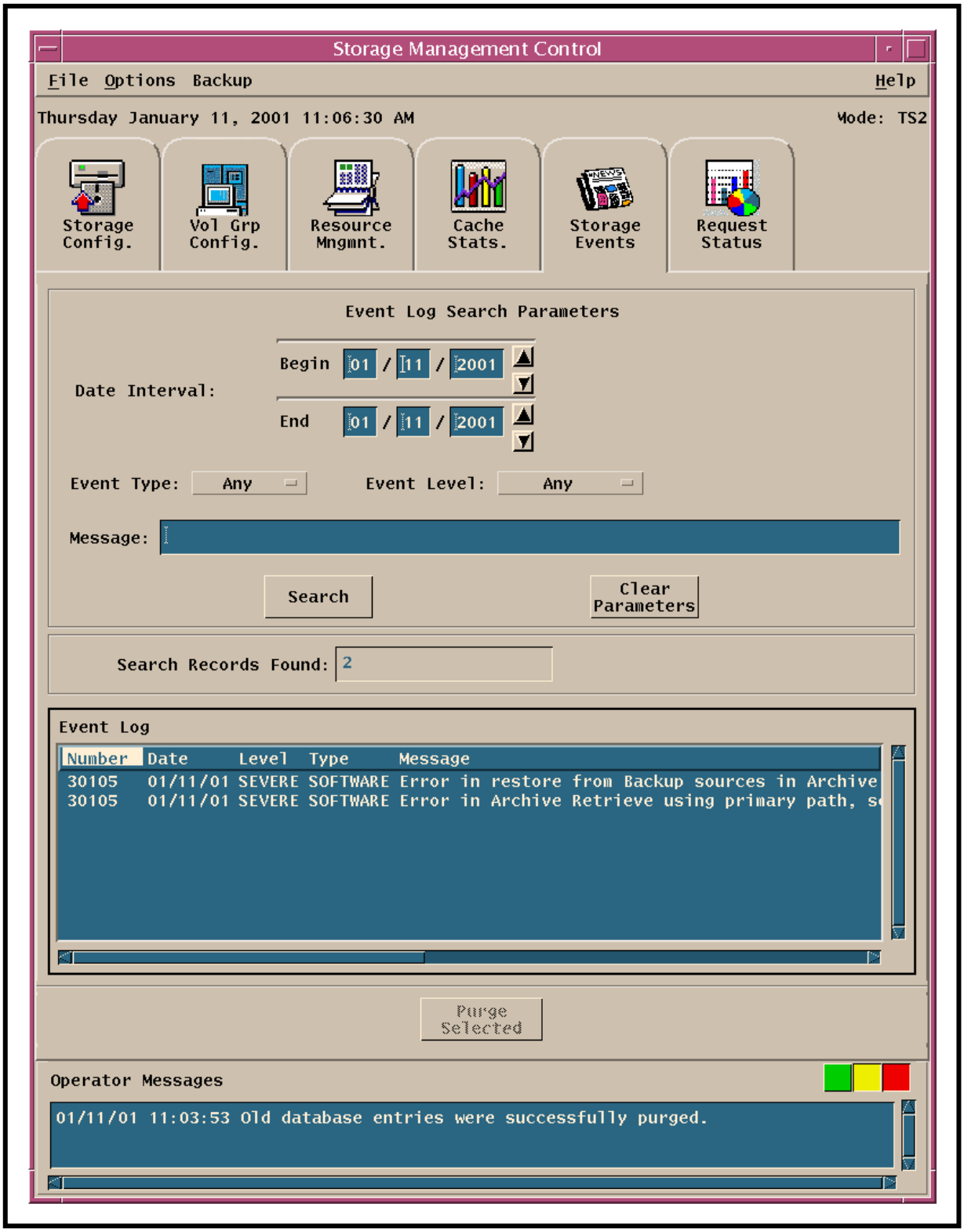


Figure 26. Storage Management, Storage Events Tab

The Storage Management GUI also provides a powerful capability for tracking the status of requests associated with archive activities (see Figure 16 and associated procedure). This is the **Request Status** tab, which permits filtering requests to specific Storage Management request types and quick determination of the status of an individual request. Identification of the status of an archive request using this tool can be a helpful first step in troubleshooting apparent archive problems.

There are also several troubleshooting tools provided with AMASS that can assist you in monitoring archive activity and in responding to fault notifications. The AMASS *System Administrator's Guide* includes instructions on using these tools. Some of the most useful ones are addressed here.

AMASS Commands, Utilities, and Scripts for Monitoring and Fault Response

The AMASS file system needs to have the following daemons running at all times:

- **amassmain**.
- **daemons/lm_ip -a fslock.**
- **qset**.
- **klogd**.
- **amass_iocomp**.
- **libsched** (one instance for each virtual library).
- **libio_tape** (at least one instance for each drive in each jukebox).

The UNIX process search provides an easy check for these daemons. If they are up, the AMASS **healthcheck** command provides a useful check on the health of AMASS while it is running. Checking for these daemons and running **healthcheck** can be an appropriate initial response to an indication of a potential problem with AMASS. Use the following procedure.

Check Daemons and Use *healthcheck*

---

1 Log in as **amass** at the FSMS host.

2 Type **ps -ef | grep amass** and press the **Return/Enter** key.

- UNIX returns running AMASS processes in a format similar to the following:

```
amass   7214464   7208385  0   Sep 19 ?      0:00 libio_tape 2 1
amass   7208385      1  0   Sep 19 ?    10:33 /usr/amass/daemons/amassmain 0
amass   7214747   7208385  0   Sep 19 ?      0:10 amass_iocomp
amass   7282853   7215637  0   Sep 20 ?      1:47 libio_tape 1 1
amass   7282868   7215637  0   Sep 20 ?      0:00 libio_tape 1 1
amass   6949087   7215637  0   Sep 20 ?      1:47 libio_tape 1 1
```

```
amass    7214915    7208385  0   Sep 19 ?      0:00 klogd
amass    7214972    7208385  0   Sep 19 ?     50:54 libio_tape 1 2
amass    5539722    7217884  0   Sep 20 ?      0:23 libio_tape 1 3
amass    7301726    7215964  0   Sep 20 ?      1:05 libio_tape 3 1
amass    7215313        1  0   Sep 19 ?      9:34 /usr/amass/daemons/lm_ip -a
fslock1 -u 128 -f 256 -q 128
amass    7357656    7216363  0   Sep 20 ?      0:00 libio_tape 3 3
amass    7215637    7208385  0   Sep 19 ?     84:10 libio_tape 1 1
amass    7215638    7208385  0   Sep 19 ?      2:43 libsched 3
amass    7277545    7214972  0   Sep 20 ?      0:41 libio_tape 1 2
amass    7215870    7208385  0   Sep 19 ?      2:52 libsched 1
amass    7215964    7208385  0   Sep 19 ?    109:25 libio_tape 3 1
amass    7216363    7208385  0   Sep 19 ?     84:16 libio_tape 3 3
amass    6950984    7217884  0   Sep 20 ?      0:23 libio_tape 1 3
amass    8175053    7212410  0   Sep 26 ?      0:00 libio_tape 1 4
amass    7340525    7217134  0   Sep 20 ?      1:19 libio_tape 3 2
amass    7278745    7217884  0   Sep 20 ?      0:23 libio_tape 1 3
amass    7216941    7208385  0   Sep 19 ?      0:32 qset
amass    7340710    7217134  0   Sep 20 ?      1:19 libio_tape 3 2
amass    7217134    7208385  0   Sep 19 ?    138:26 libio_tape 3 2
amass    7359550    7216363  0   Sep 20 ?      0:52 libio_tape 3 3
amass    7217388    7208385  0   Sep 19 ?      0:00 libio_tape 2 2
amass    7285477    7215637  0   Sep 20 ?      1:47 libio_tape 1 1
amass    7285537    7215637  0   Sep 20 ?      1:47 libio_tape 1 1
amass    7217884    7208385  0   Sep 19 ?     17:37 libio_tape 1 3
amass    7279821    7214972  0   Sep 20 ?      0:41 libio_tape 1 2
amass    6878049    7208385  0   Sep 19 ?      2:36 libsched 2
amass    7279907    7214972  0   Sep 20 ?      0:41 libio_tape 1 2
amass    7335573    7217134  0   Sep 20 ?      1:19 libio_tape 3 2
    .
    .
    .
```

- If the running processes do not include **amassmain**, **daemons/lm_ip -a fslock**, **qset**, **klogd**, **amass_iocomp**, **libsched**, and **libio_tape**, it may be necessary to restart AMASS [refer to the **Reboot AMASS** procedure (previous section of this lesson)].

3 To check the AMASS database integrity, check the availability of write resources FNODEs and cache blocks, and to verify cache partitions, type **/usr/amass/bin/healthcheck -viwc** and press the **Return/Enter** key.

- AMASS returns information on its health in format similar to the following:

  --- STARTING DATABASE INTEGRITY CHECK ---

  -api has been opened properly and AMASS is running.
  -verifying pathnames.

-got locks on database
-unlocking database tables and exiting

--- CHECK COMPLETED!! ---

--- CHECKING AVAILABILITY OF WRITE RESOURCES FNODEs AND CACHE BLOCKS ---

-api has been opened properly and AMASS is running.
-Initializing the passed arguments.
-Returning the passed arguments.
-Restoring signals.
-exiting.

--- CHECK COMPLETED!! ---

--- RUNNING CACHE TEST ---

-api has been opened properly and AMASS is running.
-Validating the raw cache.
-Restoring signals.
-exiting.

--- TEST COMPLETED!! ---

If an error message is returned, it may be necessary to restart AMASS [refer to the **Reboot AMASS** procedure (previous section of this lesson)].

4 To check library components, type **/usr/amass/bin/healthcheck -vl 1 0** ***volumenumber*** and press the **Return/Enter** key.

- The argument **-l** (lower-case l) specifies the library components check, and requires specification of a jukebox (**1** in this case), a drive number (entering **0** as in this case checks all active drives), and a volume number (***volumenumber*** is the volume ID of an available volume in the specified jukebox; it may be helpful to use the **vollist** command [refer to the **Display Volume Data Using** ***vollist*** procedure (subsequent section of this lesson)] to identify a suitable volume, such as a volume in the Space Pool, to use for this test).
- AMASS returns information on the health of library components in the following format:

--- CHECKING LIBRARY COMPONENTS ---

-api has been opened properly and AMASS is running.
-mapping shared memory.
-verifying the juke number.
-validating volume number.

-validating drive number and checking for active drive/s/.
-saving the volume's status before inactivating it.
-proceeding with physical test.
-restoring signals and exiting.

--- CHECK COMPLETED!! ---

- If an error message is returned, it may be necessary to restart AMASS [refer to the **Reboot AMASS** procedure (previous section of this lesson)] and/or to check for possible hardware problems with drives or other components.

A command provided to display the status of the AMASS I/O activity is **sysperf**. This command returns several items:

- the number of reads and writes that are outstanding.
- the number of volumes (for reads) or volume groups (for writes) that are going to be used by those reads and writes.
- the current volumes in the drives.
- the I/O rate in Kb per second since the last update. This value first appears as a zero. Then AMASS continues to update the information at intervals based on a value for *updateinterval* entered by the operator.

Sysperf can often show the first sign of trouble. For example, if there are reads and writes in process but throughput is always 0, a problem is indicated. The most common problems are volumes and drives that go off line and/or inactive.

To run **sysperf**, use the following procedure.

Display the Status of AMASS I/O Activity Using *sysperf*

---

1 Log in to the FSMS host (e.g., e0drg11, g0drg01, l0drg01, or n0drg01).

2 Type **/usr/amass/bin/sysperf -k 5** and press the **Return/Enter** key.

- The screen updates every 5 seconds and display information on the amass kernel (**-k**) in a form similar to the following example:

```
SYSTEM STATISTICS        -  Thu Sep 27 08:17:33
UPDATE INTERVAL          -  10 SEC
AVERAGE THROUGHPUT -   0 KBYTES/SEC

READ REQUESTS        # OF VOLUMES
        0                  0

WRITE REQUESTS       # OF VOL GROUPS
        0                  0
```

CACHE BLOCKS 2957 Total 2957 Free 0 Dirty
FNODES 800 Total 796 Free 4 Used

JUKE DRIVE VOLFLAGS VOLUME VOLGRP KBYTES/SEC

- A different number of seconds may be entered to specify a different refresh rate.

3 To break out of the command, use **ctrl-c** (while holding down the **Control Key**, press **c**).

- The screen stops updating and displays a UNIX prompt.

Volumes are monitored using the **vollist** command. For example, to show data on a particular volume (e.g., volume 100) use the following procedure.

Display Volume Data Using *vollist*

---

1 Log in to the FSMS host (e.g., e0drg11, g0drg01, l0drg01, or n0drg01).

2 Type **/usr/amass/bin/vollist ###** then press the **Return/Enter** key.

- ### is the volume number.
- For example, to specify volume 100 type the following command.

  /usr/amass/bin/vollist 100

- AMASS displays the following:

VOL NUM	VOL GRP	JUKE NUM	POS	VOL LABEL	FLAGS	USED (MB)	AVAIL (MB)	DEAD (%)	ERRS
100	500	3	NET	SD0060	O	99213	3167	0	0

- In this example, the **O** in the **FLAGS** column indicates that the volume is offline.

3 To put a volume back on line, type **/usr/amass/bin/volloc -n ###** then press the **Return/Enter** key.

- For example, to put volume 100 back on line type the following command.

  /usr/amass/bin/volloc –n 100

4 Type **/usr/amass/bin/vollist ###** then press the **Return/Enter** key.

- AMASS displays the following:

VOL NUM	VOL GRP	JUKE NUM	POS	VOL LABEL	FLAGS	USED (MB)	AVAIL (MB)	DEAD (%)	ERRS
100	500	3	NET	SD0060	A	99213	3167	0	0

- In this example, the **A** in the **FLAGS** column indicates that the volume is now on line and Active.

If the output of **vollist** indicates that the volume is inactivated (i.e., there is an **I** in the **FLAGS** column), use the **amass_log** script to determine the nature of the problem. The **amass_log** script displays AMASS messages from the system log file. This script can provide helpful information under several circumstances, such as when a command gives unexpected results or when AMASS appears not to be functioning properly in other ways. Use the following procedure to run **amass_log**.

Display AMASS Errors Using the *amass_log* Script

---

1 Log in to the FSMS host (e.g., e0drg11, g0drg01, l0drg01, or n0drg01).

2 To change to the AMASS tools directory, type **cd /usr/amass/tools** and then press the **Return/Enter** key.

- The working directory is changed to **/usr/amass/tools**.

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

- ***logfilepath*** is the full pathname of the system log file to scan for AMASS messages.
- On a Sun, the ***logfilepath*** is likely to be **/var/adm/messages**; on an SGI, the ***logfilepath*** is likely to be **/var/adm/SYSLOG**. Any AMASS error messages in the scanned log file are displayed.

4 Perform the action recommended for the error message in the log.

- The AMASS *System Administrator's Guide* [available electronically on **drg** servers (e.g., g0drg01, e0drg11, l0drg01, or n0drg01) in directory **/usr/amass/books**] provides detailed information concerning error messages. An error message informs of critical problems that prevent AMASS from functioning. An error message is usually followed by a correction message, which provides instructions for correcting the situation. Sometimes, there is a previous warning message that may provide an accompanying correction message. Other messages may be identified by number only; the *System Administrator's Guide* provides a reference list, with accompanying corrective actions.

Unless use of the **amass_log** script shows that there are many errors on a volume that has been inactivated, you can reactivate the volume using the command:

/usr/amass/bin/volstat -a 100

(for this example, to reactivate volume 100).

Just as **vollist** provides information on the status of volumes, the command **drivelist** displays the status of drives available to AMASS. Active drives are noted by an **A**, and inactive drives are noted by an **I**. The command is **/usr/amass/bin/drivelist**. If AMASS inactivates a drive, use the **amass_log** script as described previously. Unless there is a hardware problem and several attempts have been made to ready the drive, it is usually appropriate to reactivate the drive using the **drivestat** command. For example, to reactivate drive 1 in jukebox 1, type the following command:

/usr/amass/bin/drivestat -a 1 1

A useful library utility included with AMASS is **quedisplay**. This utility permits the operator to see what is in the queue, and to diagnose problems such as the following:

- During an attempt to write to a file, the drive light does not illuminate.
- The system is slowing down.
- An AMASS command does not complete.

Figure 27 shows an example of the form of the output of the **quedisplay** utility. The output shows the queue, which consists of read and write requests, AMASS administration commands, and a list of libraries, drives, and what volumes they manage.

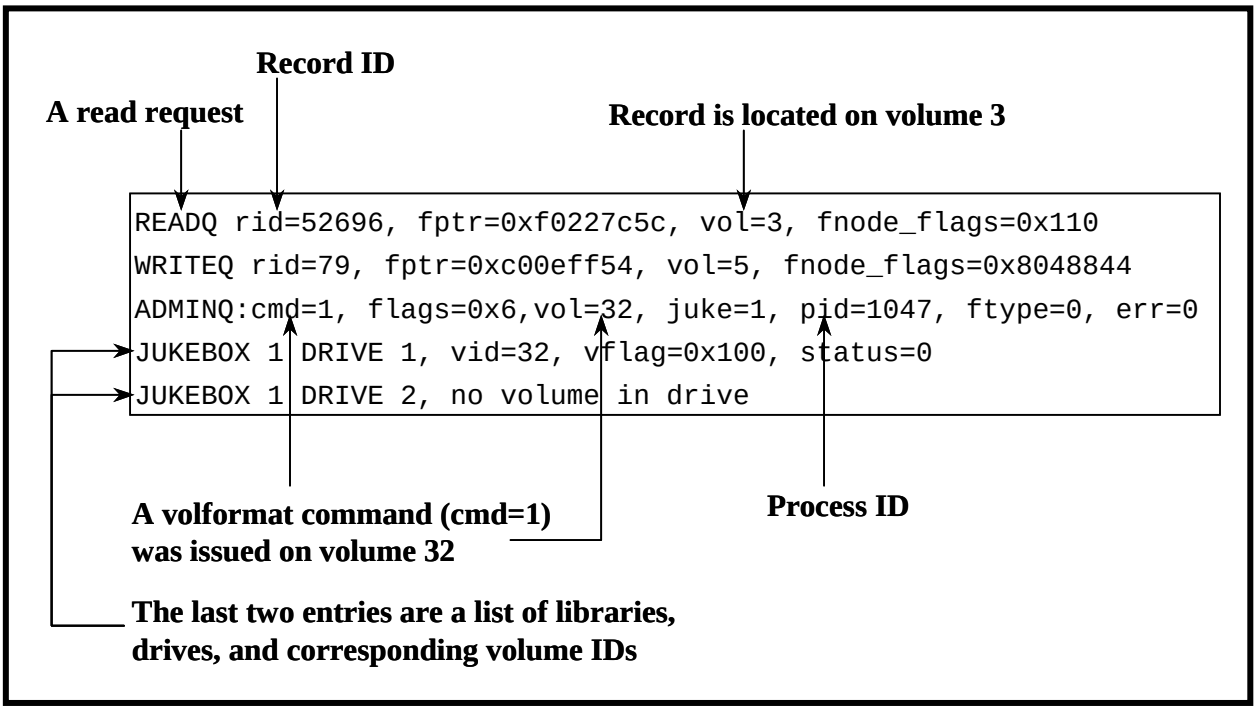


Figure 27. Sample Output from AMASS* quedisplay *Utility

Use the following procedure to monitor what is in the queue.

View What Is in the AMASS Queue Using *quedisplay*

1 Log in to the FSMS host (e.g., e0drg11, g0drg01, l0drg01, or n0drg01).

2 To change to the utilities directory, type **cd /usr/amass/utils** and then press the **Return/Enter** key.

- The working directory is changed to **/usr/amass/utils**.

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

- The AMASS queue is displayed in the following form:

  READQ rid=52696, fptr=0xf0227c5c, vol=3, fnode_flags=0x110
  WRITEQ rid=79, fptr=0xc00eff54, vol=5, fnode_flags=0x8048844
  ADMINQ:cmd=1, flags=0x6,vol=32, juke=1, pid=1047, ftype=0, err=0
  JUKEBOX 1 DRIVE 1, vid=32, vflag=0x100, status=0
  JUKEBOX 1 DRIVE 2, no volume in drive

- In the output, "rid" = Record ID, "pid" = Process ID

If there are **READQ** or **WRITEQ** entries, the name(s) of the file(s) being processed can be determined by using the **filepath** command and the first number in the entry. For example, type **/usr/amass/utils/filepath 52696** for the first file number in the sample output shown previously in Figure 27.

Occasionally, a robot may lose synchrony with AMASS as to the location of media. The best way to verify this is to compare **quedisplay** and **medialist**. The **medialist** utility is a standalone program that communicates with the robot controller in the Powderhorn to determine the robot's view of media and their slot locations. If the two programs disagree, you can bring the two programs into synchrony using **mediamove**. Use the following procedure.

Establish Synchrony Between *quedisplay* and *medialist* Using *mediamove*

1 Log in as **amass** at the FSMS host (e.g., e0drg11, g0drg01, l0drg01, or n0drg01).

2 Type **/usr/amass/utils/quedisplay** and then press the **Return/Enter** key.

- AMASS displays the following information (for example of incorrect status).

  . . .
  JUKEBOX 1 DRIVE 1, no volume in drive
  JUKEBOX 1 DRIVE 2, vid=50, vflags-0x4, status=0

3 Type **/usr/amass/utils/medialist** and then press the **Return/Enter** key.

- AMASS displays the following information (for example of actual status).

  . . .
  SLOT VSD0098 FULL
  DRIVE 1 FULL FROM VSD0096
  DRIVE 2 FULL FROM VSD0097

- Note that the **medialist** result shows that drive 1 actually is occupied, although **quedisplay** registers that drive 1 is empty.

4 Type **/usr/amass/utils/mediamove 1 VSD0096 1** and then press the **Return/Enter** key.

- AMASS moves the volume from the ***source*** (drive **1** in this example) to the ***destination*** (slot **VSD0096** in this example) in the specified ***jukeboxnumber*** (jukebox **1** in this example), thereby bringing the actual status of drive 1 (as known by ***medialist***) to the status reflected by ***quedisplay***.

Recovery from Failure to Store or Retrieve Data

Successful data storage and retrieval functions are the heart of the system. Successful ingest of data or processing of data to produce new science data granules require that Storage Management (STMGT) is inserting the product into the archive and that Science Data Server (SDSRV) is inserting the associated metadata into the inventory. Staging disks and cache managers for the Archive server and the FTP server are also involved in this process. To check the functioning of these elements, it is necessary that the ESDTs for the data to be inserted are installed and available, and that subscriptions have been registered.

Troubleshooting failures to store or retrieve data (as well as other failures) often requires review of server or application log files. This section contains a general procedure for reviewing log files to check for proper start-up and communications. It also has a procedure for a special case of reviewing log files for the Storage Management Request Manager server, and a procedure for reviewing the current **tac** log file of interactions between AMASS and ACSLS. Separate procedures then address recovery from failure to insert (store) data and recovery from failure to acquire (retrieve) data.

Checking Server Log Files

Use the following procedure for checking server log files for nominal start-up and communications.

Check Server Log Files

---

1 Log in to the host for the server and log(s) to be examined.

2 Type **cd /usr/ecs/<*MODE*>/CUSTOM/logs** and then press the **Return/Enter** key.

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

3 To view a server log, type **pg *filename*** and then press the **Return/Enter** key.

- ***filename*** refers to the log file to be reviewed (e.g., **EcDsScienceDataServer.ALOG**, **EcDsScienceDataServerDebug.log**).
- The first page of the log file is displayed; additional sequential pages can be displayed by pressing the **Return/Enter** key at the **:** prompt.
- Although this procedure has been written for the **pg** command, any UNIX editor or visualizing command (e.g., **vi**, **view**, **more**, **tail**) can be used to review the log file.
- Typically, the **<*server*>Debug.log** captures more detailed information than the **<*server*>.ALOG**. However, for some servers (e.g., **SDSRV**), there may be significant detail in the **<*server*>.ALOG**. It is also important to note that the **DebugLevel** parameter setting in the **Configuration Registry** determines the level of detail captured in the **<*server*>Debug.log** (**0** is off, a setting of **1** captures status and errors, a setting of **2** captures major events, and a setting of **3** is a full trace recording of all activity). If the **DebugLevel** has been set to one of the lower levels during operations, the System Administrator may set it to **3** during troubleshooting. Similarly, the **AppLogLevel** parameter setting determines the level of detail captured in the **<*server*>.ALOG** (**0** provides a full trace recording of all events, **1** provides messages related to all major events, **2** yields just records of errors, and **3** turns recording off).
  - There are other debug levels available for some logs; Storage Management (STMGT) offers "enhanced" debugging based on bitmasks. Level **7** provides a four-bit level for detailed database debugging. Level **15** provides an eight-bit level that repeatedly dumps the in-memory request queue in the STMGT Request Manager.

4 Review the log file(s) to determine if there are any indications of connection problems or errors at start up.

- The log file for the called server may contain an error message indicating a problem at start-up. The debug log should indicate a typical start sequence, including (sample log entries in the following material were taken from a debug log showing start-up for **EcDsStFtpServer**):
  - Get parameters from registry (log entries similar to the following).

    DSS EcDsStFtpServer Server Debug log on f2acg01 starting at Mon Jun 4 07:57:45 EDT 2001
    EcAgInstanceID Sequence Number is 3870
    Setting up environment variables needed for DCE:
    RPC_UNSUPPORTED_NETIFS = ""
    /usr/ecs/DEV07/CUSTOM/bin/DSS/EcDsStFtpServer ConfigFile

/usr/ecs/DEV07/CUSTOM/cfg/EcDsStFtpServer.CFG ecs_mode DEV07 StartTemperature cold
Started process EcDsStFtpServer in mode DEV07 with PID 2709893
EcRgRegistry_1_0::ctor this = 0x104eef38
EcRgRegistry_1_0::ctor this = 0x104eef88
FoIpPtToPtPortalImp::Send sent 20/20
FoIpPtToPtPortalImp::Send sent 219/219
FoIpPtToPtPortalImp::Receive got 20
FoIpPtToPtPortalImp::Receive got 1024
FoIpPtToPtPortalImp::Receive got 246
********* After Retrieving of RGY: Name = EcDsStFtpServerNONE**
ProgramID = 4645102
ApplicationID = 4600000
Release = B
DeltaTime = 0
Site = RBD
SubSysName = DSS
MajorVersion = 1
MinorVersion = 0
DebugLevel = 3
AppLogLevel = 0
AppLogSize = 3000000
DBServer = f2acg01_srvr
DBLoginName = EcDsStFtpServer
DBName = stmgtdb1

– Load resource catalogs (log entries indicate the loading, or that the loading did not complete, similar to the following).

06/04/01 07:57:47: Thread ID : 65536 : loading resource catalog file from /usr/ecs/DEV07/CUSTOM/data/DSS/ResourceCatalogs/DsMdResource.dat .rcat
06/04/01 07:57:48: Thread ID : 65536 : loading resource catalog file from /usr/ecs/DEV07/CUSTOM/data/DSS/ResourceCatalogs/EcDsSdHr.dat.rcat
06/04/01 07:57:48: Thread ID : 65536 : loading resource catalog file from /usr/ecs/DEV07/CUSTOM/data/DSS/ResourceCatalogs/DsSrResource.dat.r cat
06/04/01 07:57:48: Thread ID : 65536 : loading resource catalog file from /usr/ecs/DEV07/CUSTOM/data/DSS/ResourceCatalogs/DsGlResource.dat. rcat
06/04/01 07:57:48: Thread ID : 65536 : loading resource catalog file from /usr/ecs/DEV07/CUSTOM/data/DSS/ResourceCatalogs/DsShResource.dat. rcat
06/04/01 07:57:48: Thread ID : 65536 : loading resource catalog file from /usr/ecs/DEV07/CUSTOM/data/DSS/ResourceCatalogs/EcDsSdHc.dat.rcat

– Pre-cache errors associated with database connectivity (log entries similar to the following).

06/04/01 07:57:48: Thread ID : 65536 : User Name : EcDsStFtpServer | Thread 65536
06/04/01 07:57:48: Thread ID : 65536 : Database Name : stmgtdb1_DEV07 | Thread 65536
06/04/01 07:57:49: Thread ID : 65536 : Server Name : f2acg01_srvr | Thread 65536
06/04/01 07:57:49: Thread ID : 65536 : DsShTSStorage: creating the MutexVec for this thread
06/04/01 07:57:49: Thread ID : 65536 : SEARCHING FOR: 30141 (Not found) | Thread 65536
06/04/01 07:57:49: Thread ID : 65536 : CACHING: DsEStUnknownError (30141) | Thread 65536
06/04/01 07:57:49: Thread ID : 65536 : SEARCHING FOR: 30143 (Not found) | Thread 65536
06/04/01 07:57:49: Thread ID : 65536 : CACHING: DsEStUnknownError (30143) | Thread 65536
06/04/01 07:57:49: Thread ID : 65536 : SEARCHING FOR: 30139 (Not found) | Thread 65536
06/04/01 07:57:49: Thread ID : 65536 : CACHING: DsEStUnknownError (30139) | Thread 65536
06/04/01 07:57:49: Thread ID : 65536 : SEARCHING FOR: 30142 (Not found) | Thread 65536
06/04/01 07:57:49: Thread ID : 65536 : CACHING: DsEStUnknownError (30142) | Thread 65536
06/04/01 07:57:49: Thread ID : 65536 : SEARCHING FOR: 30148 (Not found) | Thread 65536
06/04/01 07:57:49: Thread ID : 65536 : CACHING: DsEStUnknownError (30148) | Thread 65536
06/04/01 07:57:49: Thread ID : 65536 : SEARCHING FOR: 30144 (Not found) | Thread 65536
06/04/01 07:57:49: Thread ID : 65536 : CACHING: DsEStUnknownError (30144) | Thread 65536
06/04/01 07:57:49: Thread ID : 65536 : SEARCHING FOR: 30145 (Not found) | Thread 65536
06/04/01 07:57:49: Thread ID : 65536 : CACHING: DsEStUnknownError (30145) | Thread 65536
06/04/01 07:57:49: Thread ID : 65536 : SEARCHING FOR: 30147 (Not found) | Thread 65536
06/04/01 07:57:49: Thread ID : 65536 : CACHING: DsEStUnknownError (30147) | Thread 65536
06/04/01 07:57:49: Thread ID : 65536 : SEARCHING FOR: 30146 (Not found) | Thread 65536

06/04/01 07:57:49: Thread ID : 65536 : CACHING: DsEStUnknownError (30146) | Thread 65536
06/04/01 07:57:49: Thread ID : 65536 : SEARCHING FOR: 30211 (Not found) | Thread 65536
06/04/01 07:57:49: Thread ID : 65536 : CACHING: DsEStUnknownError (30211) | Thread 65536
06/04/01 07:57:49: Thread ID : 65536 : SEARCHING FOR: 30140 (Not found) | Thread 65536
06/04/01 07:57:49: Thread ID : 65536 : CACHING: DsEStUnknownError (30140) | Thread 65536

– Get server configuration parameters from the database (log entries similar to the following).

06/04/01 07:57:49: Thread ID : 65536 : BaseReal::Ctor: Server Name is - EcDsStFtpServerNONE | Thread 65536
06/04/01 07:57:50: Thread ID : 65536 : User Name : EcDsStFtpServer | Thread 65536
06/04/01 07:57:50: Thread ID : 65536 : Database Name : stmgtdb1_DEV07 | Thread 65536
06/04/01 07:57:50: Thread ID : 65536 : Server Name : f2acg01_srvr | Thread 65536
06/04/01 07:57:50: Thread ID : 65536 : myTransactionList[0]: use stmgtdb1_DEV07 | Thread 65536
06/04/01 07:57:50: Thread ID : 65536 : DBIF:Execute: Ultimate SQL: use stmgtdb1_DEV07 | Thread 65536
06/04/01 07:57:50: Thread ID : 65536 : myTransactionList[1]: exec DsStCPSelectByName "EcDsStFtpServerNONE" | Thread 65536
06/04/01 07:57:50: Thread ID : 65536 : DBIF:Execute: Ultimate SQL: exec DsStCPSelectByName "EcDsStFtpServerNONE" | Thread 65536
06/04/01 07:57:50: Thread ID : 65536 : DBIF:Fetched:[8.000000][EcDsStFtpServerNONE][1][10][FTP][][0][FTPA][NONE][4194304] | Thread 65536

– Spawn receptionist thread and register server in the database (log entries similar to the following).

06/04/01 07:57:50: Thread ID : 65536 : DsStReceptionist:BindSocketGetInfo: Port assigned is 13441 | Thread 65536
06/04/01 07:57:50: Thread ID : 65536 : myTransactionList[0]: use stmgtdb1_DEV07 | Thread 65536
06/04/01 07:57:50: Thread ID : 65536 : DBIF:Execute: Ultimate SQL: use stmgtdb1_DEV07 | Thread 65536
06/04/01 07:57:50: Thread ID : 65536 : myTransactionList[1]: exec DsStCPRegisterServer 8, 13441, "f2acg01" | Thread 65536

06/04/01 07:57:50: Thread ID : 65536 : DBIF:Execute: Ultimate SQL: exec DsStCPRegisterServer 8, 13441, "f2acg01" | Thread 65536
06/04/01 07:57:50: Thread ID : 65536 : Ftp:Ctor: EcDsStFtpServerNONE | Thread 65536
06/04/01 07:57:50: Thread ID : 65536 : myTransactionList[0]: use stmgtdb1_DEV07 | Thread 65536
06/04/01 07:57:50: Thread ID : 65536 : DBIF:Execute: Ultimate SQL: use stmgtdb1_DEV07 | Thread 65536
06/04/01 07:57:50: Thread ID : 65536 : myTransactionList[1]: exec DsStCPSelectById 8 | Thread 65536
06/04/01 07:57:50: Thread ID : 65536 : DBIF:Execute: Ultimate SQL: exec DsStCPSelectById 8 | Thread 65536
06/04/01 07:57:50: Thread ID : 65536 : DBIF:Fetched:[8.000000][EcDsStFtpServerNONE][1][10][FTP][f2acg01][13441][FTPA][NONE][4194304] | Thread 65536
06/04/01 07:57:50: Thread ID : 65536 : Ftp:Ctor: Leaving | Thread 65536

– Spawn service threads (log entries similar to the following).

06/04/01 07:57:50: Thread ID : 65536 : Ftp:Startup: temperature = cold | Thread 65536
06/04/01 07:57:50: Performing startup processing | Thread 65536
06/04/01 07:57:50: Thread ID : 65536 : Spawning service threads | Thread 65536
06/04/01 07:57:50: Thread ID : 65536 : BR:GetThreadPoolConfiguration | Thread 65536
06/04/01 07:57:51: Thread ID : 65536 : myTransactionList[0]: use stmgtdb1_DEV07 | Thread 65536
06/04/01 07:57:51: Thread ID : 65536 : DBIF:Execute: Ultimate SQL: use stmgtdb1_DEV07 | Thread 65536
06/04/01 07:57:51: Thread ID : 65536 : myTransactionList[1]: exec DsStSTCSelectForServer 8, "ThreadPool" | Thread 65536
06/04/01 07:57:51: Thread ID : 65536 : DBIF:Execute: Ultimate SQL: exec DsStSTCSelectForServer 8, "ThreadPool" | Thread 65536
06/04/01 07:57:51: Thread ID : 65536 : DBIF:Fetched:[ThreadPool][10][0][0][0][0][10] | Thread 65536
06/04/01 07:57:51: Thread ID : 65536 : 3_2709893_0757-1125858625_155062001_f2acg01:FTPA: BR:GetThreadPoolConfiguration Returning | Thread 65536
06/04/01 07:57:51: Thread ID : 65536 : Ftp: Spawning a service thread | Thread 65536
06/04/01 07:57:51: Starting a new service thread | Thread 65536
06/04/01 07:57:51: Thread ID : 65536 : Ftp: Spawning a service thread | Thread 65536
06/04/01 07:57:51: 06/04/01 07:57:51: Thread ID : 65554 : Waiting for work | Thread 65554

06/04/01 07:57:51: Thread ID : 65554 : DsShTSStorage: creating the MutexVec for this thread
06/04/01 07:57:51: Thread ID : 65554 : Waking up manager thread | Thread 65554

.
.
.

06/04/01 07:57:51: Starting a new service thread | Thread 65536
06/04/01 07:57:51: Thread ID : 65536 : Ftp: Spawning a service thread | Thread 65536
06/04/01 07:57:51: Starting a new service thread | Thread 65536
06/04/01 07:57:51: Thread ID : 65559 : Waiting for work | Thread 65559
06/04/01 07:57:51: Thread ID : 65560 : Waiting for work | Thread 65560
06/04/01 07:57:51: Thread ID : 65561 : Waiting for work | Thread 65561

– Process Restart Notification for server restart ("Ready to accept requests") (log entries similar to the following).

06/04/01 07:57:51: Thread ID : 65536 : myTransactionList[0]: use stmgtdb1_DEV07 | Thread 65536
06/04/01 07:57:51: Thread ID : 65536 : DBIF:Execute: Ultimate SQL: use stmgtdb1_DEV07 | Thread 65536
06/04/01 07:57:51: Thread ID : 65536 : myTransactionList[1]: BEGIN TRANSACTION OUTER_278888352 | Thread 65536
06/04/01 07:57:51: Thread ID : 65536 : DBIF:Execute: Ultimate SQL: BEGIN TRANSACTION OUTER_278888352 | Thread 65536
06/04/01 07:57:51: Thread ID : 65536 : myTransactionList[2]: exec DsStGRRestartNotification "10_2709893_0757-1125858625_15506 2001_f2acg01:FTPA:Server restart", "EcDsStFtpServerNONE", "cold" | Thread 65536
06/04/01 07:57:51: Thread ID : 65536 : DBIF:Execute: Ultimate SQL: exec DsStGRRestartNotification "10_2709893_0757-1125858625 _155062001_f2acg01:FTPA:Server restart", "EcDsStFtpServerNONE", "cold" | Thread 65536
06/04/01 07:57:52: Thread ID : 65536 : DBIF:Fetched:[] | Thread 65536
06/04/01 07:57:52: Thread ID : 65536 : DBIF:Fetched:[8.000000] | Thread 65536
06/04/01 07:57:52: Thread ID : 65536 : DBIF:Execute: Ultimate SQL: COMMIT TRANSACTION OUTER_278888352 | Thread 65536
06/04/01 07:57:52: Thread ID : 0 : No servers to awaken -- get status | Thread 0
06/04/01 07:57:52: Thread ID : 65536 : Spawning manager thread | Thread 65536
06/04/01 07:57:52: Ready to accept requests | Thread 65564

– Check queue for requests ("Waiting for an event" means there is nothing else in the queue.) (log entries similar to the following).

06/04/01 07:57:52: Thread ID : 65564 : BR:ProcessCancelledRequests | Thread 65564
06/04/01 07:57:52: Thread ID : 65564 : DsShTSStorage: creating the MutexVec for this thread
06/04/01 07:57:52: Thread ID : 65564 : myTransactionList[0]: use stmgtdb1_DEV07 | Thread 65564
06/04/01 07:57:52: Thread ID : 65564 : DBIF:Execute: Ultimate SQL: use stmgtdb1_DEV07 | Thread 65564
06/04/01 07:57:52: Thread ID : 65564 : myTransactionList[1]: exec DsStGRSelectCancelled 8 | Thread 65564
06/04/01 07:57:52: Thread ID : 65564 : DBIF:Execute: Ultimate SQL: exec DsStGRSelectCancelled 8 | Thread 65564
06/04/01 07:57:52: Thread ID : 65564 : BR:ProcessCancelledRequests: Nothing cancelled | Thread 65564
06/04/01 07:57:52: Thread ID : 65564 : BR:ProcessCancelledRequests Returning | Thread 65564
06/04/01 07:57:52: Thread ID : 65564 : Ftp: Getting next request | Thread 65564
06/04/01 07:57:52: Thread ID : 65564 : myTransactionList[0]: use stmgtdb1_DEV07 | Thread 65564
06/04/01 07:57:52: Thread ID : 65564 : DBIF:Execute: Ultimate SQL: use stmgtdb1_DEV07 | Thread 65564
06/04/01 07:57:52: Thread ID : 65564 : myTransactionList[1]: exec DsStFRGetNextRequest 8 | Thread 65564
06/04/01 07:57:52: Thread ID : 65564 : DBIF:Execute: Ultimate SQL: exec DsStFRGetNextRequest 8 | Thread 65564
06/04/01 07:57:52: Thread ID : 65564 : GetNextRequest: No requests available | Thread 65564
06/04/01 07:57:52: Thread ID : 65564 : Waiting for an event | Thread 65564

- The log file for the server from which the call originated may indicate a problem completing a connection. The log should indicate successful awakening of a remote host, with entries similar to the following:

06/04/01 07:57:52: Thread ID : 65536 : DsStPatron:AwakenRemoteServer: Hostname - f2acg01 | Thread 65536
06/04/01 07:57:52: Thread ID : 65536 : DsStPatron:AwakenRemoteServer: Port Number - 13441 | Thread 65536
06/04/01 07:57:52: Thread ID : 65536 : Patron: Creating new entry for EcDsStFtpServerNONE | Thread 65536
06/04/01 07:57:52: Thread ID : 65536 : Trying gethostbyname_r() 0 of 5 attempts | Thread 65536

06/04/01 07:57:52: Thread ID : 65536 : Waking up EcDsStFtpServerNONE | Thread 65536

and should indicate completion of a connection to the called server, with entries similar to the following:

06/04/01 07:57:52: Thread ID : 65553 : DsStReceptionist:WaitForConnections: A connection has been accepted | Thread 65553
06/04/01 07:57:52: Thread ID : 65564 : BR:ProcessCancelledRequests | Thread 65564
06/04/01 07:57:52:
06/04/01 07:57:52: Thread ID : 65553 : Waking up manager thread | Thread 65553
06/04/01 07:57:52: Thread ID : 65564 : : 06/04/01 07:57:52: read ID : 7:57:52: DsShTSStorageDsShTSStorage: creating the MutexVec for this thread: creating the MutexVec for this thread665553: 53 : DsShTSStorage: creating the MutexVec for this thread
06/04/01 07:57:52: Thread ID : 65564 : myTransactionList[0]: use stmgtdb1_DEV07 | Thread 65564
06/04/01 07:57:52: Thread ID : 65564 : DBIF:Execute: Ultimate SQL: use stmgtdb1_DEV07 | Thread 65564
06/04/01 07:57:52: Thread ID : 65564 : myTransactionList[1]: exec DsStGRSelectCancelled 8 | Thread 65564
06/04/01 07:57:52: Thread ID : 65564 : DBIF:Execute: Ultimate SQL: exec DsStGRSelectCancelled 8 | Thread 65564
06/04/01 07:57:52: Thread ID : 65564 : BR:ProcessCancelledRequests: Nothing cancelled | Thread 65564
06/04/01 07:57:52: Thread ID : 65564 : BR:ProcessCancelledRequests Returning | Thread 65564.

- This procedure is applicable for reviewing logs for different types of errors and events on system servers.

5 Exit the log file (e.g., from **pg**, type **q** and then press the **Return/Enter** key).

A Special Case: Checking the Request Manager Server Debug Log

The Request Manager server in the Storage Management computer software configuration item of the Data Server Subsystem processes requests from external clients (processes outside of Storage Management). Requests between Storage Management servers are passed directly from one server to another.

- Requests that require one of the Storage Management servers to perform processing are checkpointed (except requests that can be serviced solely through SQL).
  - Checkpointing involves recording the request's state (e.g., "checkpointed," "failed," "completed") in the database to assist in error recovery.

- Requests that can be serviced solely through SQL are considered "trivial" requests.
  - Trivial requests are not checkpointed.
  - Examples include attaching to a staging disk, getting capacity, and getting block size.
  - Trivial requests submitted from outside Storage Management are serviced by the Request Manager server.
  - Trivial requests originating within Storage Management are passed directly from the client to the database server.

The Request Manager server (like other Storage Management servers) can manage several concurrent activities. This is accomplished through the use of threads. There are several different kinds of threads:

- Manager thread.
  - One per Storage Management server.
  - Responsible for dequeuing requests and assigning them to service threads.
  - Checks for cancelled requests.
- Service thread.
  - Multiple threads per Storage Management server.
  - Responsible for the actual servicing of requests.
  - Logs all progress including all changes of request state.
  - Notifies submitter when request has been completed.
- Receptionist thread.
  - One per Storage Management server.
  - Registers the server as "up" in the database.
  - Sits on a socket, waiting for connections from other Storage Management servers.
  - Unregisters the server at shutdown.
- Inbound RPC thread.
  - Spawned by a request from a Storage Management client.
  - Hands off the request to the manager thread and waits for completion of the request.

- Housekeeper thread.
  - Watches for completed requests that have not previously been seen and processed.

Information concerning Request Manager server processing of requests (identified by thread) is recorded in the Request Manager server debug log (assuming some level of debug log recording is specified in the Registry database).

Trivial requests typically involve the following types of activities:

- Inbound RPC thread appears with a request.
- Manager thread dequeues the request and assigns it to a service thread.
- Service thread recognizes the thread as "trivial."
  - A "No checkpointing required -- going straight to responded" message is recorded in the Request Manager server debug log.
- Service thread executes the database transaction for results.
  - When the request is completed, a "Done servicing" message is recorded in the Request Manager server debug log.
  - If the request fails, an "Unable to service" message is recorded in the Request Manager server debug log.
- Service thread hands the results to the inbound RPC thread.
  - A "Notifying the client" message is recorded in the Request Manager server debug log.
- Inbound RPC thread silently returns to the client with the results.

Non-trivial requests are forwarded to the appropriate Storage Management server (e.g., EcDsStFtpServer, EcDsStStagingDiskServer, EcDsStArchiveServer) for processing.

- Some of the same types of entries are made in the Request Manager server debug log for non-trivial requests as for trivial requests.
  - For example:
    - "Waking up service thread" (Request Manager is preparing to process the request).
    - "Done servicing" (request processing has been completed).
    - "Unable to service" (the request has failed).

- Although some trivial requests include "token" statements, tokens are characteristic of non-trivial requests.
  - A token includes request information that varies with the type of operation to be performed.
  - For example, a token for an ftp request might include the following types of data:
    - Stored procedure (e.g., DsStFRInsert) [other types of stored procedures include DsStSDRInsert and DsStGRMapLogicalArchiveId].
    - RPC ID (e.g., RPCId=1821_535_1109-1124464729_171062001_x0ins01.xdc.ecs.nasa.gov:SBSVSDSV1DSDD1DSDD4:).
    - Username.
    - Encrypted password.
    - Host.
    - Source path.
    - Destination path.
    - External request ID.
    - Server name (e.g., EcDsStFtpServerNONE) [other types of operations might involve the EcDsStStagingDiskServerDRP1 for example].
    - Type of operation (e.g., FtpPush) [other types of operations include ArRetrieve, SDAllocateDisk, SDLinkFile].
    - Submitter (e.g., DSDD) [other types of operations might involve SDSV].
    - Priority.
  - The server to which the request was sent is identified by name (ServerName).
  - Transaction ID is embedded in the RPC ID (the portion before the first colon in the RPC ID).

A "transaction" may involve multiple operations on a host or several hosts. Consequently, multiple threads may be used on each relevant host. Use the following procedure to check the debug log for the Storage Management Request Manager server.

Check the Request Manager Server Debug Log

1 Log in to the Sun internal server host (e.g., e0acs11, g0acs11, l0acs03, and n0acs04).

2 To change to the logs directory, type **cd /usr/ecs/<*MODE*>/CUSTOM/logs** then press the **Return/Enter** key.

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

3 Type **pg *filename*** then press the **Return/Enter** key.

- ***filename*** refers to the appropriate Request Manager debug log.
- For example:

  pg EcDsStRequestManagerServerDebug.log

- The content of the first page of the specified file is displayed.
- Although this procedure has been written for the **pg** command, any UNIX editor or visualizing command (e.g., **vi**, **view**, **more**) can be used to review the log file.

4 At the **:** prompt type **/*date time*** then press the **Return/Enter** key.

- ***date time*** refers to the approximate date and time of the problem.
  - For example:

    /06/18/01 12:17:31

- The file is searched for the specified text.
  - If the specified text is in the log file, the following type of response is displayed.

    ...skipping forward
    06/18/01 12:17:31: Thread ID : 105 : DsShTSStorage: creating the MutexVec for this thread
    […]

  - If the specified text is not in the log file, the following type of response is displayed.

    Pattern not found:

  - If the specified text is not in the log file, verify the following aspects of Steps 3 and 4:
    - Date and time were entered correctly (Step 4).
    - Proper file was opened (Step 3).

5 At the **:** prompt type **/Unable to service** then press the **Return/Enter** key.

- **pg** searches the file for the specified text.
  - If the specified text is in the log file, the following type of response is displayed.

    ...skipping forward
    2:IngestRQ409GR1 Unable to service | Thread 52
    […]

- If the specified text is not in the log file, the following type of response is displayed.

  Pattern not found:

- If the specified text is in the file, go to Step 7.
- If the specified text is not in the file, go to Step 6.

6 Examine the contents of the log file to determine which thread is associated with the problem being investigated.

- The following **pg** commands (at the **:** prompt) are useful:
  - ***n*** then **Return/Enter** (go to Page *n*).
  - **Return/Enter** or **+1** then **Return/Enter** (go down to the next page).
  - **-1** then **Return/Enter** (go back to the preceding page).
  - **+*n*** then **Return/Enter** (go down *n* number of pages).
  - **-*n*** then **Return/Enter** (go back *n* number of pages).
  - **+*n*l** then **Return/Enter** (go down *n* number of lines).
  - **-*n*l** then **Return/Enter** (go back *n* number of lines).
  - **q** then **Return/Enter** (exit from **pg**).

7 At the **:** prompt type the appropriate text (depending on the direction of the desired search) then press the **Return/Enter** key:

8 To search back toward the beginning of the file, type **^Waking up service thread *n*^** and then press **Return/Enter**.

9 To search toward the end of the file, type **/Waking up service thread *n*** and then press **Return/Enter**.

- For example:

  ^Waking up service thread 52^

  - The file is searched back toward the beginning of the file for the specified text.

- If the specified text is in the log file, the following type of response is displayed.

  ...skipping backward
  06/18/01 12:17:31: Thread ID : 102 : Waking up service thread 52 | Thread 102
  […]

- If the specified text is not in the log file, the following type of response is displayed.

  Pattern not found:

- The entries "Waking up service thread *n*" and "Unable to service | Thread *n*" bracket the thread servicing in which an error occurred.

NOTE: Thread IDs are reused frequently. There are likely to be many processes with the same thread ID in any particular log file. It is important to follow the correct instance of the thread.

NOTE: It is likely that the Request Manager would try again to process a failed request. Subsequent request processing may use the same thread ID or a different thread ID. However, it would involve the same transaction ID.

- A "No checkpointing required -- going straight to responded" entry associated with the thread ID indicates that the request is "trivial."

10 At the **:** prompt type **/SEARCHING** then press **Return/Enter**.

- The file is searched for the specified text.
  - If the specified text is in the log file, the following type of response is displayed.

...skipping forward
06/18/01 12:17:31: Thread ID : 52 : SEARCHING FOR: 30148 (Found) | Thread 52
06/18/01 12:17:31: Thread ID : 52 : SEARCHING FOR: 30148 (Found) | Thread 52
06/18/01 12:17:31: Thread ID : 52 : DsStStoredProcedures::Execute - ERROR: Could not execute stored procedure | Thread 52
06/18/01 12:17:31: Thread ID : 52 : Error encountered in stored procedure | Thread 52
06/18/01 12:17:31: Thread ID : 52 : DBIF:Execute: Ultimate SQL: ROLLBACK TRANSACTION OUTER_7077776 | Thread 52
06/18/01 12:17:32: Thread ID : 52 : 1_4501810_1217-1124633447_169062001_p0icg01.pvc.ecs.nasa.gov:IPOBIPOB1INRM1IGSA15:IngestRQ409GR1 Done servicing | Thread 52
06/18/01 12:17:32: Thread ID : 52 : 1_4501810_1217-1124633447_169062001_p0icg01.pvc.ecs.nasa.gov:IPOBIPOB1INRM1IGSA15:IngestRQ409GR1 Unable to service | Thread 52
06/18/01 12:17:32: Thread ID : 52 : 1_4501810_1217-1124633447_169062001_p0icg01.pvc.ecs.nasa.gov:IPOBIPOB1INRM1IGSA15:IngestRQ409GR1 Marked as unassigned | Thread 52
06/18/01 12:17:32: Thread ID : 52 : 1_4501810_1217-1124633447_169062001_p0icg01.pvc.ecs.nasa.gov:IPOBIPOB1INRM1IGSA15:IngestRQ409GR1 Notifying the client | Thread 52
06/18/01 12:17:32: Thread ID : 52 : Waiting for work | Thread 52
06/18/01 12:17:32: Thread ID : 52 : Waking up manager thread | Thread 52
[…]

- In the preceding example the expression **SEARCHING** is associated with Thread ID 52.
- The context of the SEARCHING statement indicates the type and source of the problem; in this case there appears to be a problem executing a stored procedure.

– If the specified text is not in the log file, the following type of response is displayed.

Pattern not found:

11 If the expression **SEARCHING** is not associated with the specified thread in the lines displayed, repeat Step 8.

12 If necessary, at the **:** prompt type **-2l** [lower-case letter l] then press the **Return/Enter** key.

- **pg** simulates scrolling the screen backward two lines (or any other number of lines that is typed at the prompt).
  – The file is redisplayed to include the two lines that preceded the page previously displayed.
  – For example:

...skipping backward
06/18/01 12:17:31: Thread ID : 52 : DBIF:Execute: Ultimate SQL: exec DsStSDAttachDisk "/usr/ecs/TS2/CUSTOM/pdps/x0spg01/data/DpPrRm/x0spg01_disk", "SDSV", 0 | Thread 52
06/18/01 12:17:31: Thread ID : 52 : SEARCHING FOR: 30148 (Found) | Thread 52
06/18/01 12:17:31: Thread ID : 52 : SEARCHING FOR: 30148 (Found) | Thread 52
06/18/01 12:17:31: Thread ID : 52 : DsStStoredProcedures::Execute - ERROR: Could not execute stored procedure | Thread 52
06/18/01 12:17:31: Thread ID : 52 : Error encountered in stored procedure | Thread 52
[…]

  – The additional lines preceding "SEARCHING FOR" in the example indicate that the stored procedure in which the error was encountered is DsStSDAttachDisk.

13 At the **:** prompt type **q** then press the **Return/Enter** key.

- **pg** exits from the Request Manager server debug log file.

14 If the request is a trivial request, go to Step 22.

15 If the request is a non-trivial request, open a separate UNIX window.

- The results of related operations on the server involved in performing copy or ftp functions for the transaction are going to be checked in a separate UNIX window.

16 In the new UNIX window log in to the appropriate server host (e.g., e0drg11, g0drg01, l0drg01, n0drg01) for the server involved in performing copy or ftp functions for the transaction.

17 At the shell prompt type **grep '*TransactionId*' *filename* | grep 'LogProgress'** then press the **Return/Enter** key.

- For example:

  grep 'af610628-' EcDsStArchiveServerDebug.log | grep '*LogProgress*'

- ***filename*** refers to the name of the log file for the process involved in performing copy or ftp functions for the transaction.
- ***TransactionId*** refers to the Transaction ID associated with the applicable request.

- In this example af610628-1dd1-11b2-a047-af3a589fd88e is the relevant Transaction ID.
  - However, usually it is not necessary to use the entire Transaction ID in the command; a representative sample (e.g., af610628- from the example) should be sufficient.
  - References to other Transaction IDs and entries that do not contain the string "LogProgress" are filtered out so references to the specified Transaction ID that contain the string "LogProgress" are the only log entries displayed.
    - The string "LogProgress" is a filter for references to stored procedure DsStGRLogProgress.
  - Progress is logged for copy and ftp input/output at each block.
  - The following type of response is displayed:

    06/26/01 12:46:00: Thread ID : 65674 : myTransactionList[1]: exec DsStGRLogProgress "af610628-1dd1-11b2-a047-af3a589fd88e:PDPSSDSV1DSDD1DSDD10DSDD1DSDD1:MoPGE02#sy14182000TS2SC:MOD03.001:55732", 0, 1, "files" | Thread 65674
    06/26/01 12:46:00: Thread ID : 65674 : DBIF:Execute: Ultimate SQL: exec DsStGRLogProgress "af610628-1dd1-11b2-a047-af3a589fd88e:PDPSSDSV1DSDD1DSDD10DSDD1DSDD1:MoPGE02#sy14182000TS2SC:MOD03.001:55732", 0, 1, "files" | Thread 65674
    06/26/01 12:46:43: Thread ID : 65674 : : 06/26/01 12:46:43: read ID : 2:46:43: myTransactionmyTransactionList[1]: exec DsStGRLogProgress "af610628-1dd1-11b2-a047-af3a589fd88e:PDPSSDSV1DSDD1DSDD10DSDD1DSDD1:MoPGE02#sy14182000TS2SC:MOD03.001:55732", 60, 60, "MB"List[1]: exec

DsStGRLogProgress "af610628-1dd1-11b2-a047-af3a589fd88e:PDPSSDSV1DSDD1DSDD10DSDD1DSDD1:MoPGE02#sy14182000TS2SC:MOD03.001:55732", 60, 60, "MB"65714read 65674 : 74
06/26/01 12:46:43: Thread ID : 65674 : DBIF:Execute: Ultimate SQL: exec DsStGRLogProgress "af610628-1dd1-11b2-a047-af3a589fd88e:PDPSSDSV1DSDD1DSDD10DSDD1DSDD1:MoPGE02#sy14182000TS2SC:MOD03.001:55732", 60, 60, "MB"0DBIF:Execute: Ultimate SQL: exec DsStGRLogProgress "af610628-1dd1-11b2-a047-af3a589fd88e:PDPSSDSV1DSDD1DSDD10DSDD1DSDD1:MoPGE02#sy14182000TS2SC:MOD03.001:55732", 60, 60, "MB"06/26/01 12:46:43:
6/26/01 12:46:43: | Thread : 65714read 65674 : 74

- If no progress is indicated, go to Step 22.

18 Click in the UNIX window for the Sun internal server host.

19 Type **grep '*TransactionId*' *filename* | grep 'Done servicing'** then press **Return/Enter**.

- ***filename*** refers to the appropriate Request Manager debug log.
- For example:

  grep 'af610628-' EcDsStRequestManagerServerDebug.log | grep 'Done servicing'

- If the operation has been completed, the following type of response is displayed:

  06/26/01 12:46:00: Thread ID : 52 : af610628-1dd1-11b2-a047-af3a589fd88e:PDPSSDSV1DSDD1DSDD10DSDD1DSDD1:MoPGE02#sy14182000TS2SC:MOD03.001:55732 Done servicing | Thread 52
  06/26/01 12:46:44: Thread ID : 52 : af610628-1dd1-11b2-a047-af3a589fd88e:PDPSSDSV1DSDD1DSDD10DSDD1DSDD1:MoPGE02#sy14182000TS2SC:MOD03.001:55732 Done servicing | Thread 52
  06/26/01 12:46:45: Thread ID : 52 : af610628-1dd1-11b2-a047-af3a589fd88e:PDPSSDSV1DSDD1DSDD2DSDD1DSDD3:MoPGE02#sy14182000TS2SC:MOD03.001:55732 Done servicing | Thread 52
  06/26/01 12:46:47: Thread ID : 52 : af610628-1dd1-11b2-a047-af3a589fd88e:PDPSSDSV1DSDD1DSDD2DSDD1DSDD3:MoPGE02#sy14182000TS2SC:MOD03.001:55732 Done servicing | Thread 52
  06/26/01 12:46:47: Thread ID : 52 : af610628-1dd1-11b2-a047-af3a589fd88e:PDPSSDSV1DSDD1DSDD2DSDD1DSDD7:MoPGE02#sy14182000TS2SC:MOD03.001:55732 Done servicing | Thread 52
  06/26/01 12:46:50: Thread ID : 52 : af610628-1dd1-11b2-a047-af3a589fd88e:PDPSSDSV1DSDD1DSDD2DSDD1DSDD7:MoPGE02#sy14182000TS2SC:MOD03.001:55732 Done servicing | Thread 52
  06/26/01 12:46:51: Thread ID : 52 : af610628-1dd1-11b2-a047-af3a589fd88e:PDPSSDSV1DSDD1DSDD4:MoPGE02#sy14182000TS2SC:MOD0

3.001:55732 Done servicing | Thread 52
06/26/01 12:46:56: Thread ID : 52 : af610628-1dd1-11b2-a047-af3a589fd88e:PDPSSDSV1DSDD1DSDD4:MoPGE02#sy14182000TS2SC:MOD03.001:55732 Done servicing | Thread 52
06/26/01 12:46:56: Thread ID : 52 : af610628-1dd1-11b2-a047-af3a589fd88e:PDPSSDSV1DSDD1DSDD8:MoPGE02#sy14182000TS2SC:MOD03.001:55732 Done servicing | Thread 52
06/26/01 12:46:59: Thread ID : 52 : af610628-1dd1-11b2-a047-af3a589fd88e:PDPSSDSV1DSDD1DSDD8:MoPGE02#sy14182000TS2SC:MOD03.001:55732 Done servicing | Thread 52

  - The statement "Done servicing" shows that the operation has been completed; however, it provides no indication as to whether the operation succeeded or failed.
  - If "Done servicing" is followed by "Unable to service," (as described in Step 19) the operation failed.

- If the operation has not been completed, no file entries are displayed (the UNIX prompt is displayed).
  - It may just be slow to complete.
- If the operation has been completed, go to Step 19.
- If the operation has not been completed, go to Step 20.

20 Type **grep '*TransactionId*' *filename* | grep 'Unable to service'** then press the **Return/Enter** key.

- ***filename*** refers to the appropriate Request Manager debug log.
- For example:

  grep '2a7d4168-' EcDsStRequestManagerServerDebug.log | grep 'Unable to service'
- If the request has failed, the following type of response is displayed:

  06/26/01 12:56:22: Thread ID : 52 : 2a7d4168-1dd2-11b2-8c52-99d0f708dce5:PDPSSDSV1:MoPGE02#sy14182000TS2MOD02OBC Unable to service | Thread 52
  06/26/01 12:56:22: Thread ID : 52 : 2a7d4168-1dd2-11b2-8c52-99d0f708dce5:PDPSSDSV4:MoPGE02#sy14182000TS2MOD02OBC Unable to service | Thread 52

  - If the operation has failed, return to Step 7.
- If the operation has not failed, no file entries are displayed (the UNIX prompt is displayed).

21 If the operation has not failed, at the shell prompt type **tail -f *filename* | grep '*TransactionId*'** and then press the **Return/Enter** key.

- ***filename*** refers to the appropriate Request Manager debug log.
- ***TransactionId*** refers to the Transaction ID associated with the applicable request.
- For example:

  tail -f EcDsStRequestManagerServerDebug.log | grep 'af610628-'

- If new entries are being posted to the log, the operation has not finished yet.
  - If the same entries continue to be repeated over and over, it may be necessary to restart the server.
- If it is necessary to exit from a tailed log, type **ctrl-c** (while holding down the **Control** key, press **c**).

22 If the operation has not finished yet, monitor the tailed log for a while.

- If the operation does not seem to finish (i.e., if entries continue to be made to the tailed log) after a reasonable period of time (e.g., 30 minutes), it may be necessary to restart the Request Manager server.
- If it is necessary to exit from a tailed log, type **ctrl-c** (while holding down the **Control** key, press **c**).

23 If problems were detected in the Request Manager server debug log and/or the log file for the process involved in performing copy or ftp functions for the transaction, it may be necessary to restart the server(s) performing those functions.

- If server restart does not resolve the problem, it is appropriate to notify the Help Desk and prepare a Trouble Ticket.

24 If no problems were detected in the Request Manager server debug log or the log file for the process involved in performing copy or ftp functions for the transaction, check the Science Data Server log files.

- For detailed instructions refer to the **Check Server Log Files** procedure (previous section of this lesson).

Checking the *tac* Log

Each day a current **tac_00** log on the FSMS host records interactions between AMASS and ACSLS. This log can provide helpful information in troubleshooting problems manifested in those interactions. Use the following procedure to check the **tac_00** log.

Check the *tac* Log

---

1 Log in as **amass** at the FSMS host.

2 Type **cd /usr/amass/logs/tac** and then press the **Return/Enter** key.

- The working directory is changed to **/usr/amass/logs/tac**.

3 Use the current tac log to investigate possible problems in communication between AMASS and ACSLS. To view the current tac log, type **pg tac_00** and then press the **Return/Enter** key.

- The first page of the log file is displayed; additional sequential pages can be displayed by pressing the **Return/Enter** key at the **:** prompt.
- Although this procedure has been written for the **pg** command, any UNIX editor or visualizing command (e.g., **vi**, **view**, **more**, **tail**) can be used to review the log file.
- The log contains entries related to activities and communications associated with actions by AMASS to direct ACSLS robotic activities; the entries should appear in format similar to the following sample:

```
Sep 24 09:49:42 p0drg01 amass LIBSCHED3[7215638]:
E7003(16)<00000>:xdiStk2749:  STK Response received; Status: 0
Sep 24 09:49:42 p0drg01 amass LIBSCHED3[7215638]:
E7003(16)<00000>:xdiStk2797:  ACSLS ACK response received
Sep 24 09:49:42 p0drg01 amass LIBSCHED3[7215638]:
E7003(16)<00000>:xdiStk2742:  Waiting for ACSLS response
Sep 24 09:49:51 p0drg01 amass LIBSCHED3[7215638]:
E7003(16)<00000>:xdiStk2749:  STK Response received; Status: 0
Sep 24 09:49:51 p0drg01 amass LIBSCHED3[7215638]:
E7003(16)<00000>:xdiStk2873:  ACSLS final response received
Sep 24 09:49:51 p0drg01 amass LIBSCHED3[7215638]:
E7003(16)<00000>:xdiStk2876:  1 network packets transferred
Sep 24 10:18:52 p0drg01 amass LIBSCHED1[7215870]:
E7003(16)<00000>:xdiArch39:  Archive index : 0
Sep 24 10:18:52 p0drg01 amass LIBSCHED1[7215870]:
E7003(16)<00000>:xdiStk486:  Media Id = P10011, Drive index = 0
Sep 24 10:18:52 p0drg01 amass LIBSCHED1[7215870]:
E7003(16)<00000>:xdiStk533:  Sending a mount command
Sep 24 10:19:32 p0drg01 amass LIBSCHED1[7215870]:
E7003(16)<00000>:xdiArch39:  Archive index : 0
```

Sep 24 10:19:32 p0drg01 amass LIBSCHED1[7215870]:
E7003(16)<00000>:xdiStk486: Media Id = P20676, Drive index = 1
Sep 24 10:19:32 p0drg01 amass LIBSCHED1[7215870]:
E7003(16)<00000>:xdiStk533: Sending a mount command
Sep 24 10:34:56 p0drg01 amass LIBSCHED1[7215870]:
E1043(7)<00000>:libsched3165: Idle Eject timer expired on volume 188 in drive 2.
Sep 24 10:35:07 p0drg01 amass LIBSCHED1[7215870]:
E7003(16)<00000>:xdiArch39: Archive index : 0
Sep 24 10:35:07 p0drg01 amass LIBSCHED1[7215870]:
E7003(16)<00000>:xdiStk686: Media Id = P20676, Drive index = 1
Sep 24 10:35:07 p0drg01 amass LIBSCHED1[7215870]:
E7003(16)<00000>:xdiStk719: Sending a dismount command
Sep 24 10:35:07 p0drg01 amass LIBSCHED1[7215870]:
E7003(16)<00000>:xdiStk2742: Waiting for ACSLS response
Sep 24 10:35:07 p0drg01 amass LIBSCHED1[7215870]:
E7003(16)<00000>:xdiStk2749: STK Response received; Status: 0
Sep 24 10:35:07 p0drg01 amass LIBSCHED1[7215870]:
E7003(16)<00000>:xdiStk2758: Error unexpected sequence number: 101 - expected sequence number: 109

- Examine the sections of the log with entries near the time of any problem being investigated, looking for messages that indicate whether there was successful communication between AMASS and ACSLS regarding mounting of a tape and transfer of information. It may be useful to search the log for occurrences of the word **fail** (while viewing the log with **pg**, **view**, **vi**, or other viewing/editing tool, type **/fail** and press the **Return/Enter** key).
- If the log indicates problems in communication between AMASS and ACSLS, it may be useful to use the **quedisplay** command to obtain the AMASS view of the queue and the **medialist** command to obtain the robot view. If these commands show discrepancies indicating a lack of synchrony between AMASS and ACSLS, it may be possible to re-establish that synchrony using the **mediamove** command [refer to the **Establish Synchrony Between *quedisplay* and *medialist* Using *mediamove*** procedure (previous section of this lesson)].
- The message "Error unexpected sequence number: 101 -expected sequence number: 109" is an artifact likely to be removed in releases of AMASS subsequent to Version 5.0.0 Revision 17 and does not reflect a real error.

Handling a Data Insertion Failure

Successful data insertion requires interactions among numerous servers, and the interactions are reflected in entries in the debug logs for those servers. Detection and initial isolation of a

problem that prevents successful insertion may require tracing events across multiple log files on different hosts. The following procedure is applicable.

Handle a Data Insertion Failure

---

1 At the Sun Consolidation Internal Server host (e.g., e0acs11, g0acs11, l0acs03, or n0acs04), review the debug log **EcDsScienceDataServerDebug.log**.

- For detailed instructions refer to the **Check Server Log Files** procedure (previous section of this lesson).
- Examine the section of the log with entries near the time of the problem, looking for error messages that indicate communication failure.
- If the log file entries indicate a communication problem, note the server(s) with which **there is impairment or disruption of communication.**

2 At the APC Server host (e.g., e0acg11, g0acg01, l0acg02, or n0acg01), review the debug log **EcDsStArchiveServerDebug.log**.

- For detailed instructions refer to the **Check Server Log Files** procedure (previous section of this lesson).
- Examine the section of the log with entries near the time of the problem, looking for error messages that indicate communication failure.
- If the log file entries indicate a communication problem, note the server(s) with which there is impairment or disruption of communication.

3 If Step 1 and/or Step 2 resulted in detection of a problem in the interaction of SDSRV and/or Archive Server with other servers, at the host(s) for those servers, review the server debug log(s). These logs may include:

EcDsStStagingDiskServerDebug.log (e.g., on e0acg11, g0acg01, l0acg02, or n0acg01).

EcDsStCacheManagerServerDebug.log (e.g., on e0acg11, g0acg01, l0acg02, or n0acg01).

EcDsStRequestManagerServerDebug.log [e.g., on e0acs11, g0acs11, l0acs03, or n0acs04; use the **Check the Request Manager Server Debug Log** procedure (previous section of this lesson)].

EcSbSubServerDebug.log (e.g., on e0acs11, g0acs11, l0acs03, or n0acs04).

- If there is evidence of requests not succeeding or other communication failure, it may be necessary to have System Administrators or Engineering Support personnel resolve the problem (e.g., restart affected servers, execute **EcCsIdPingServers**, ensure that the **Name Server** is up in the mode being used and that its debug log reflects appropriate look-up activity by the application servers, mount points are intact, and database access is not impaired).

NOTE: The next three steps address running the Check Archive script, **EcDsCheckArchive**. To run this script, it is necessary to enter eight database-specific parameters when prompted during the running of the script: STMGT SQL server name, STMGT database name, STMGT SQL server userID, STMGT SQL server database password, SDSRV SQL server name, SDSRV database name, SDSRV SQL server userID, and SDSRV database password. To facilitate the smooth execution of the script, the parameters may be set as environmental variables instead. The parameters are not readily available to most operators; therefore, you will need to obtain them from the Database Administrator or have the Database Administrator run the script for you, using steps 4 through 6.

4 On the APC Server host (e.g., e0acg11, g0acg01, l0acg02, or n0acg01), type **cd /usr/ecs/<*MODE*>/CUSTOM/utilities** and then press the **Return/Enter** key.

- The prompt reflects the directory change to **/usr/ecs/<*MODE*>/CUSTOM/utilities**.

5 Type **EcDsCheckArchive <*MODE*>** then press the **Return/Enter** key.

- The Check Archive script runs; the initially displayed information should be similar to the following:

```
==============================================================
 This script is designed to validate the Inventory
 against the Archive.

 The user must select the menu option associated with the
 Volume Group to be validated

        Please press [RETURN] to continue
==============================================================
```

6 Follow the on-screen prompts for the script, entering the necessary parameters.

- The script provides indication of any discrepancies between the presence of granules in the Archive and entries in the inventory (metadata). Note that the appearance of a discrepancy is not necessarily indication of a failure (e.g., if a granule has been deleted but the inventory database has not been cleaned up, there may be inventory entries for which there are no granules in the archive), but a problem may be indicated if a discrepancy is apparent for a granule that you just inserted. Note also that this script would not reveal a problem if you attempted to insert a granule which failed to get inserted and also had its metadata fail to be inserted into the inventory (i.e., no granule and no inventory entry = no discrepancy). Therefore, if the script reveals no discrepancies, it may still be useful to conduct a direct examination to determine if the granule has been inserted.

7 On the APC Server host (e.g., e0acg11, g0acg01, l0acg02, or n0acg01), type the directory change command **cd /dss_stk1/<*MODE*>/<*data_type_directory*>** and then press the **Return/Enter** key.

- The working directory is changed to **/dss_stk1/<*MODE*>/<*data_type_directory*>**.

8 Type **ls -al | grep "<*date*>"** where "<***date***>**"** is a three-letter abbreviation for the month followed by a number indicating the day (e.g., **"Apr 21"**) for the granule being inserted, and then press the **Return/Enter** key.

- If the inserted file is displayed, with date and time of entry, go to Step 9.
- If the inserted file is not displayed, have the Ingest/Distribution Technician insert the file again. If this succeeds (i.e., the file is now listed), go to Step 9; otherwise, perform the procedure for **Diagnosing/Investigating Write Errors**, Document 611-EMD-001, *Release 7.11 Mission Operation Procedures for the EMD Project*, Procedure 17.7.3).

9 Determine if the inserted file is reflected in the Inventory Database (Database Administrator function) by logging into Sybase on the Sun Consolidation Internal Server host (e.g., e0acs11, g0acs11, l0acs03, or n0acs04) and then selecting the data type for the granule being inserted.

- If the inserted file is reflected in the Inventory Database, go to Step 10.
- If the inserted file is not reflected in the Inventory Database, ensure that database access is not impaired (Database Administrator function).

10 Determine if the directory from/to which the copy is being made is visible on the machine being used; have the System Administrators or Engineering Support personnel check the mount points on the APC Server host (e.g., e0acg11, g0acg01, l0acg02, or n0acg01) and the Sun Consolidation Internal Server host (e.g., e0acs11, g0acs11, l0acs03, or n0acs04).

- If the mount points are OK, go to Step 11.
- If necessary, have the System Administrators or Engineering Support personnel re-establish the mount point(s).

11 If you inserted the file with the DSS Driver, go to Step 13. If you used Ingest to insert the file, on the Ingest Server host (e.g., e0icg11, g0icg01, l0acg02, or n0acg01) examine the **drp**- or **icl**-mounted staging directory to determine if a staging disk was created. To do this, first type **cd /usr/ecs/<*MODE*>/CUSTOM/drp/<*host*>/data/staging/cache** (or type **cd /usr/ecs/<*MODE*>/CUSTOM/icl/<*host*>/data/StagingArea/cache**), then press the **Return/Enter** key.

- The prompt reflects a change to the specified directory.
- Be sure that you are checking the correct mount/host. Most ingests use Ingest subsystem staging areas (i.e., **icl**), but others may not. Media ingest (e.g., from tape) typically involves staging in a **dip** area. For a polling ingest for data from EDOS, the polling directory may serve as the staging area. Some data are staged directly to working storage in the Data Server subsystem. If in doubt, consult Ingest/Archive personnel.

12 Type **ls -al | more** and then press the **Return/Enter** key.

- Any staging areas are listed in output similar to the following sample:

  -rw-rw-r-- 1 cmshared cmshared 499804704 Feb 6 11:49
  :SC:MOD000.001:11856:1.CCSDS
  -rw-rw-r-- 1 cmshared cmshared 320663592 Feb 6 11:51
  :SC:MOD000.001:11856:2.CCSDS
  -rw-rw-r-- 1 cmshared cmshared 540 Feb 6 11:51
  :SC:MOD000.001:11856:3.CCSDS.

- If a staging area for the inserted file appears at the end of the list, go to Step 13.
- If no staging area appears for the inserted file, it is possible that the ingest failed and that the staging area was immediately removed as part of clean-up. Check the Ingest logs (e.g., **EcInReqMgrDebug.log**, **EcInAutoDebug.log**, **EcInGranDebug.log**, or **EcInGranDebug**.log, depending on the type of Ingest) (refer to procedures for troubleshooting Ingest problems, Document 611-EMD-001, *Release 7.11 Mission Operation Procedures for the EMD Project*, Chapter 16) to determine if a staging disk was created. If no staging disk was created, it may be necessary to resolve a communications failure as described in Step 7.

13 Ensure that the Archive volume groups are set up correctly.

- For detailed instructions refer to the **Display Archive Path Information Using Storage Management GUIs** procedure (previous section of this lesson).

14 Ensure that the volume groups are on line.

- For detailed instructions refer to the **Display Volume Data Using *vollist*** procedure (previous section of this lesson).
- If the volume groups are set up correctly and their volumes are on line, and insertion still fails, it is appropriate to contact the Help Desk and prepare a trouble ticket (see Document 611-EMD-001, *Release 7.11 Mission Operation Procedures for the EMD Project*, Chapter 8).

Handling a Data Acquire Failure

As a first check, it is appropriate to determine if the acquire request appears in the list of System Requests on the Science Data Server GUI. If the acquire request does not appear on the Science Data Server GUI, you will need to determine where the breakdown occurred. Diagnosing an acquire failure requires detailed examination of the following system log files and directories associated with the process:

- Science Data Server log file (EcDsScienceDataServerDebug.log).
- Archive Server log file (EcDsStArchiveServerDebug.log).
- STMGT Request Manager Server log file (EcDsStRequestManagerDebug.log)

- Staging Area.
  - Presence of the relevant file.
  - Staging Disk log files (EcDsStStagingDiskServerDebug.log or EcDsStCacheManagerServerDebug.log).
  - Space available in the staging area.

In addition, note that a number of servers, clients, or other software running on various hosts, as reflected in Table 5, may be involved at various times in processing an acquire request. More information useful in troubleshooting may appear in related logs on these hosts. Use the following procedure for handling a failure to retrieve data.

Table 5. Hosts, Servers, Clients and Other Software Relevant to Acquires

HOST	SERVER/CLIENT/OTHER SOFTWARE
Sun Consolidation Internal Server (e.g., e0acs11, g0acs11, l0acs03, n0acs04)	DDIST Command Line Interface (EcDsDdDCLI) Distribution Server (EcDsDistribution Server) Science Data Server (EcDsScienceDataServer) Storage Management Request Manager (EcDsStRequestManagerServer) Staging Disk Server (EcDsStStagingDiskServer) Subscription Server (EcSbSubServer)
Access/Process Coordinators (APC) Server (e.g., e0acg11, g0acg01, l0acg02, n0acg01)	Archive Server (EcDsStArchiveServer) Cache Manager Server (EcDsStCacheManagerServer) Copy Server (EcDsStCopyServer.pl) FTP Server (EcDsStFtpServer) Pull Monitor Server (EcDsStPullMonitorServer) Staging Disk Server (EcDsStStagingDiskServer)
FSMS Server (e.g., e0drg11, g0drg01, l0drg01, n0drg01)	Archive Server (EcDsStArchiveServer) Cache Manager Server (EcDsStCacheManagerServer) Copy Server (EcDsStCopyServer.pl) FTP Server (EcDsStFtpServer) Staging Disk Server (EcDsStStagingDiskServer)

Handle a Data Acquire Failure

1 Launch the Science Data Server GUI.

- For detailed instructions refer to the **Launch DSS GUIs Using UNIX Commands** procedure (previous section of this lesson).

2 Click on the **System Requests** tab.

- The **System Requests** window is displayed.

3 Examine the requests displayed in the **System Management Requests** field to determine if SDSRV received the acquire request.

- If the number of request is large, the **Find** button and field below the **System Management Requests** field may be used to enter and search for information in the request, such as the Requester, or the **Filter . . .** button ca be used to launch a **System Management Filter Requests** window to limit the number of entries that appear in the **System Management Requests** field.

4 On the Sun Consolidation Internal Server host (e.g., e0acs11, g0acs11, l0acs03, or n0acs04), review the server logs **EcDsScienceDataServer.ALOG** and **EcDsScienceDataServerDebug.log**.

- For detailed instructions refer to the **Check Server Log Files** procedure (previous section of this lesson).
- Examine the section of the log with entries near the time of the problem, looking for messages that indicate whether the relevant file was successfully acquired.
- The **EcDsScienceDataServer.ALOG** file should contain entries identifying the file to be acquired by the ShortName of the corresponding ESDT; entries should be similar to the following:

  PID : 29168:MsgLink :0 meaningfulname :DsSrSessionExecuteRequestStart0
  Msg: Request ID b5156038-03d3-11d3-8d16-c676e82eaa77:????: executing: DsSrRequest (1): DsShSciRequestImp: [svr: ScienceDS, pri: NORMAL domain:]: (DsShSciCommandImp: service: INSERT num parameters: 3 category: Parameters are:
  -UnnamedPL[SHORTNAME(AST_L1BT) VERSIONID(001)
  --MAINGROUP[SHORTNAME(AST_L1BT) VERSIONID(001)

  METAFILEGROUP[METADATAFILE(/home/cmops/data/SCAST_L1BT.0011279.met)]
  ---DATAFILEGROUP[DATAFILE(/home/cmops/data/tahoe-north-middle.MTA)]
  ---DATAFILEGROUP[DATAFILE(/home/cmops/data/tahoe-north-middle.hdf)]]]
  WC)

- The **EcDsScienceDataServerDebug.log** file should contain entries regarding the acquire activity. The following types of messages should be included in the log file:

  05/06/99 12:52:01:
  About to execute Statement: exec ProcInsertReqDomain 2205, "UR:10:DsShESDT UR:UR:15:DsShSciServerUR:13:[VTC:DSSDSRV]:20:SC:AST_L1BT.001:2201"

  05/06/99 12:52:01:
  About to execute Statement: ProcInsertAcquireCmd 2206, 2205, 3, null, null,

"tester", "FtpPush", "MAIL", "FILEFORMAT", null, "jrattiga", "abc123", "t1dps04", "/home/jrattiga/push", null, null

- If the ShortName does not appear in the file, with a timestamp corresponding to the time of the attempted acquire, SDSRV may not be running, or may not be communicating with other servers. Have the System Administrator or Operations Controller check to be sure the server is up and, if appropriate, resolve the problem (e.g., restart affected servers, execute **EcCsIdPingServers**, ensure that the **Name Server** is up in the mode being used and that its debug log reflects appropriate look-up activity by the application servers, mount points are intact, and database access is not impaired).
- If the log file does contain entries for the relevant ShortName, and indicates that two files (the file and its associated metadata file) are being distributed, SDSRV has completed its role in the acquire. Go to the next step.
- If the ALOG contains the ShortName, and also contains an error showing that the data file time stamp does not match the time stamp required by the acquire, the data file needs to be removed from the Science Data Server and reinserted.
  - This is usually done using a script called DsDbCleanGranules.

5 To inspect the Archive Server log and Request Manager Server log for error messages associated with the acquire, on the Archive host (e.g., **e0drg11**, **g0drg01**, **l0drg01**, **n0drg01**), review the respective server logs (**EcDsStArchiveServerDebug.log**, **EcDsStRequestManagerServerDebug.log**).

- For detailed instructions refer to the **Check Server Log Files** and **Check the Request Manager Server Debug Log** procedures (previous section of this lesson).
- Examine the sections of the logs with entries near the time of the problem, looking for messages that indicate whether the Request Manager handled the request and whether the Archive Server log shows that the relevant file was successfully acquired.
- If the logs indicate that the relevant file was successfully acquired, go to the next step.
- If the file was not successfully acquired, it may be necessary to reboot AMASS [see the **Reboot AMASS** procedure (previous section of this lesson)] and investigate the possibility of read errors (see Document 611-EMD-001, *Release 7.11 Mission Operation Procedures for the EMD Project*, Procedure 17.7.4, **Diagnosing/Investigating Read** Errors).

6 To determine whether the file being acquired (or a link to it) and its associated metadata file arrived in the Data Distribution staging area, on the Sun internal server (e.g., **e0acs11, g0acs11, l0acs03,** or **n0acs04**) type **cd /usr/ecs/<*MODE*>/CUSTOM/drp/<*archivehost*>/data/staging/cache** and then press the **Return/Enter** key.

- The working directory is changed to the specified directory.

7 Type **ls -lrt** and then press the **Return/Enter** key.

- The contents of the directory are displayed.

8 Review the listing to determine whether the relevant file and its metadata file arrived in the staging area.

- The display should contain entries similar to the following:

  lrwxrwxr-x 1 cmshared cmshared 75 Apr 26 12:52 L7CPF19980518_19980518.01 -> /usr/ecs/TS1/CUSTOM/drp/raven/data/staging/cache/:SC:L7CPF.001:1427:1.ASCII
  -rw-rw-rw- 1 cmshared cmshared 14802 Apr 26 12:52 SCL7CPF.0011427.met
  -rw-rw-r-- 1 cmshared cmshared 111 Apr 26 13:01 staging.disk.filename.list
  -rw-rw-r-- 1 cmshared cmshared 2044 Apr 26 13:01 PACKING.LST.115124935248431

- If the relevant files were not successfully staged, the staging log files may reveal the cause; go to Step 9.
- If the relevant files were successfully staged, an acquire failure could be a result of problems with related servers or software (see Table 5). Have the System Administrator or Operations Controller ensure that the necessary hosts and servers are up.

9 To inspect the Staging Disk log for error messages associated with the acquire, on the APC Server host (e.g., **e0acg11**, **g0acg01**, **l0acg02**, **n0acg01**), review the server logs (e.g., **EcDsStStagingDiskServerDebug.log**; **EcDsStCacheManagerServer Debug.log**).

- For detailed instructions refer to the **Check Server Log Files** procedure (previous section of this lesson).
- Examine the section of each log with entries near the time of the problem, looking for messages that indicate whether the relevant files were successfully staged.
- If the relevant files were not successfully staged, the cause may be a lack of space in the staging area; go to Step 10.
- If the relevant files were successfully staged, an acquire failure could be a result of problems with related servers or software (see Table 5). Have the System

Administrator or Operations Controller ensure that the necessary hosts and servers are up.

10 To check the space available in the staging area, on the Sun internal server (e.g., **e0acs11, g0acs11, l0acs03**, or **n0acs04**) type **cd /usr/ecs/<*MODE*>/CUSTOM/drp/<*archivehost*>/data** and then press the **Return/Enter** key.

- The working directory is changed to the specified directory.

11 Type **df -k .** (be sure to include the "**.**") and then press the **Return/Enter** key.

- The filesystem, staging disk space capacity in kbytes, amount used, amount available, and percent of capacity are displayed, as in the following example:

  Filesystem kbytes used avail capacity Mounted on
  t1drg01:/usr/ecs/TS1/CUSTOM/drp/t1drg01/data
  225209856 173253056 51956800 77%
  /data1/ecs/TS1/CUSTOM/drp/t1drg01/data
  If there is not adequate space for staging the relevant files, it will be necessary to free up additional space (e.g., by purging expired files from cache).

Checksum Verification

The checksum verification utility (CVU) verifies that the data on a tape is not corrupted. It operates off a database that contains a list of tape volumes requiring verification. The database is populated by a script that is run once a day to identify tape volumes that were closed out that day. Each entry in the database contains the following types of data:

- Silo ID.
- Tape volume ID.
- Date/time the volume was closed out.
- Status.
- Date/time tape volume verification was started.
- Date/time tape volume verification was completed.

The status is one of the following values:

- Requires Verification.
- Verification in Progress.
- Verification Successful.
- Verification Failed.

Multiple CVUs can operate concurrently against the database. The maximum number is specified by the value assigned to the NUM_CONCURRENT_CHECKSUMS configuration parameter. The normal range is one to five.

The CVU is comprised of two executable components. The first component takes two parameters (i.e., a silo ID and the number of tape volumes to verify) as input. The second CVU component accepts a silo ID and tape volume ID as input and verifies the checksum of every file on the tape. The second CVU component can be invoked from the command line for manual verification of a tape. Each CVU instance can recover from failures without having to re-verify the entire tape volume.

When started, the first CVU component searches the database for the matching silo ID, then finds the tape volume with the earliest closeout time that is in the Requires Verification state. Then it passes the silo ID and tape ID to the second executable CVU component.

The second CVU component accepts the silo ID and tape volume ID as input and verifies the checksum of every file on the tape. It uses AMASS file system commands and the AMASS API to perform its functions.

The CVU currently supports the following three types of checksumming:

- CKSUM.
- ECS.
- MD5.

The procedure for manually invoking the CVU starts with the assumption that the Science Data Server/Storage Management database server is currently running.

Invoke the Checksum Verification Utility (CVU)

---

1 Log in at the machine where the Checksum Verification Utility (CVU) is installed (e.g., e0drg11, g0drg01, l0drg01, or n0drg01).

2 To change to the directory for starting the CVU, type **cd /usr/ecs/<*MODE*>/CUSTOM/utilities** and then press the **Return/Enter** key.

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

3 To identify the mode for the utility at the UNIX prompt type **setenv MODE <*MODE*>** and then press the **Return/Enter** key.

- **<*MODE*>** is the operating mode in which the utility is to be executed (e.g., OPS, TS1, or TS2).

4 To run the CVU, at the UNIX prompt type **EcDsStCheckVolumeFiles <*MODE*> <*Silo ID*> <*Volume ID*>** and then press the **Return/Enter** key.

- The CVU runs and records state information in an error recovery file designated CHECKSUM<*silo*>:<*volume*> file (e.g., CHECKSUM1:52, which has state

information on Silo 1, Volume 52) in the /usr/ecs/OPS/CUSTOM/data/DSS/checksum directory (or as specified by the value assigned to the CHECKSUM_DATA_DIR configuration parameter). State can be PASS, FAIL, WARN, or PENDING.

  - PASS indicates that the checksum has been performed on that file, and it was OK.
  - FAIL indicates that the checksum validation failed.
  - PENDING indicates that the checksum has not yet been performed.
  - WARN indicates that the checksum could not be performed for a non-fatal reason, e.g. either no checksum exists for the file, or the file itself does not exist in the SDSRV database.

- The CVU records new operational-related information in the **EcDsStCheckVolumeFiles.ALOG** file (/usr/ecs/<*MODE*>/CUSTOM/logs directory).

Checksum De-activation

The system design incorporates calculation of a checksum when a granule is inserted into the archive. If such a checksum is calculated, it can then be used as an indicator to determine if there is data corruption within the archive. Comparison of the original checksum with one calculated, for example, when the granule is retrieved (e.g., for processing or distribution) can detect whether the inserted file and the retrieved file are the same. If the checksums do not match, then the operator can investigate (e.g., by using the **Storage Events** tab of the Storage Management GUI). The checksums are set in the configuration for the archive server, with variables that set calculation on granule insert and calculation on retrieval. The Storage Management GUIs provide an easy way to set these and other STMGT configuration parameters. The settings are available from the **Storage Config.** tab, by highlighting the Archive Server and clicking on the **Modify Server** button. This opens the **Archive Server Configuration** window, as illustrated in Figure 28. As the figure shows, the window includes option buttons to **Enable Checksumming On Store:** and **Enable Checksumming On Retrieve:**. Setting checksum calculation variables is addressed in Document 611-EMD-001, *Release 7.11 Mission Operation Procedures for the EMD Project*, Procedure 17.3.5.

Calculation of checksums can be time consuming. System throughput may be significantly improved if checksum calculation on granule insert is turned off, and therefore the default reflects checksum calculation turned off. Unfortunately, turning checksums off compromises the ability to detect data corruption in the archive. This problem may be alleviated somewhat by calculating a checksum when a granule is first retrieved from the archive and storing that checksum to be compared with one calculated upon a later retrieval. However, this approach will not guard against the possibility of data corruption on initial insertion (e.g., through I/O read errors).

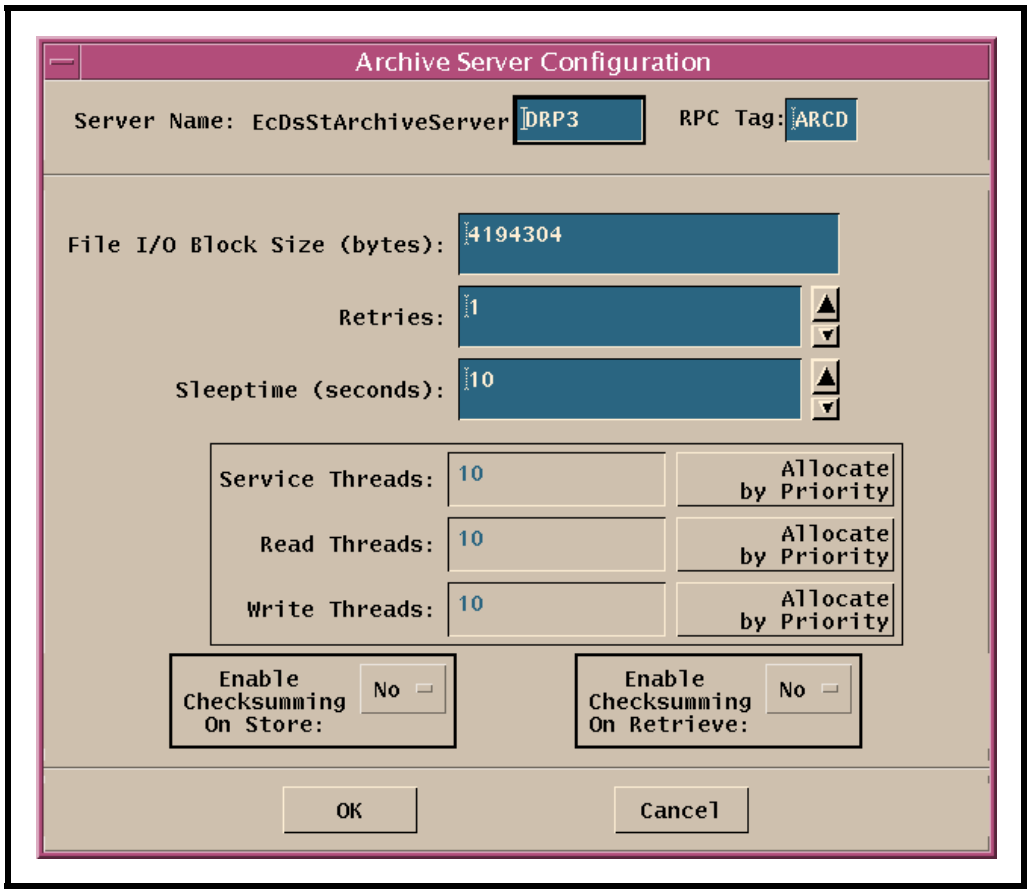


Figure 28. Archive Server Configuration (from Storage Management GUI)

This page intentionally left blank.

Data Pool Management

Features of the Data Pool Maintenance GUI

Most Archive or support personnel tasks for monitoring and maintaining the Data Pool require the use of the **Data Pool Maintenance (DPM) GUI**. The **DPM GUI** permits an operator to perform tasks in the following general areas:

- Monitoring Data Pool Active Insert Processes and Insert Actions.
- Managing Data Pool File Systems.
- Enabling/Disabling Data Compression.
- Managing Compression Algorithms.
- Managing Cloud Cover Information.
- Checking the Status of Batch Inserts.
- Checking the Data Pool Insert Queue.
- Managing Data Pool Configuration Parameters and Data Pool Tuning.
- Managing Data Pool Collection Groups.
- Managing Data Pool Collections within Collection Groups.
- Managing Themes.

Other tasks are supported by scripts or utilities. For example, a Data Pool Update Expiration Script (Update Granule Utility) is available for extending the period of retention for selected science granules already in the Data Pool. There is a Data Pool cleanup utility that is typically run in a cron job, but may be invoked manually. Similarly, a utility for accumulating Data Pool access statistics is usually run in a cron job but may be invoked manually. There is a command line utility that permits operators to execute batch inserts of data from the archive into the Data Pool.

Distribution of data from the Data Pool is supported by the **HDF-EOS to GeoTIFF Conversion Tool** (**HEG**). There are two versions of HEG:

- Data Pool HEG.
- Standalone HEG.

The Standalone HEG is a tool that an end user downloads and runs on his/her own workstation to convert EOS data products on the workstation from one format to another. The Data Pool HEG, which is accessed through the DAAC **Data Pool Web Access GUI** interface, is used to convert EOS data products before they are downloaded or shipped from the DAAC.

With the implementation of Synergy V (Release 7.10) many of the Data Pool HEG functions that previously had been controlled through the Data Pool (e.g., tracking of HEG orders) were transferred to the **Order Manager GUI** (**OM GUI**). However, each Data Pool collection's eligibility for HEG processing is still enabled or disabled using the **DPM GUI**.

Finally, the **Spatial Subscription Server GUI** is a major Data Pool management tool. Although used primarily by User Services or science personnel, Archive or engineering support personnel may use it to extend the period of retention in a Data Pool insert subscription, and to view statistics on the processing of events and actions by the Spatial Subscription Server.

Both the **DPM GUI** and the **Spatial Subscription Server GUI** are web applications, accessed through the Netscape web browser. For Synergy V, the **DPM GUI** and the **Spatial Subscription Server GUI** are certified for use with **Netscape 7** (and higher) browsers (Figure 29) and it is strongly recommended that Netscape 7-compatible browsers only be used when accessing the GUIs. The GUIs do **not** work properly with **Netscape 4.78** or earlier versions.

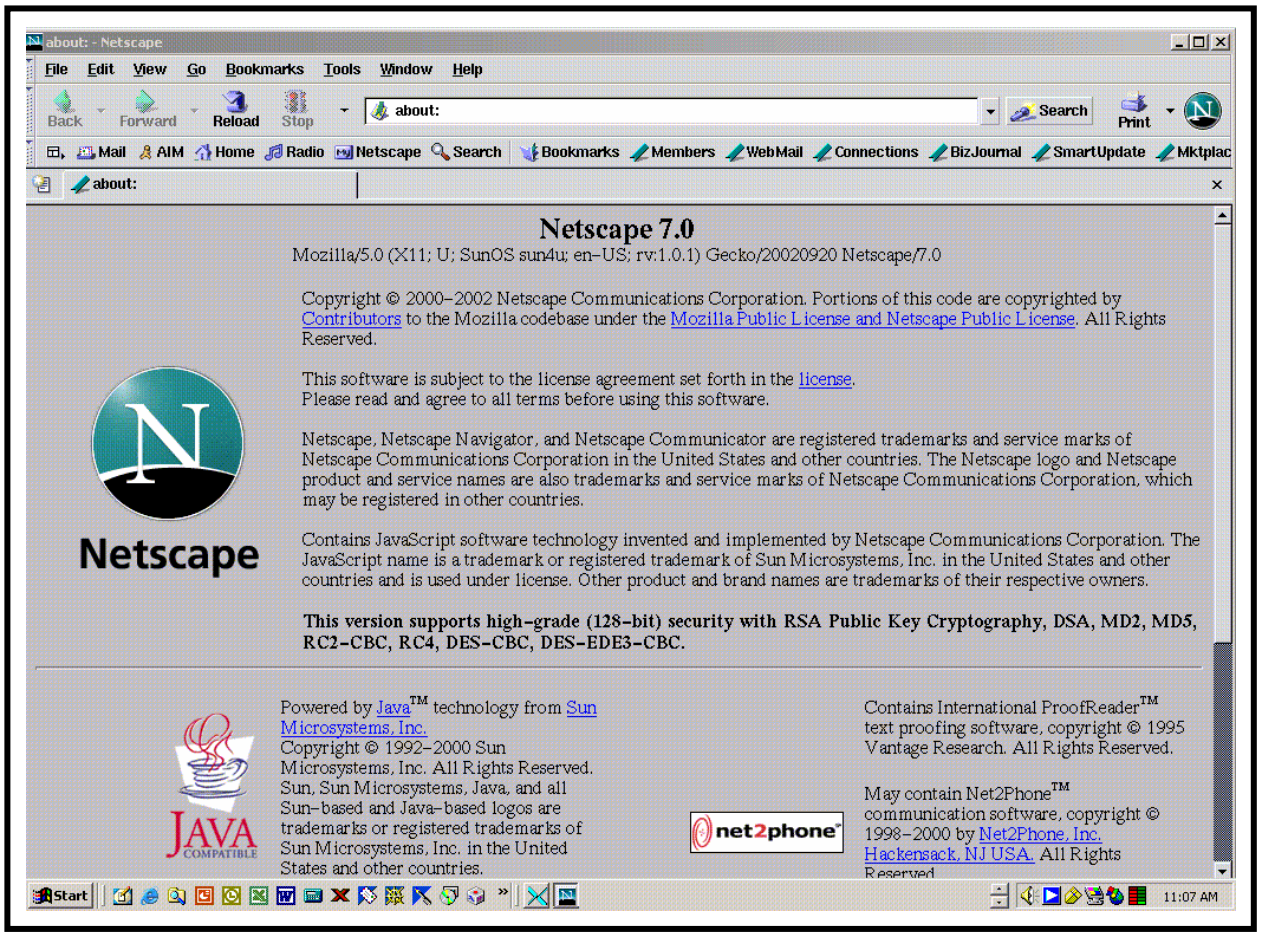


Figure 29. Netscape 7.0 Web Browser

New operator GUI security standards require the following two levels of permissions for the **DPM GUI** and the **Spatial Subscription Server GUI**:

- Full Capability.
- Limited Capability.

An operator's level of permission is determined when the operator logs in to the GUI using the security login prompt (Figure 30).

Full-capability operators have the ability to configure parameters and perform all other actions that can be accomplished with the GUIs. 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.

This lesson provides instruction in the full-capability version of the GUIs. However, the functions that are available to limited-capability operators as well as the functions that are not available to limited-capability operators are identified.

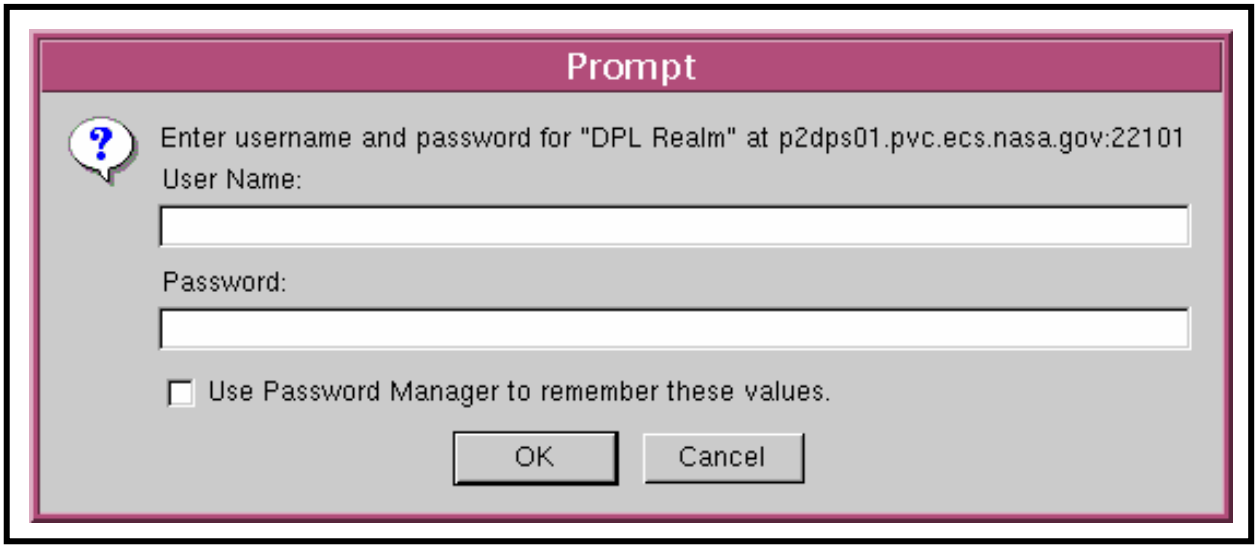


Figure 30. Security Login Prompt

The Synergy V **DPM GUI** provides both full-capability and limited-capability operators with the ability to perform the following functions:

- Launch the DPM GUI.
- Shut Down the DPM GUI (End a DPM GUI Session).
- Monitor Data Pool Active Insert Processes.
- View a List of Data Pool File Systems.

- View a List of Compression Algorithms.
- View Cloud Cover Information.
- Check the Status of Batch Inserts.
- Check the Data Pool Insert Queue.
- View DPM Configuration Parameter Values.
- View Collection Group and Collection Information.
- View a List of Themes.
- Filter a List of Themes.

In addition to the preceding actions, full-capability operators can perform the following actions:

- Modify a Data Pool File System.
- Add a Data Pool File System.
- Enable/Disable Data Compression.
- Modify Compression Algorithms.
- Add a Compression Algorithm.
- Deactivate a Compression Algorithm.
- Add New Cloud Cover Information.
- Modify Cloud Cover Source Descriptions.
- Delete Cloud Cover Information.
- Cancel a Data Pool Insert Action.
- Modify DPM Configuration Parameter Values.
- Modify Collection Groups.
- Add a Collection Group.
- Add an ECS Collection to a Collection Group.
- Add a NON-ECS Collection to a Collection Group.
- Modify an ECS Collection.
- Modify a NON-ECS Collection.
- Modify a Theme.
- Add a Theme.
- Delete a Theme.

Figure 31 illustrates the **DPM GUI Home Page**, from which the operator can perform some monitoring and maintenance tasks and from which there is access to other pages supporting other tasks.

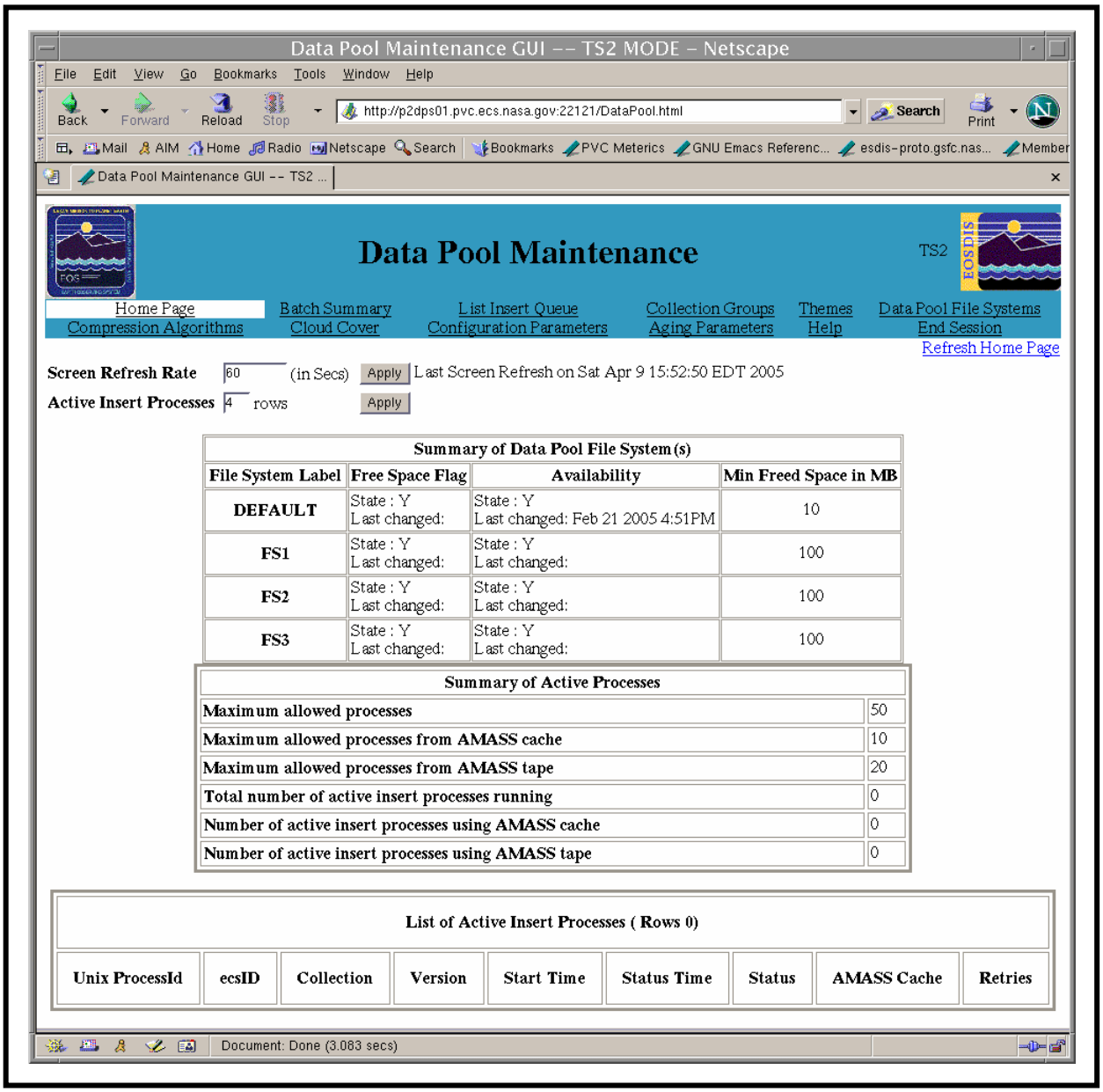


Figure 31. DPM GUI Home Page

Home Page

The **DPM GUI Home Page** (Figure 31) displays the state of several parameters and allows an operator to make changes. It also lists active insert processes. Near the top of the **Home Page** are links allowing an operator to access other functions including the following items:

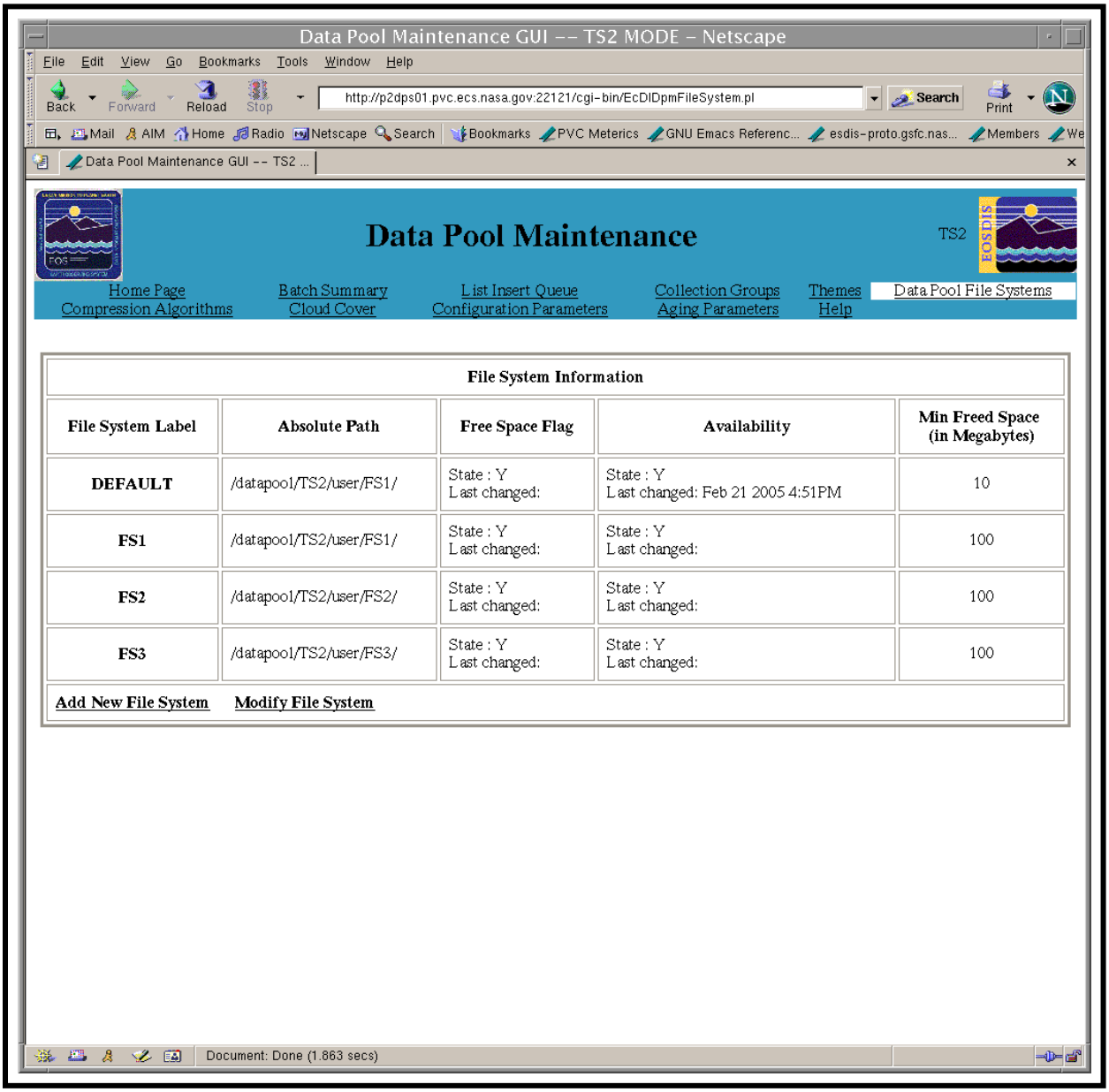
- **Data Pool File Systems.**
- **Compression Algorithms.**
- **Cloud Cover.**
- **List Insert Queue.**
- **Batch Summary.**
- **Collection Groups.**
- **Themes.**
- **Configuration Parameters.**
- **Aging Parameters.**

There is also a **Help** page for assistance in navigation of the GUI and an **End Session** link for logging out of the GUI.

Data Pool File Systems

Figure 32 illustrates the **Data Pool File System Information** page that allows both full-capability and limited-capability operators to view a list of Data Pool file systems and obtain information on Free Space Flag, Availability for insert, and Minimum Freed Space. From this page, the full-capability operator is able to configure a new file system or modify an existing file system (which may include assigning Availability and/or No Free Space status).

Clicking on the **Add New File System** link takes the full-capability operator to the **Add New File System** page shown in Figure 33. The operator needs to add data in five fields --- 1) [File System] Label: a label representing an existing Data Pool file system; 2) Absolute Path: the path to the directory where the file system is located (the basic ftp root path is provided and the operator completes the path name if necessary); 3) Free Space Flag: a value that needs to be set to either “ON” or “OFF” (ON means free space is available for inserts; OFF means free space is not available); 4) Availability: a value that needs to be set to either “YES” or “NO” (YES means the file system is currently available for Data Pool insert; NO means the file system is not available for Data Pool insert); 5) Min Freed Space (in Megabytes): an integer value that represents the minimum amount of freed space in the file system in megabytes; it is an amount of space must remain free in order to make the file system available for insert.



File System Information				
File System Label	**Absolute Path**	**Free Space Flag**	**Availability**	**Min Freed Space (in Megabytes)**
DEFAULT	/datapool/TS2/user/FS1/	State : Y Last changed:	State : Y Last changed: Feb 21 2005 4:51PM	10
FS1	/datapool/TS2/user/FS1/	State : Y Last changed:	State : Y Last changed:	100
FS2	/datapool/TS2/user/FS2/	State : Y Last changed:	State : Y Last changed:	100
FS3	/datapool/TS2/user/FS3/	State : Y Last changed:	State : Y Last changed:	100
Add New File System **Modify File System**				

Figure 32. Data Pool File System Information Page

Selecting the **Modify File System** link takes the full-capability operator to the **Modify File System Information** page shown in Figure 34. The operator can change the Absolute Path, Free Space Flag, Availability flag, or the Min Freed Space on this page. There are check boxes associated with each file system. The operator can change multiple file systems at one time by checking the desired file systems' checkboxes and clicking on the **Apply Change** button.

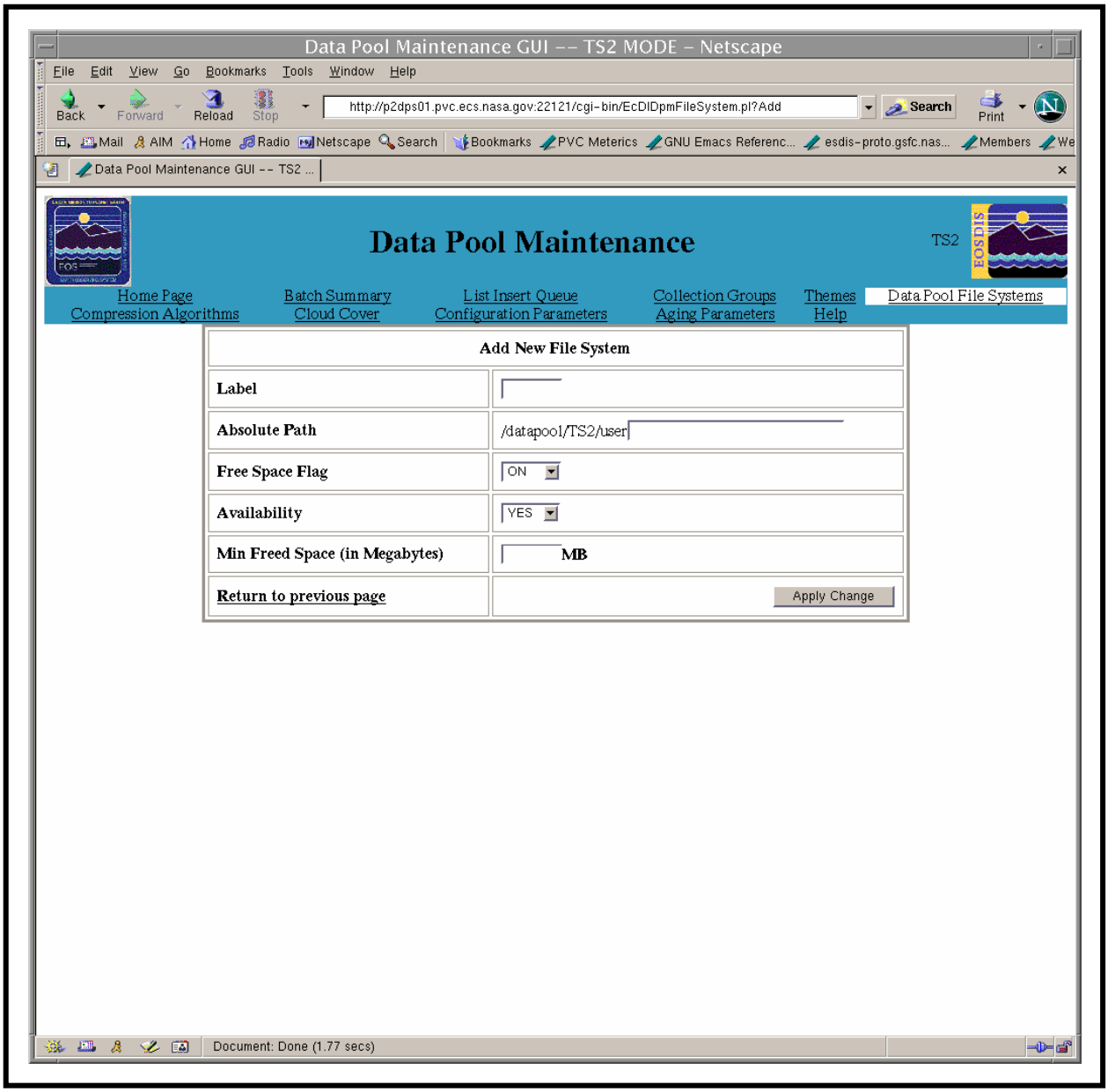


Figure 33. Add New File System Page

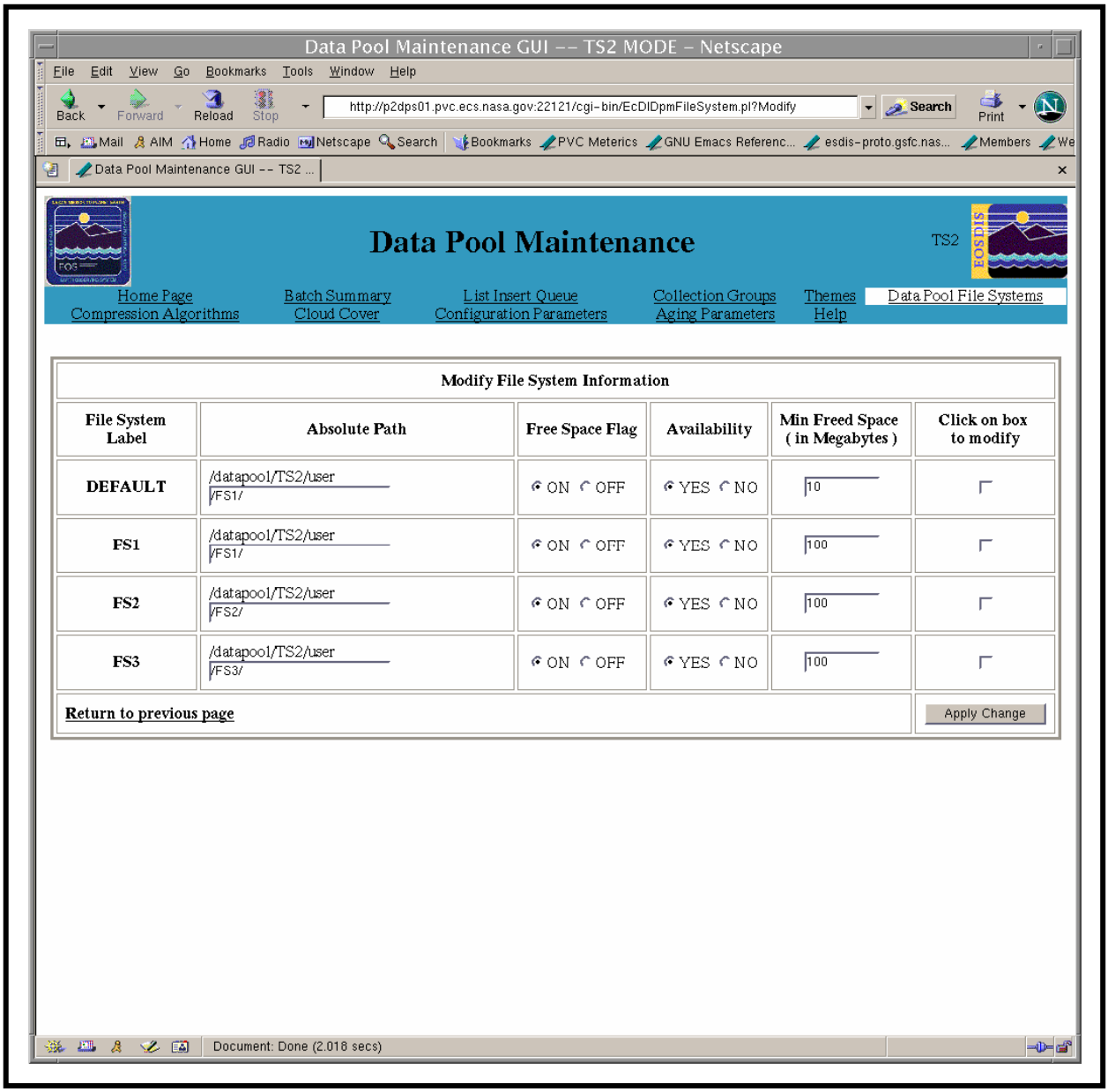
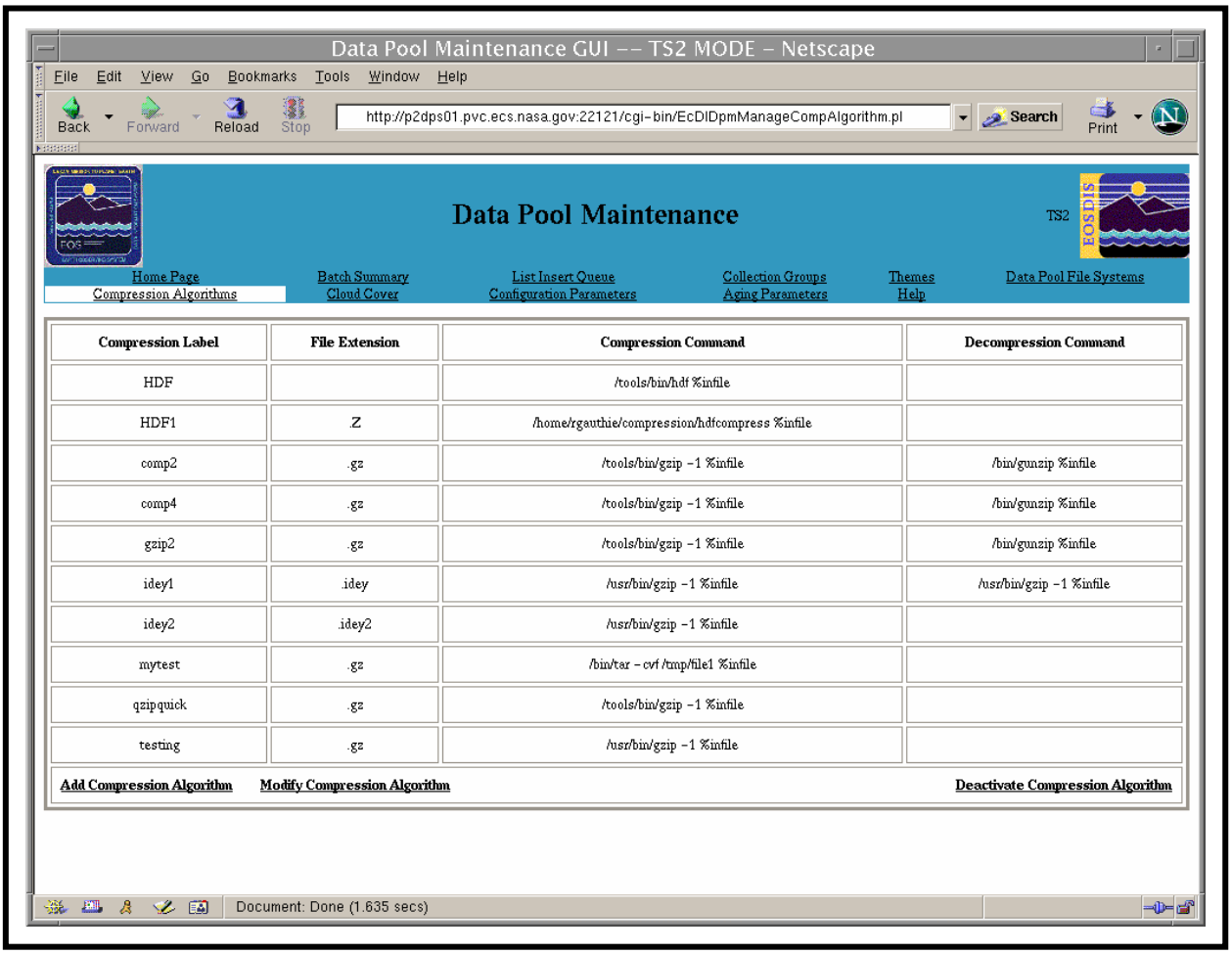


Figure 34. Modify File System Information Page

Compression Algorithms

Both full-capability and limited-capability operators can view existing compression algorithms by clicking on the **Compression Algorithms** link shown in Figure 31. The link takes the operator to the **Compression Algorithms** page shown in Figure 35. The page displays the Compression Label, File Extension, Compression Command, and Decompression Command for the compression algorithm.



Compression Label	File Extension	Compression Command	Decompression Command
HDF		/tools/bin/hdf %infile	
HDF1	.Z	/home/rgauthie/compression/hdfcompress %infile	
comp2	.gz	/tools/bin/gzip -1 %infile	/bin/gunzip %infile
comp4	.gz	/tools/bin/gzip -1 %infile	/bin/gunzip %infile
gzip2	.gz	/tools/bin/gzip -1 %infile	/bin/gunzip %infile
idey1	.idey	/usr/bin/gzip -1 %infile	/usr/bin/gzip -1 %infile
idey2	.idey2	/usr/bin/gzip -1 %infile	
mytest	.gz	/bin/tar -cvf /tmp/file1 %infile	
qzipquick	.gz	/tools/bin/gzip -1 %infile	
testing	.gz	/usr/bin/gzip -1 %infile	

Add Compression Algorithm Modify Compression Algorithm Deactivate Compression Algorithm

Figure 35. Compression Algorithms Page

The full-capability operator can modify an existing compression algorithm by clicking on the **Modify Compression Algorithm** link shown in Figure 35. The link takes the operator to the **Modify Compression Algorithm** page shown in Figure 36. The operator can modify the File Extension, Compression Command, and/or Decompression Command for any or all compression algorithms. After making desired changes, the operator clicks on the checkbox(es) adjacent to the compression algorithm(s) to be modified and (after all desired changes have been made) clicks on the **Apply Change** button. The changes are applied to the Data Pool database and the **Compression Algorithms** page is refreshed.

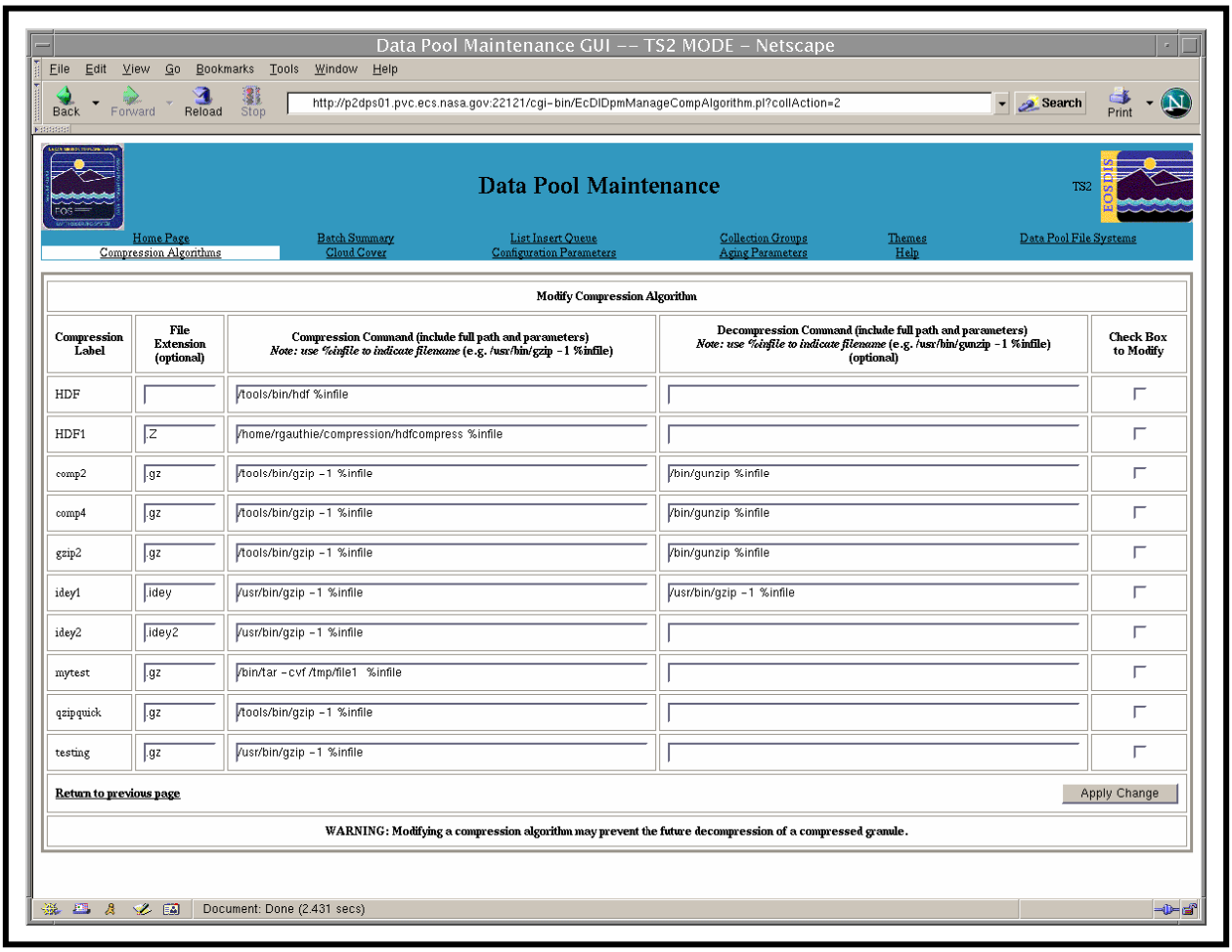


Figure 36. Modify Compression Algorithms Page

The full-capability operator can add a new compression algorithm by clicking on the **Add Compression Algorithm** link shown in Figure 35. The link takes the operator to the **Add Compression Algorithm** page shown in Figure 37. After entering Compression Label, File Extension, Compression Command, and Decompression Command data for the new compression algorithm, the operator clicks on the **Apply Change** button. The new compression algorithm is added to the Data Pool database and the **Compression Algorithms** page is refreshed.

The full-capability operator can deactivate a compression algorithm by clicking on the **Deactivate Compression Algorithm** link shown in Figure 35. The link takes the operator to the **Deactivate Compression Algorithm** page shown in Figure 38. To deactivate compression algorithms the operator clicks on the checkbox(es) adjacent to the compression algorithm(s) to be deactivated then clicks on the **Deactivate Selected** button. The selected algorithm(s) is (are) deactivated in the Data Pool database and the **Compression Algorithms** page is refreshed.

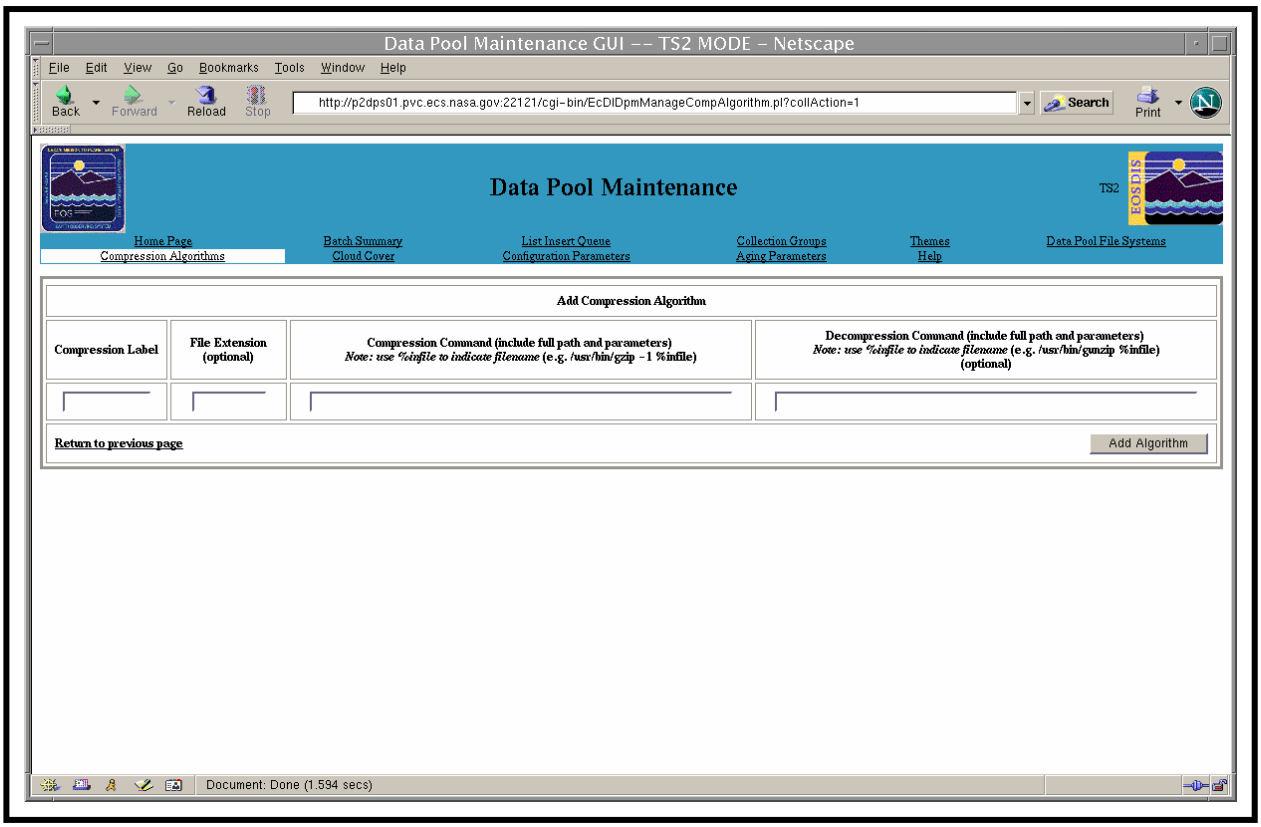


Figure 37. Add New Compression Algorithm Page

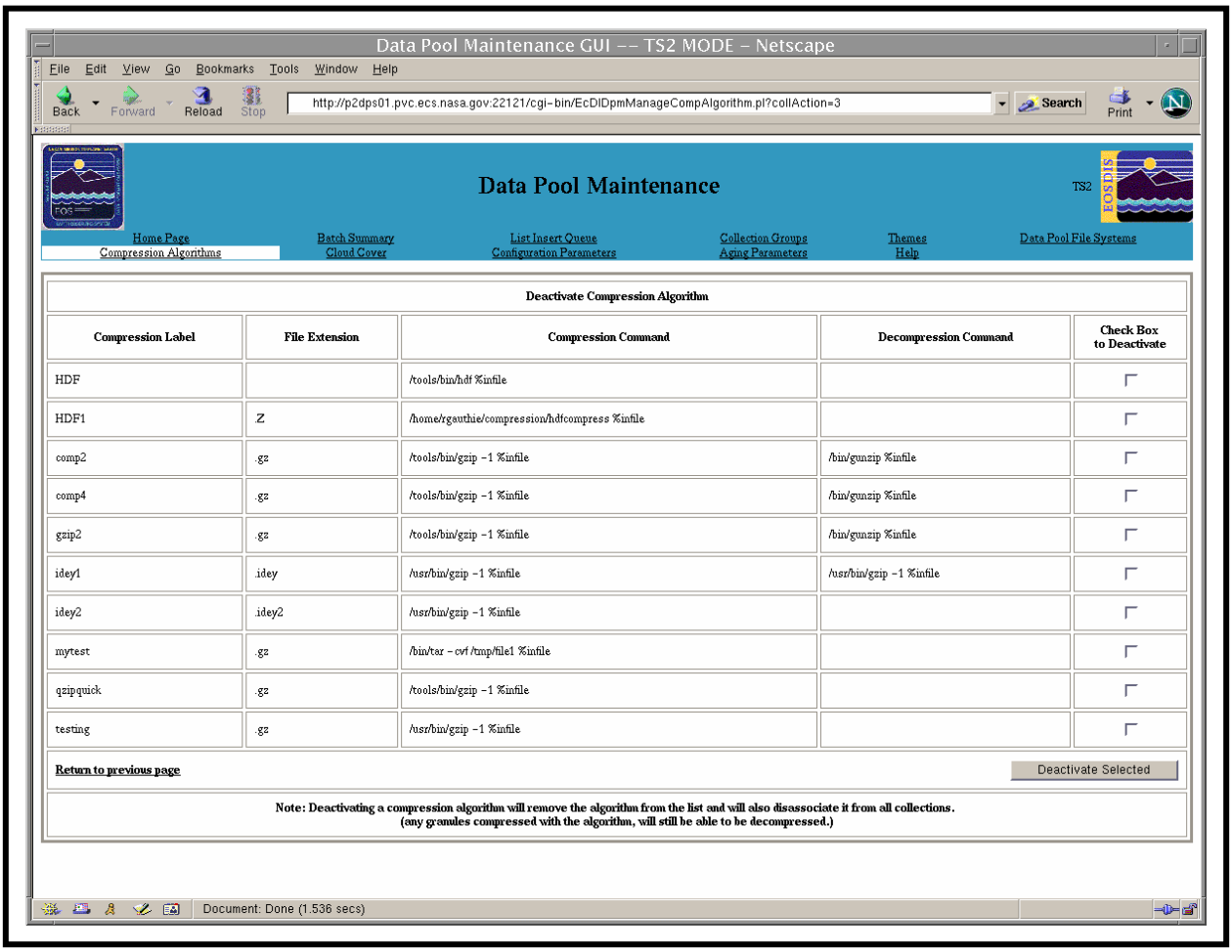


Figure 38. Deactivate Compression Algorithm Page

Cloud Cover

Both full-capability and limited-capability operators can view existing cloud cover information in the Data Pool database by clicking on the **Cloud Cover** link shown in Figure 31. The link takes the operator to the **Cloud Cover Information** page shown in Figure 39. The page displays the information concerning the sources of cloud cover; i.e., the Source Type, Source Name, and Source Description.

The full-capability operator can delete source types by clicking on the checkbox(es) adjacent to the source(s) to be deleted then clicking on the **Apply Change** button. The selected source(s) is (are) deleted from the Data Pool database and the **Cloud Cover Information** page is refreshed. If any cloud cover information is associated with any collection, it will not be deleted.

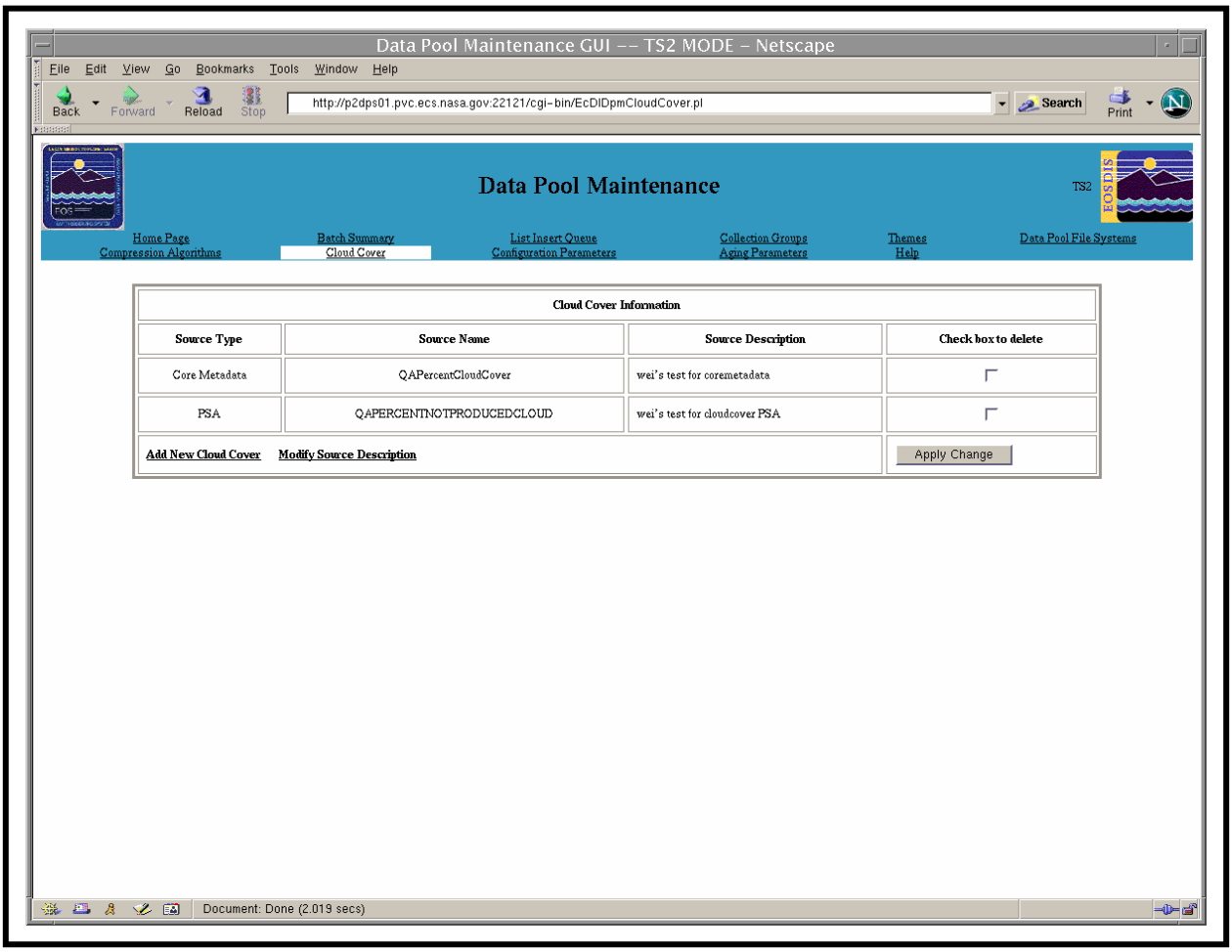


Figure 39. Cloud Cover Information Page

The full-capability operator can add a new cloud cover source by clicking on the **Add New Cloud Cover** link shown in Figure 39. The link takes the operator to the **Add New Cloud Cover Information** page shown in Figure 40. After selecting the Source Type from an option list and entering the Source Name and Source Description, the operator clicks on the **Apply Change** button. All Source Names are validated against the Science Data Server database. The new cloud cover source is added to the Data Pool database and the **Cloud Cover Information** page is refreshed.

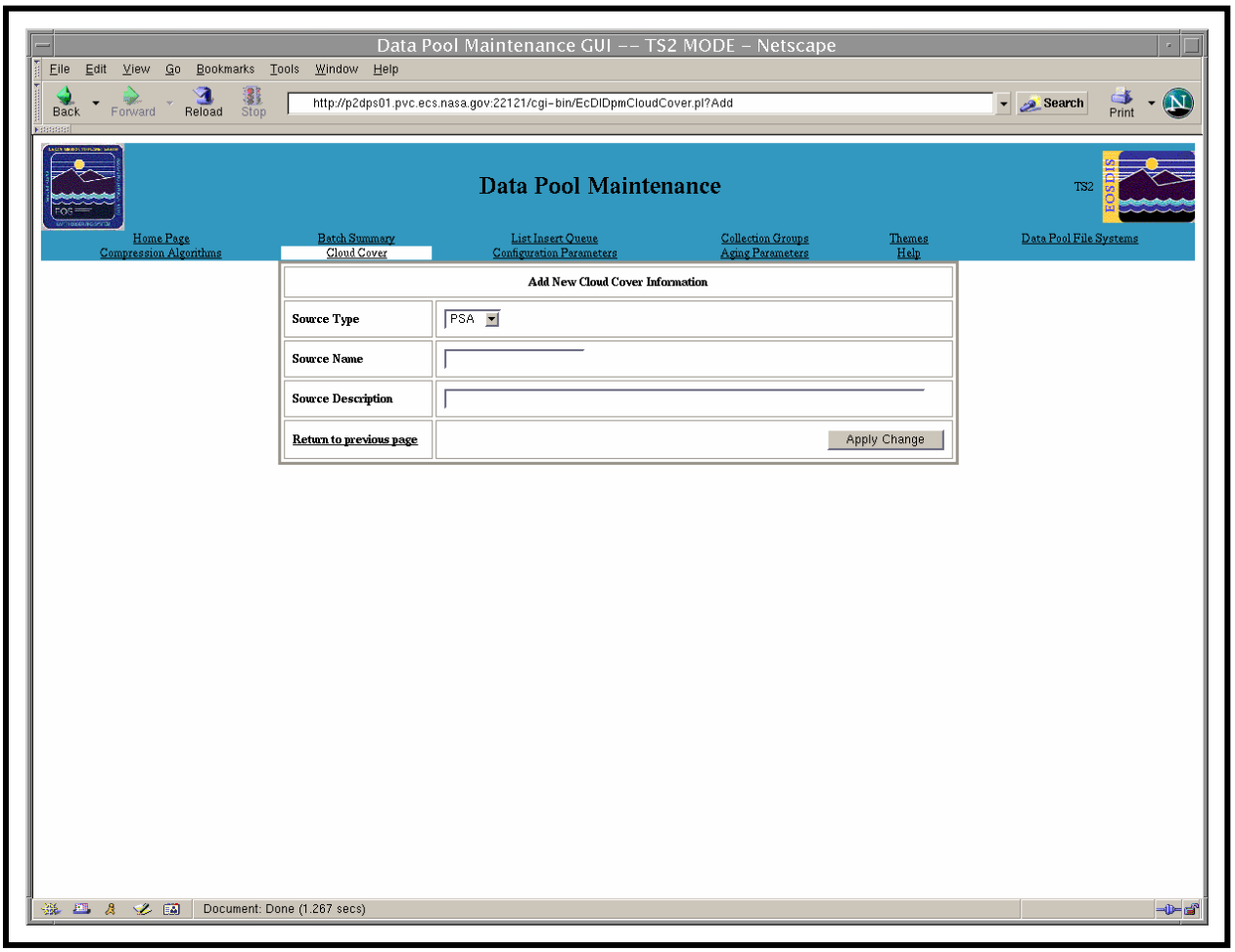


Figure 40. Add New Cloud Cover Information Page

The full-capability operator can modify an existing cloud cover Source Description by clicking on the **Modify Source Description** link shown in Figure 39. The link takes the operator to the **Modify Source Description** page shown in Figure 41. The operator can modify the Source Descriptions only. (To modify a Source Type or Source Name the operator must delete the applicable cloud cover information row and add a new one with the correct information.) After making desired changes, the operator clicks on the checkbox(es) adjacent to the source(s) to be modified and clicks on the **Apply Change** button. The changes are applied to the Data Pool database and the **Cloud Cover Information** page is refreshed.

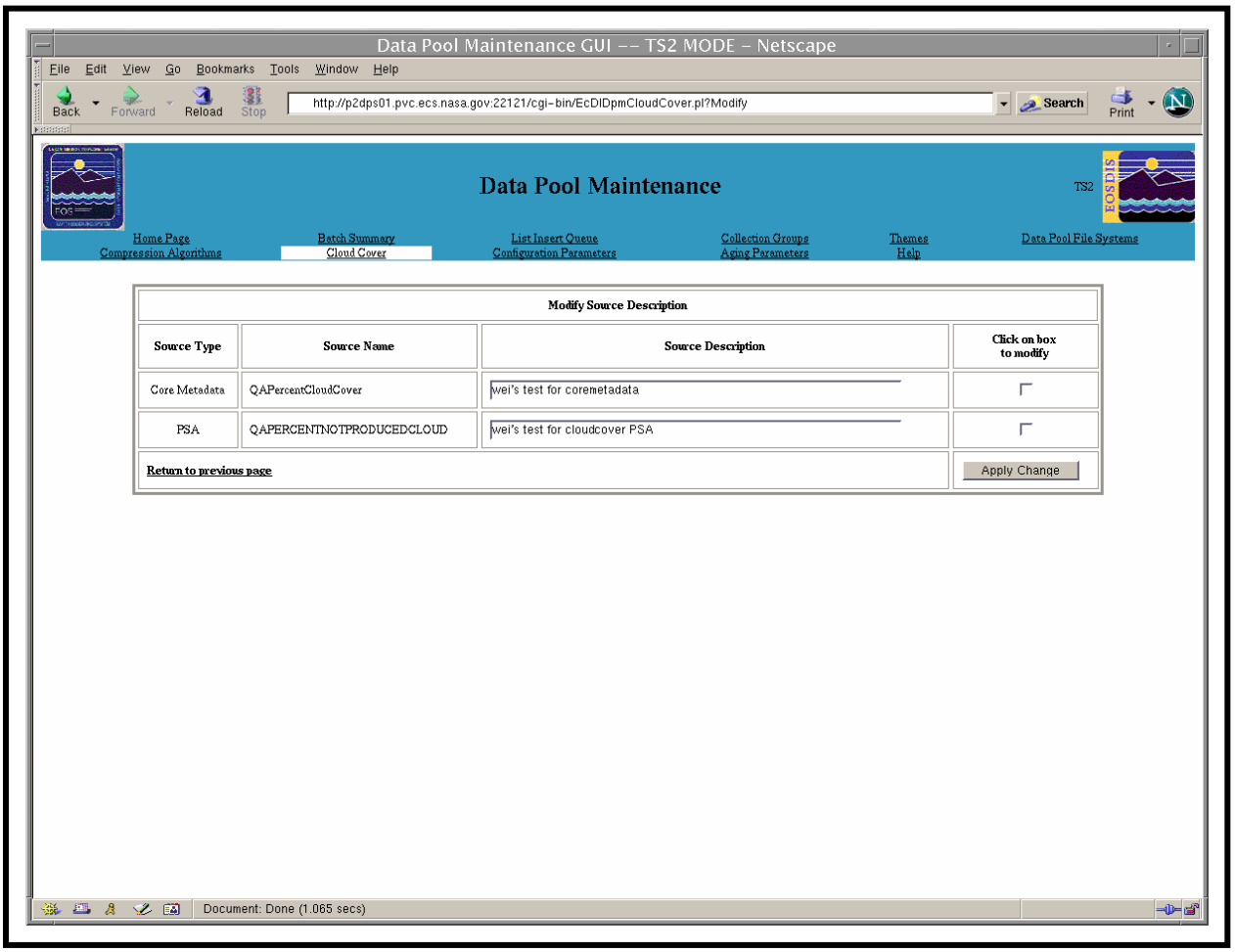


Figure 41. Modify Source Description Page

Batch Summary

Figure 42 illustrates the **Batch Summary** page, which is accessible from the **Batch Summary** link on the **DPM GUI Home Page** (Figure 31). The **Batch Summary** page displays information on inserts made with the command line utility that permits operators to execute batch inserts of data from the archive into the Data Pool. In addition, it displays a summary of the status of Data Pool inserts for each batch label. Insert statuses include "new," "completed," "failed," "retry," and "canceled." The information is accessible to both full-capability and limited-capability operators.

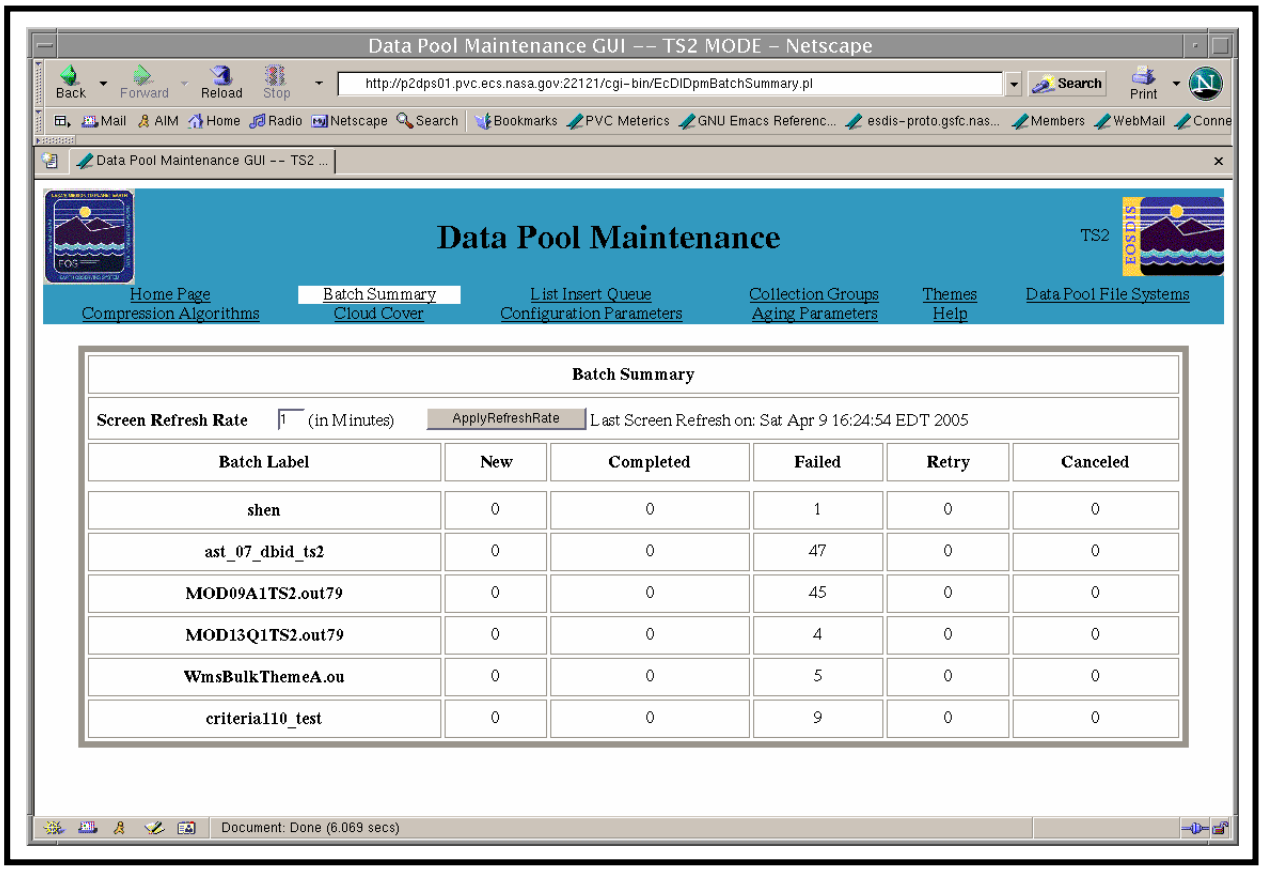


Figure 42. Batch Summary Page

List Insert Queue

Figure 43 illustrates the **List Insert Queue** page, which is accessible from the **List Insert Queue** link on the **DPM GUI Home Page** (Figure 31). The page provides a list with detailed information on inserts left to process. The information is accessible to both full-capability and limited-capability operators.

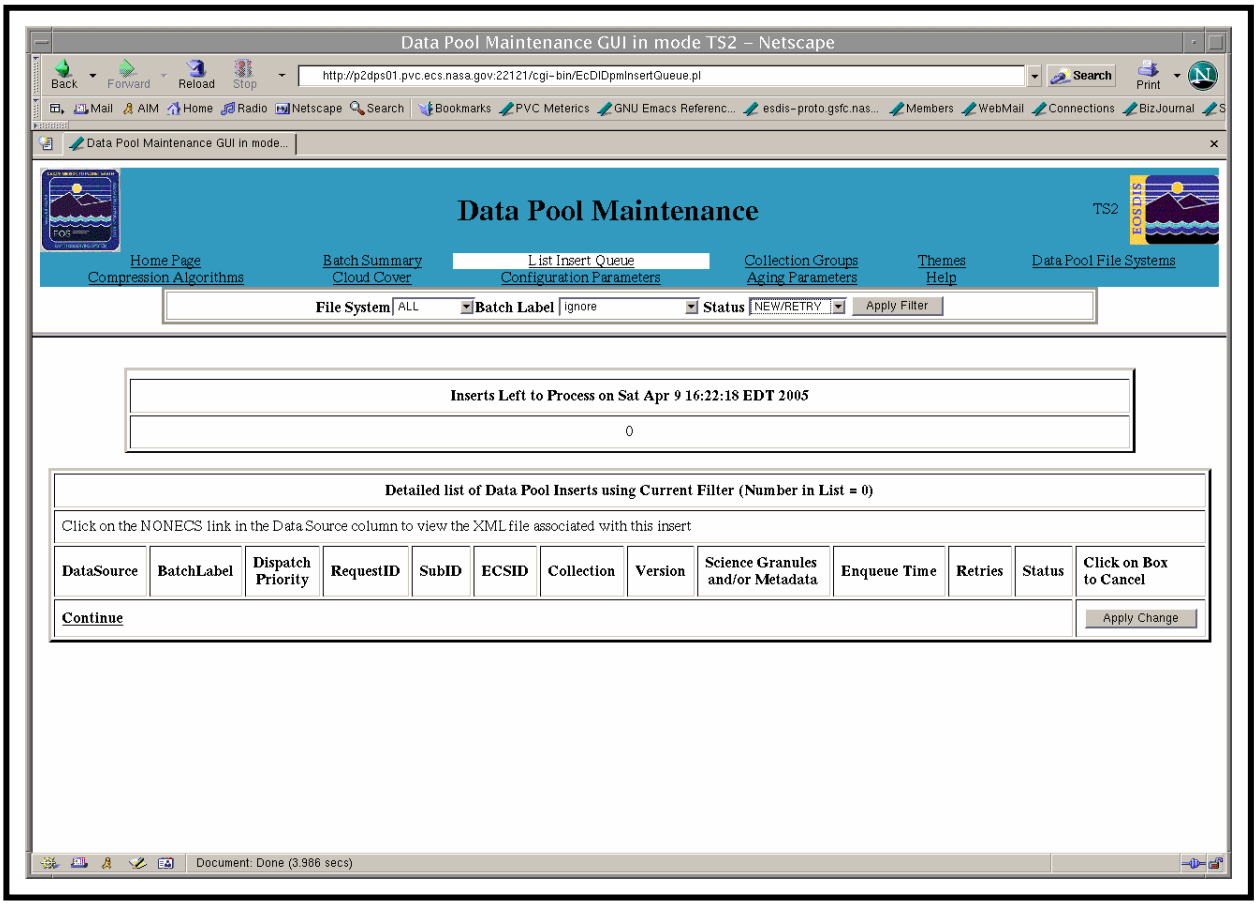


Figure 43. List Insert Queue Page

Insert actions shown on the **Home Page** list of active insert processes appear on the **List Insert Queue** page also. The operator can filter the list by choosing a specific file system from the **File System** pull-down list or a specific batch label from the **Batch Label** pull-down list. The Insert Queue list can also be filtered by Status. For example the operator can choose "Completed" from the **Status** pull-down list, "ALL" from the **File System** pull-down list, and "ALL" from the **Batch Label** pull-down list, which would show all the completed inserts for each batch label for all file systems. After selecting the filter options, the operator clicks on the **Apply Filter** button to display the filtered list. For each listed insert, a full-capability operator can use a check box in the last column to mark the insert for cancellation. The action to cancel the insert is implemented by a click on an **Apply Change** button at the bottom of the page.

The batch insert utility can be used to insert non-ECS data into the Data Pool. On the **DPM GUI List Insert Queue** page, non-ECS data insert actions are identified by the entry “NONECS” in the **Data Source** column. XML file and path name for a non-ECS granule insert action can be viewed by clicking on "NONECS" in the **Data Source** column. Figure 44 shows the appearance of the pages displaying absolute .XML file path and file content.

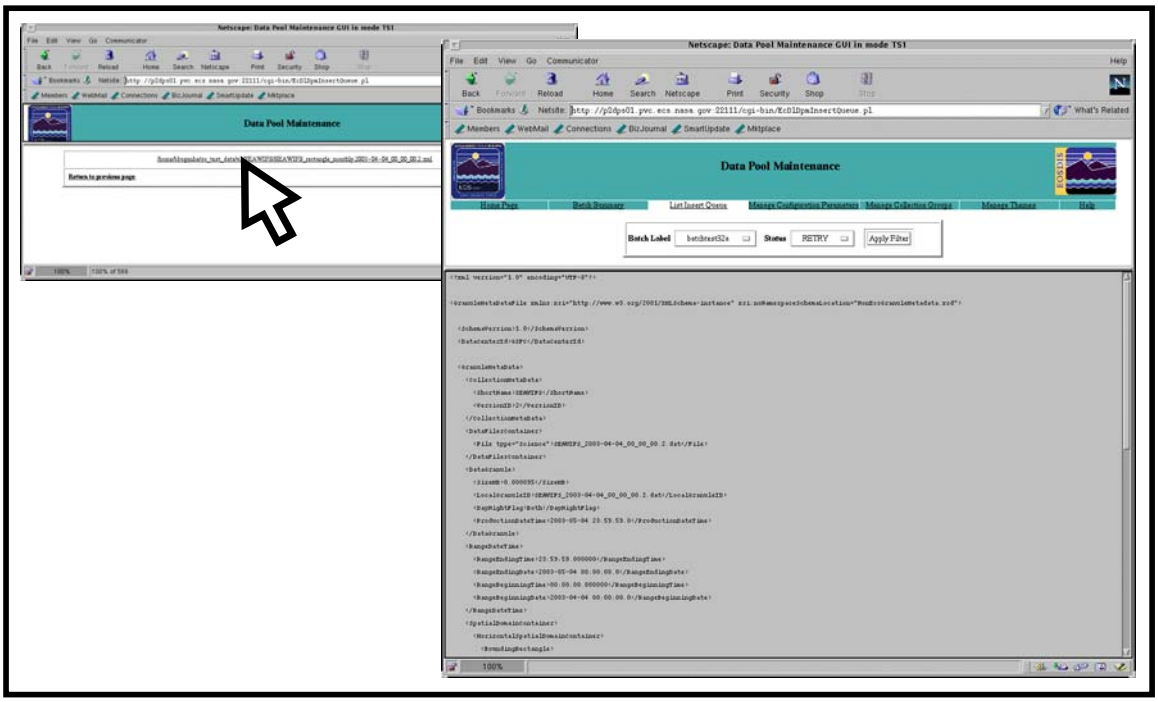


Figure 44. Sample DPM GUI Pages for .XML File Path and Content

Configuration Parameters

Figure 45 shows the **List of Configuration Parameters** page, which is accessible from the **Configuration Parameters** link on the **DPM GUI Home Page** (Figure 31). The page lists numerous Data Pool configuration parameters with their settings and a brief description of each. The information is accessible to both full-capability and limited-capability operators.

For each parameter there is a text box or option list in the **Parameter Value** column so the full-capability operator can assign a new value to the parameter when necessary. In addition, there is a check box that the full-capability operator uses to mark parameters with values to be modified. At the bottom of the page is an **Apply Change** button for implementing the change(s).

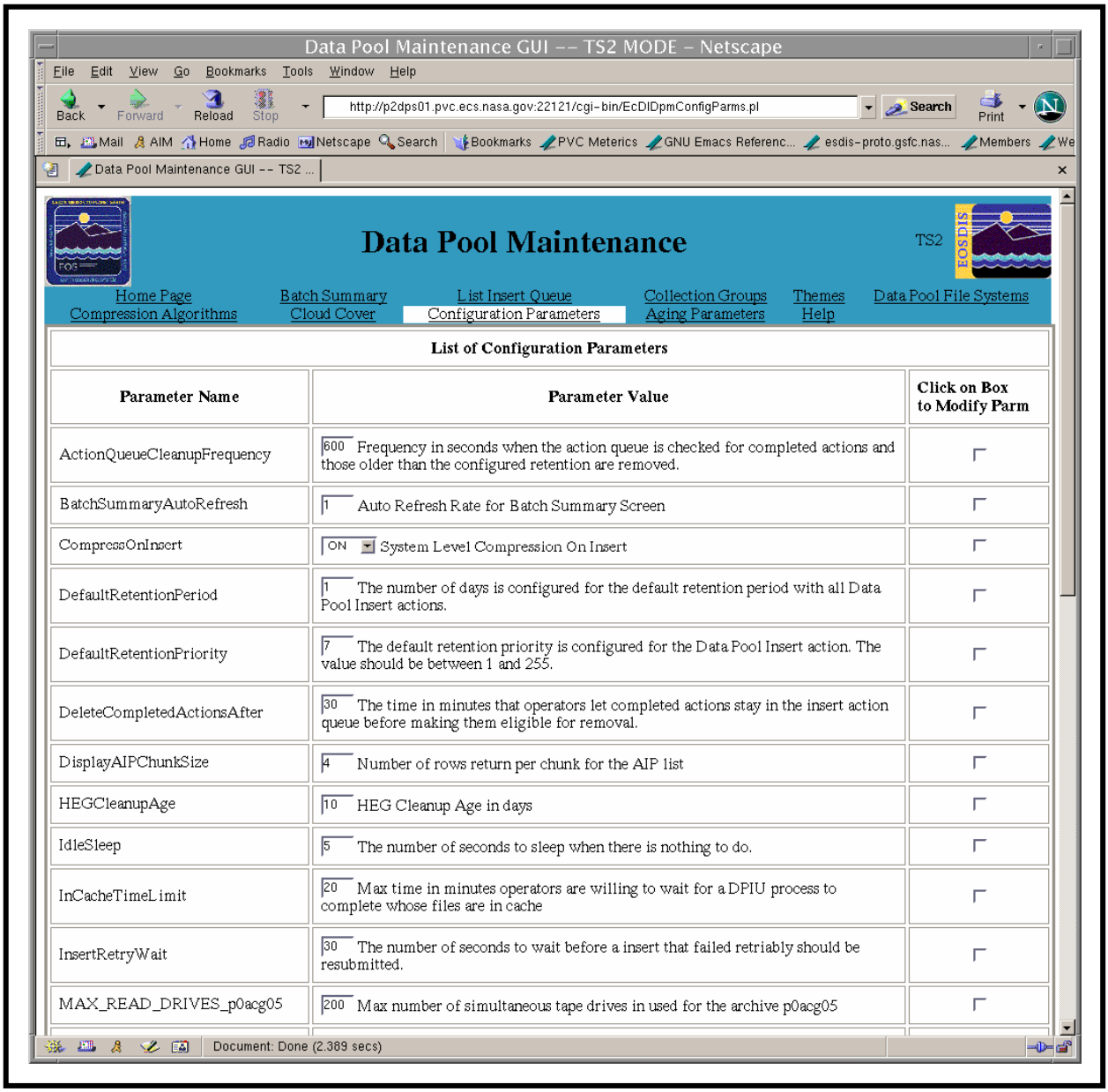


Figure 45. List of Configuration Parameters Page

Aging Parameters

Figure 46 shows the **Aging Parameters** page, which is accessible from the **Aging Parameters** link on the **DPM GUI Home Page** (Figure 31). The page lists the starting priority, aging step, and maximum priority associated with each ECS priority. The information is accessible to both full-capability and limited-capability operators.

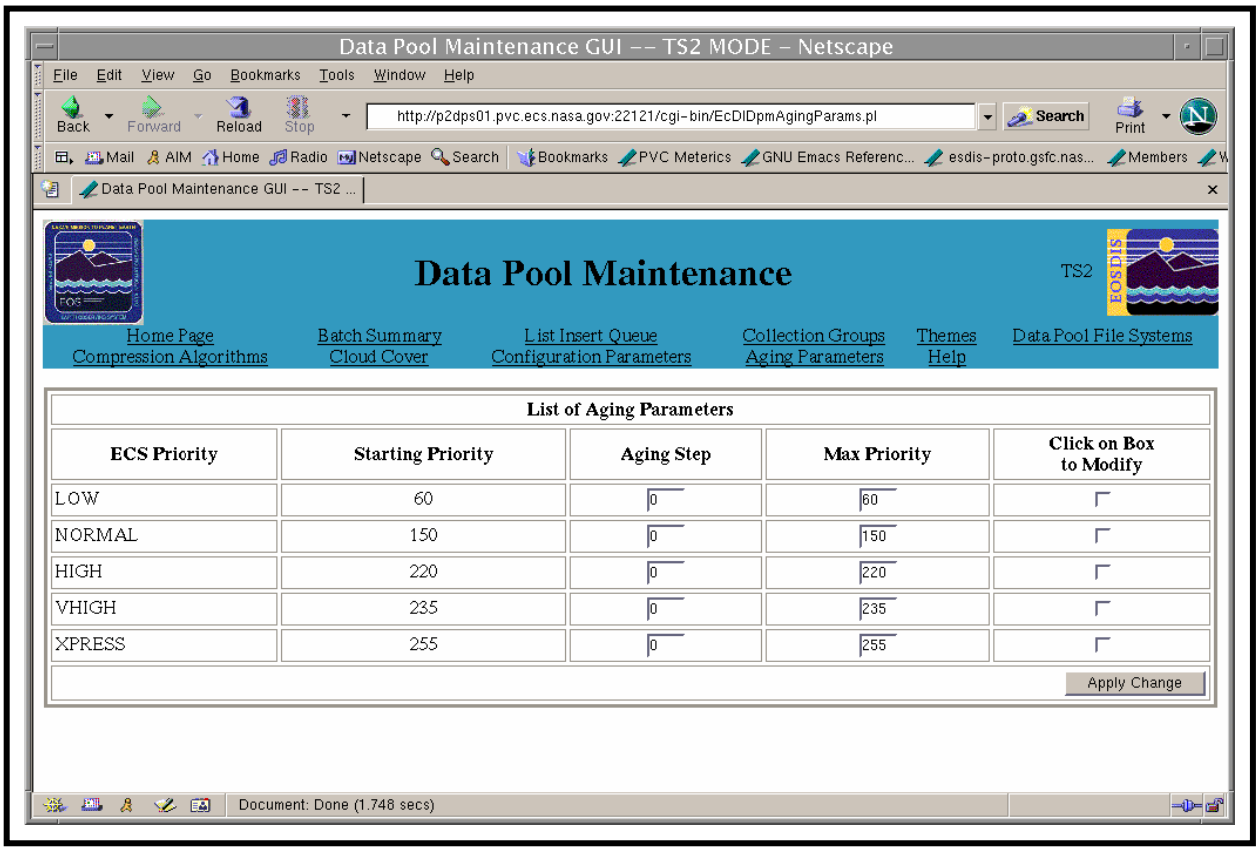


Figure 46. Aging Parameters Page

For the aging step, and maximum priority associated with each ECS priority there is a text box so the full-capability operator can assign a new value to the parameter when necessary. In addition, there is a check box that the full-capability operator uses to mark parameters with values to be modified. At the bottom of the page is an **Apply Change** button for implementing the change(s).

Collection Groups

Figure 47 illustrates the **Collection Groups** page, which is accessible from the **Collection Groups** link on the **Home Page** (Figure 31). The page lists the collection groups, providing for each the Data Source (ECS or NON-ECS), Group ID, Display Name, and a brief description of the collection group. The information is accessible to both full-capability and limited-capability operators.

At the bottom of the **Collection Groups** page, there are links that permit a full-capability operator to **Add Collection Group** or **Modify Collection Group**. A click on one of the **Group ID** links brings up a **Collection Group Detail** page (Figure 48) listing the collection(s) in that group along with a link to add a collection to the group.

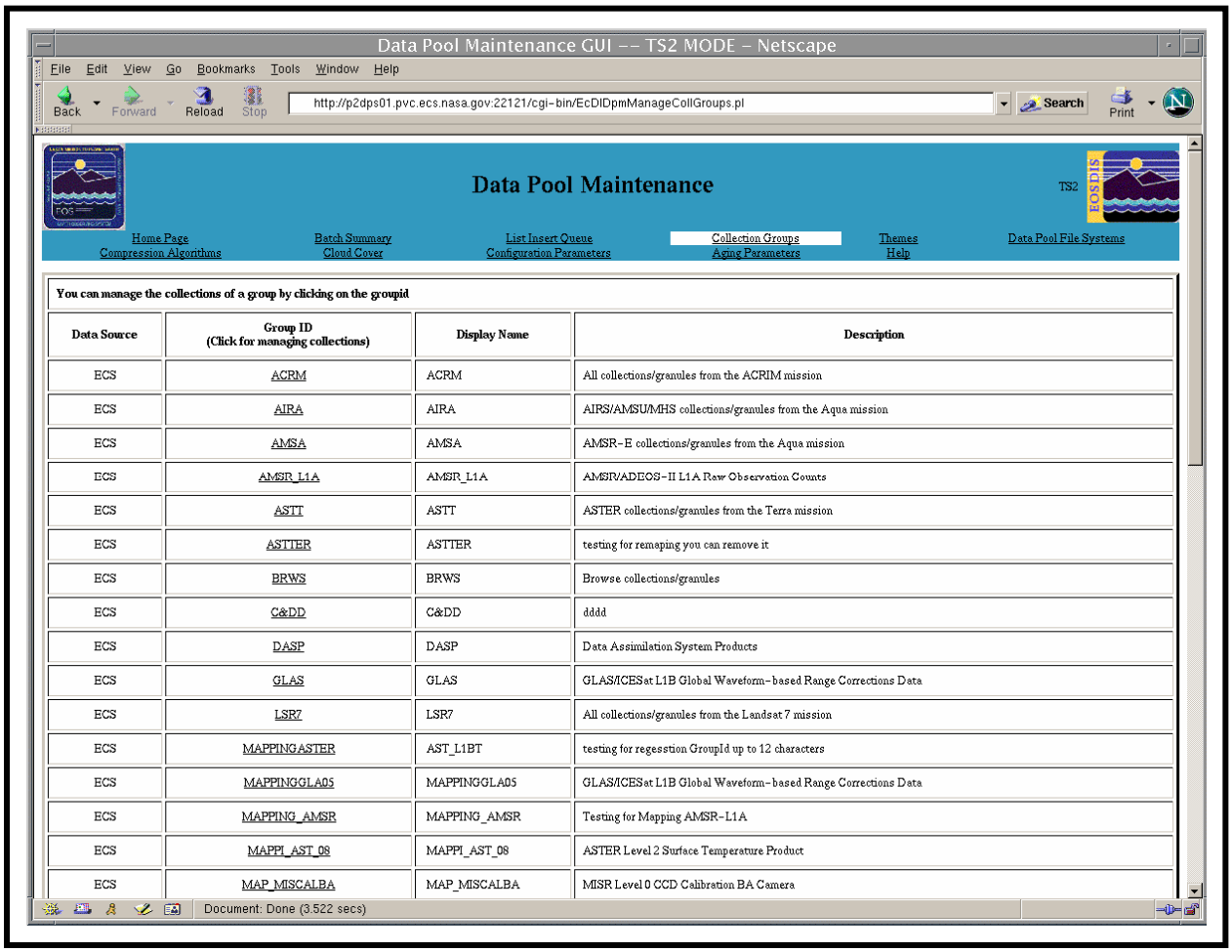


Figure 47. Collection Groups Page

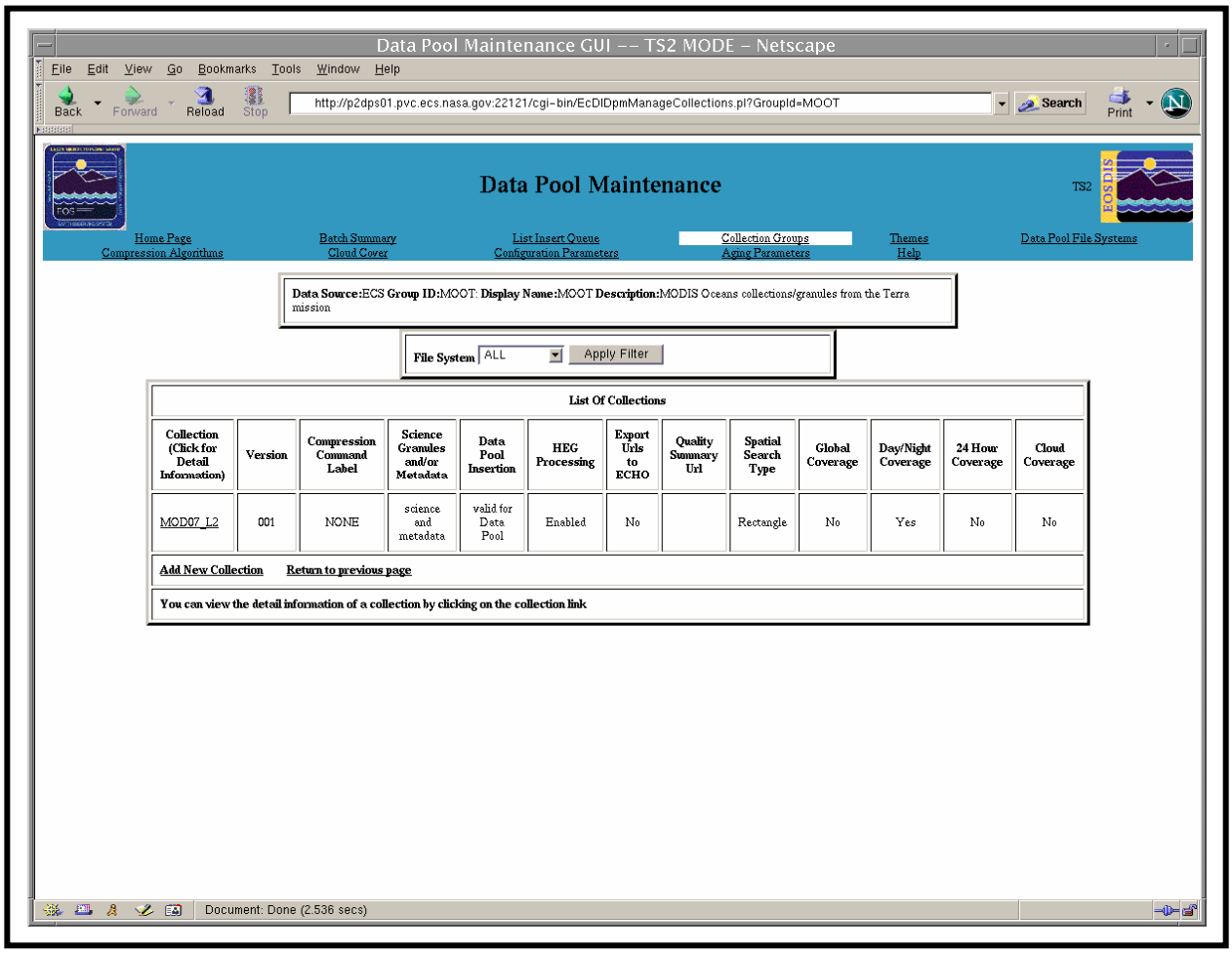


Figure 48. Collection Group Detail Page

Figure 48 shows a **Collection Group Detail (List of Collections)** page obtained by clicking on one of the **Group ID** links on the **Collection Groups** page. The **Collection Group Detail (List of Collections)** page lists the collections in the collection group, providing for each collection information (as applicable) concerning the Version, Compression Command Label, Science Granules and/or Metadata, Data Pool Insertion, HDF-EOS to GeoTIFF Conversion Tool (HEG) Processing, Export URLs to ECHO, Quality Summary URL, Spatial Search Type, Global Coverage, Day/Night Coverage, 24-Hour Coverage, and Cloud Coverage characteristics of the collection. The information is accessible to both full-capability and limited-capability operators.

At the bottom of the **Collection Group Detail (List of Collections)** page, there is a link that permits a full-capability operator to **Add New Collection** to the collection group. A click on one of the **Collection** ID links brings up a **Collection Detail** page (Figure 49) listing the same information for the collection as was displayed on the **Collection Group Detail** page plus some additional information. The additional information includes a Description, File System, Cloud Cover Type, Cloud Cover Source, and Cloud Cover Description.

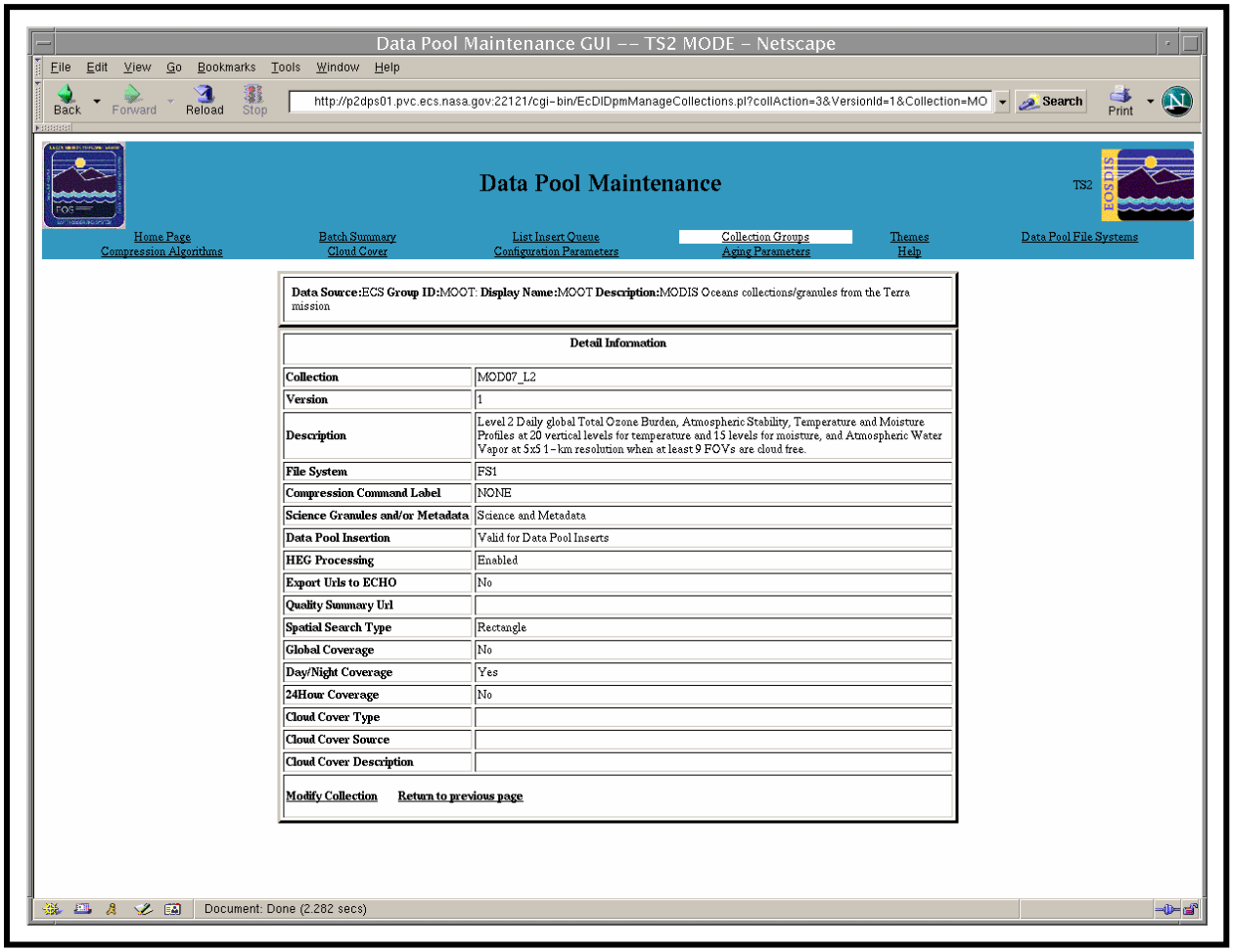


Figure 49. Collection Detail Page

The **Collection Group Detail** and **Collection Detail** pages provide a means of determining what collections within a collection group have been designated valid for Data Pool insertion and whether the insertion is for science granules and metadata or metadata only.

At the bottom of the **Collection Detail** page, there is a link that permits a full-capability operator to **Modify Collection**. Figure 50 shows the **Modify Collection** page obtained by clicking on a **Modify Collection** link. On this page, a full-capability operator can modify many of the characteristics of the collection then implement the changes with a click on the **Apply Change** button at the bottom.

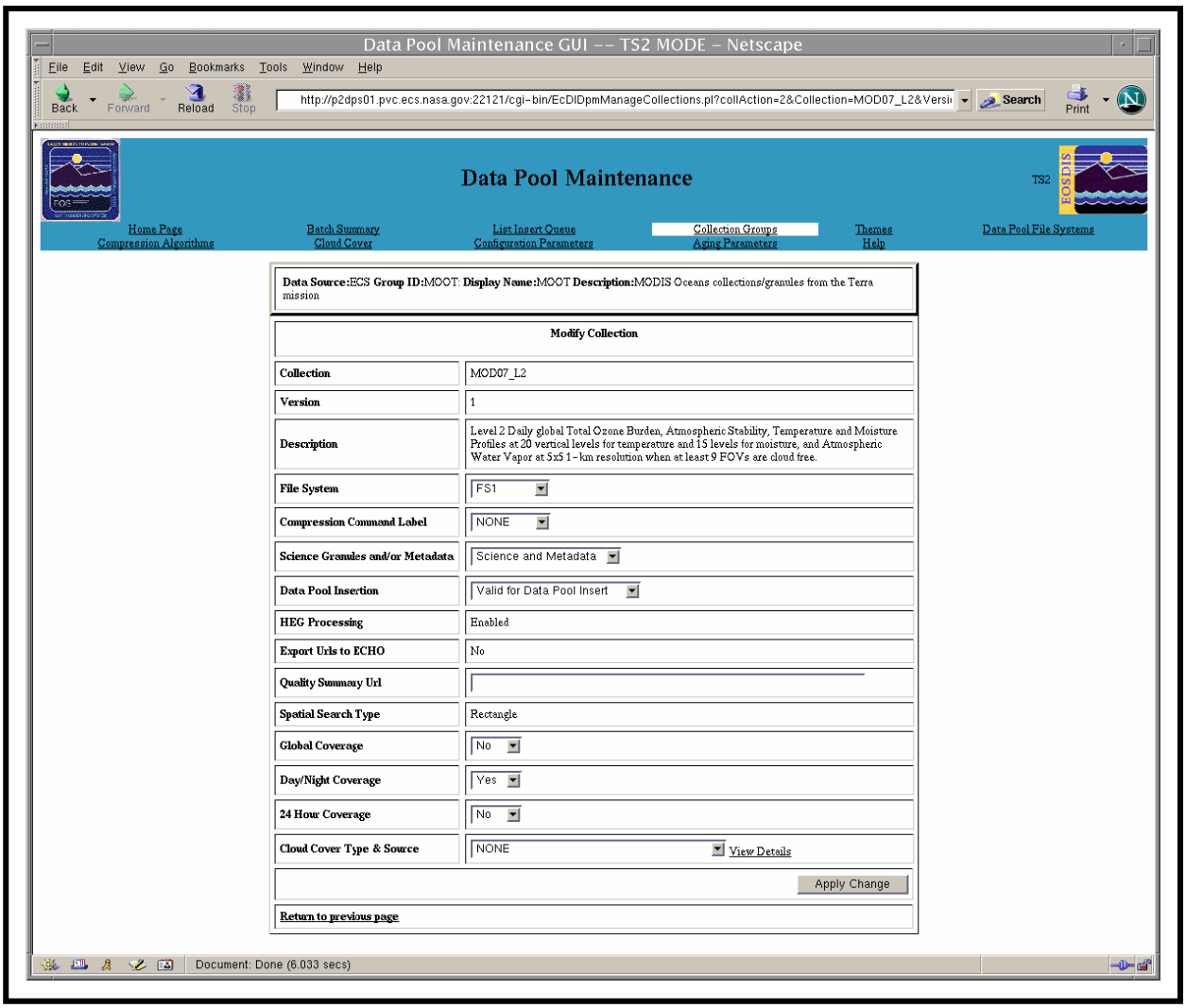


Figure 50. Modify Collection Page

If a full-capability operator clicks on an **Add New Collection** link at the bottom of a **Collection Group Detail** page for an ECS collection, a **Collections Not in Data Pool** page (Figure 51) is displayed. The page lists ECS collections that are not currently part of a Data Pool collection group. The operator can select an ECS collection to add to the collection group by clicking on the link in the **Collection (Click on collection to add)** column of the table on the page. That causes an **Add New Collection** page (Figure 52) to be displayed. The Collection, Version, Description, and Spatial Search Type fields are filled in when the page comes up. The page has fields and option lists for entering the remaining data concerning the collection (e.g., File System, Compression Command Label, and Science Granules and/or Metadata). After the operator enters the appropriate data concerning the ECS collection, clicking on the **Apply Change** button at the bottom of the page applies the changes to the Data Pool database and refreshes the **Collection Group Detail** page.

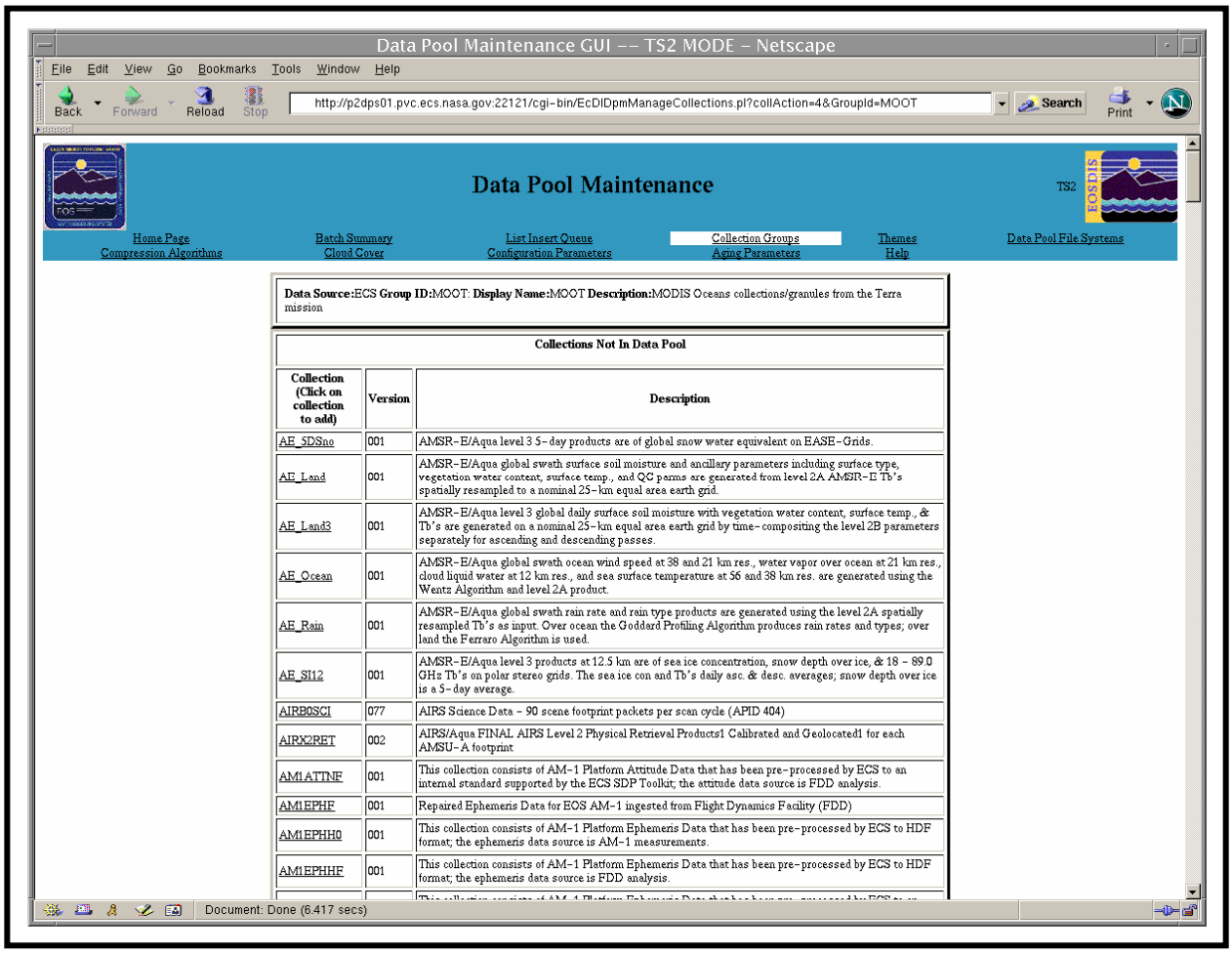


Figure 51. Collections Not in Data Pool Page

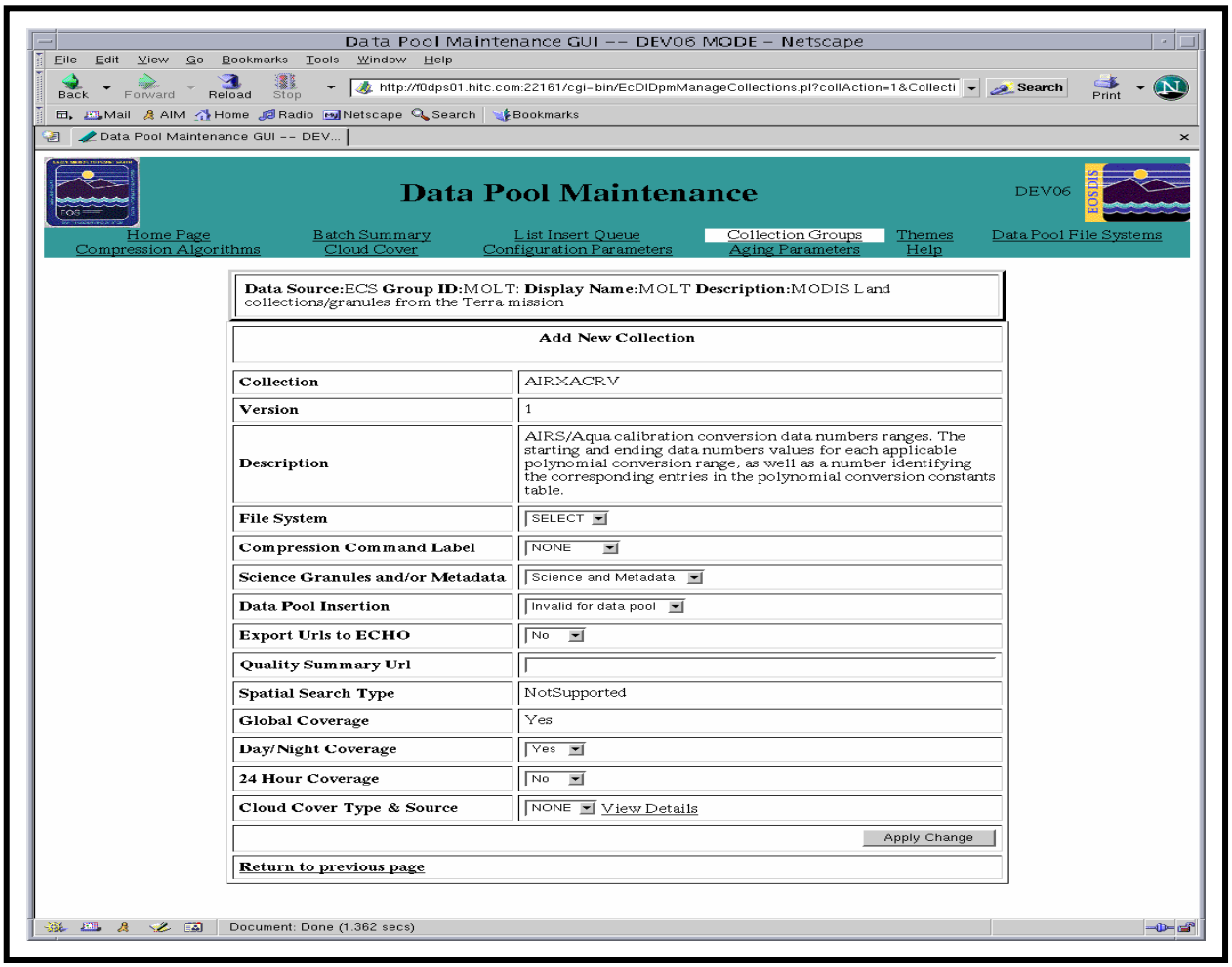


Figure 52. Add New [ECS] Collection Page

Adding a new collection to a non-ECS collection group is somewhat different from adding a new collection to an ECS collection group. To add a new collection to a non-ECS collection group a full-capability operator clicks on the **Add New Collection** link at the bottom of the **Collection Group Detail** page for the relevant non-ECS collection group, causing an **Add New [NON-ECS] Collection** page (Figure 53) to be displayed. The page has fields and option lists for entering data concerning the characteristics (e.g., name, Version, Description, File System, and Compression Command Label) of the non-ECS collection to be added. After the operator enters the appropriate data concerning the non-ECS collection, clicking on the **Apply Change** button at the bottom of the page adds the new collection to the collection group.

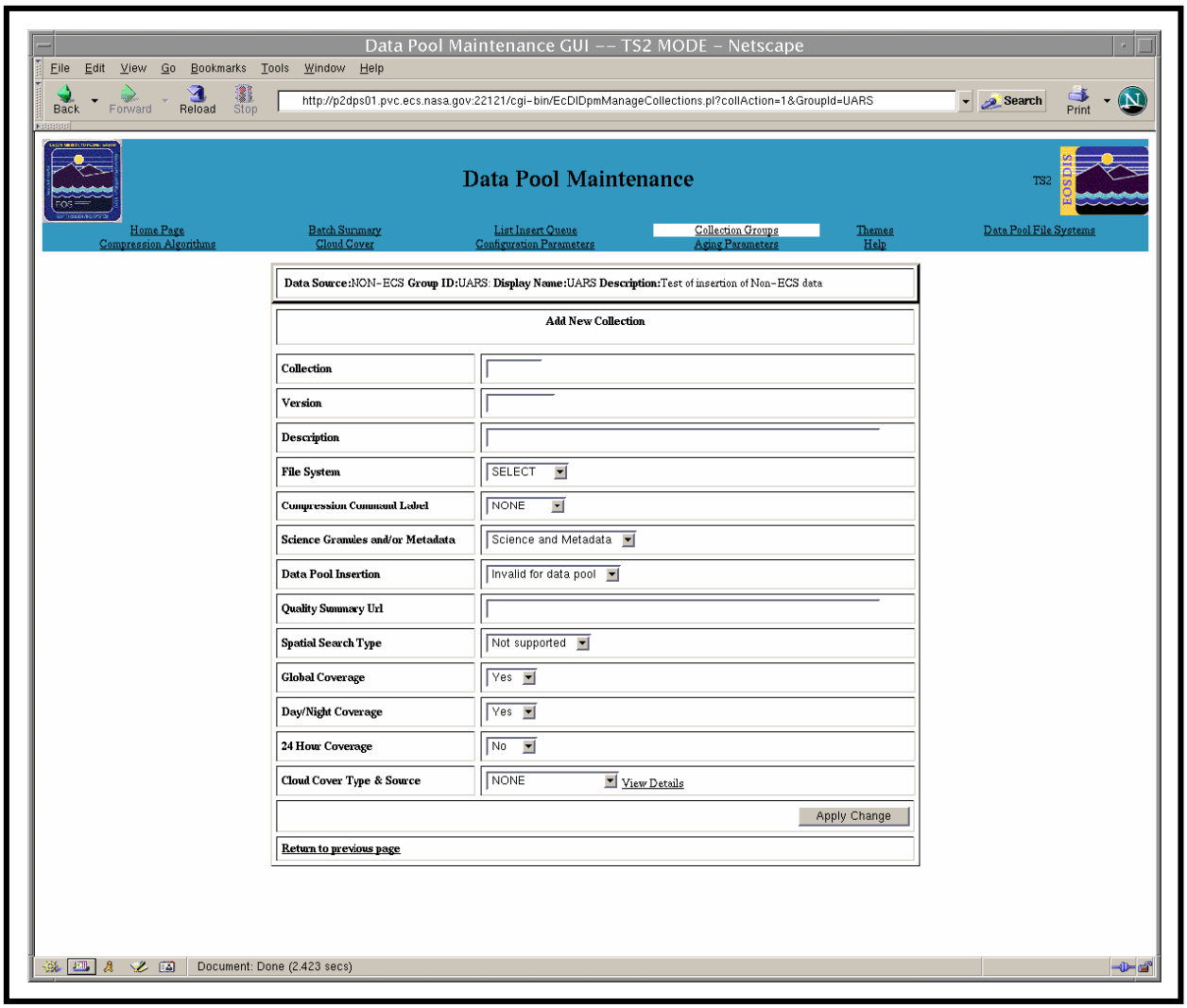


Figure 53. Add New [NON-ECS] Collection Page

Figure 54 shows the **Add Collection Group** page that a full-capability operator obtains by clicking on the **Add Collection Group** link at the bottom of the **Collection Groups** page. On the **Add Collection Group** page the full-capability operator can pick from an option list to specify the Data Source (i.e., ECS or NON-ECS). Then the operator enters a Group ID, Display Name, and Description for the new collection group. Selecting the **Apply Change** button at the bottom of the page implements the addition of the new collection group.

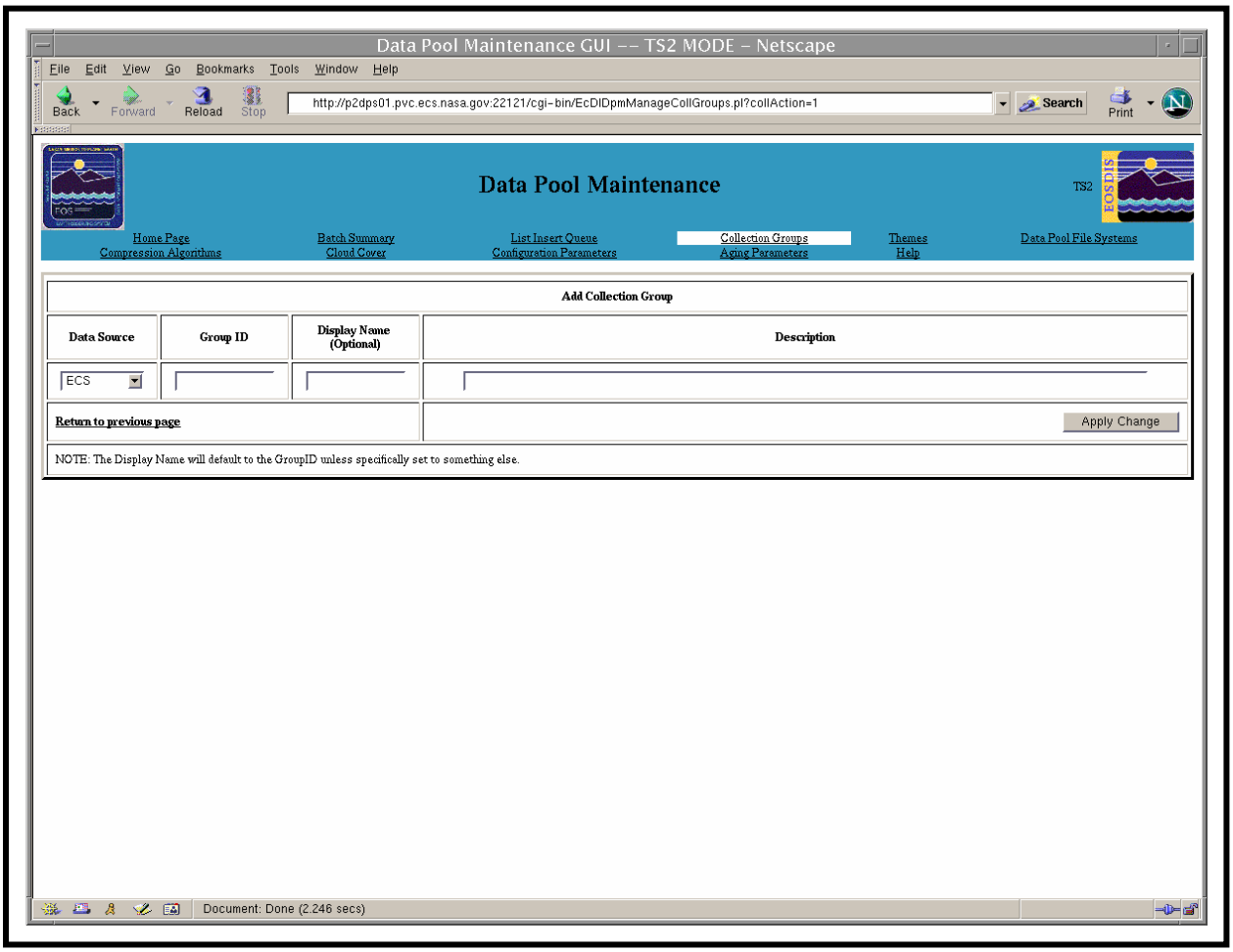


Figure 54. Add Collection Group Page

Figure 55 shows the **Modify Collection Group** page that a full-capability operator obtains by clicking on the **Modify Collection Group** link at the bottom of the **Collection Groups** page. On the **Modify Collection Group** page, the full-capability operator can modify the Display Name and/or Description of one or more collection groups then mark the group for change by checking the box(es) in the last column of the table. The operator implements the change(s) with a click on the **Apply Change** button at the bottom of the page.

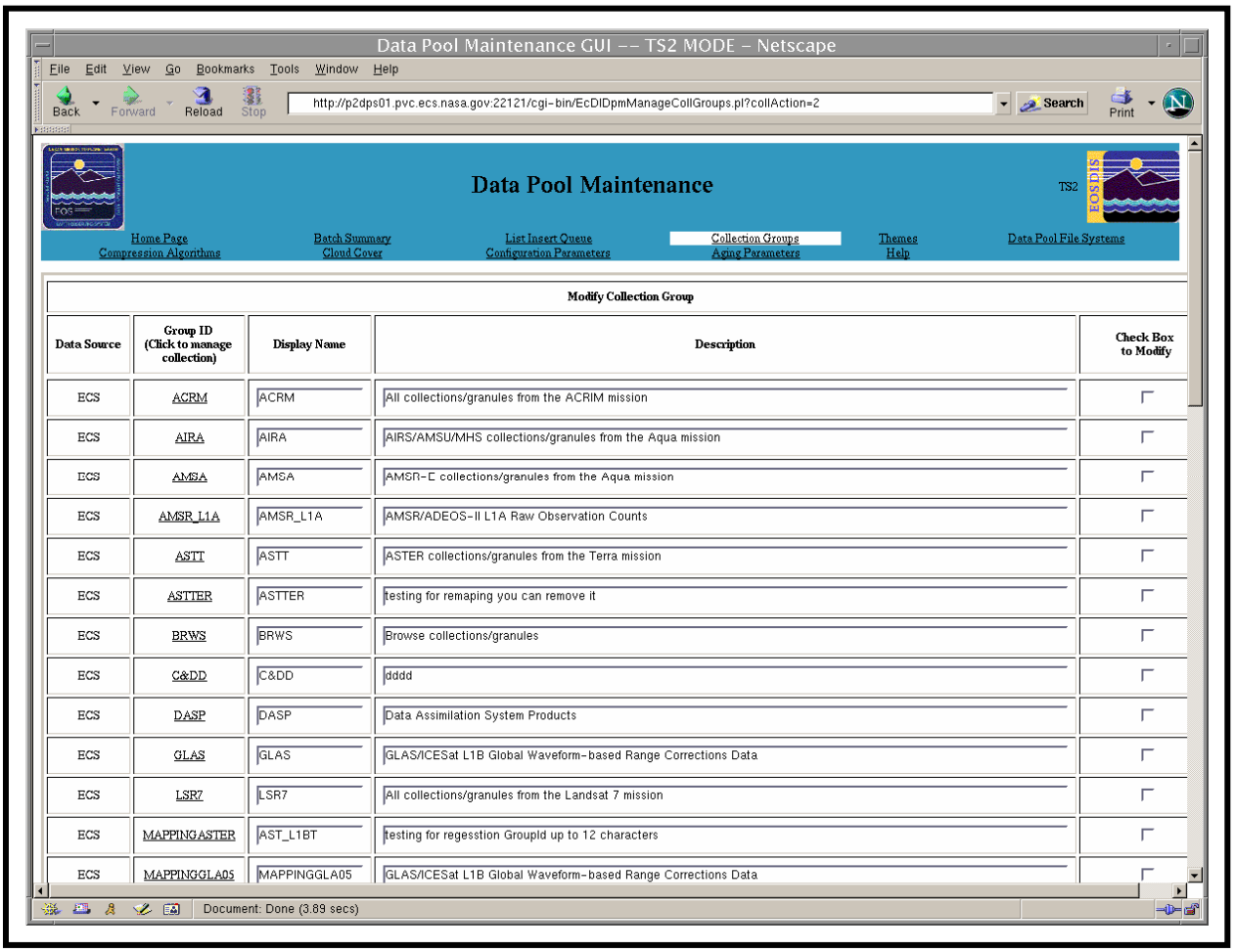


Figure 55. Modify Collection Group Page

Themes

Figure 56 illustrates the **Detailed List of Data Pool Themes** page. The page allows either the full-capability operator or the limited-capability operator to view a list of themes in alphabetical order. The list can be filtered using the option lists for **Web Visible**, **Insert Enabled**, **WCS**, **WMS**, and/or **PreConvert**, and/or typing **Beginning Letters** (of the theme name). After selecting the options, a click on the **Apply Filter** button displays the filtered list of themes. The full-capability operator can delete a theme by selecting the corresponding check box and clicking on the **Apply Change** button. There are **Add New Theme** and **Modify Theme** links providing access to pages for managing those functions. After the operator completes adding a new theme or modifying a theme by clicking on the **Apply Change** button at the pages for those functions, the changes take effect in the Data Pool database and the changes are also reflected in the **Detailed List of Data Pool Themes** page (Figure 56).

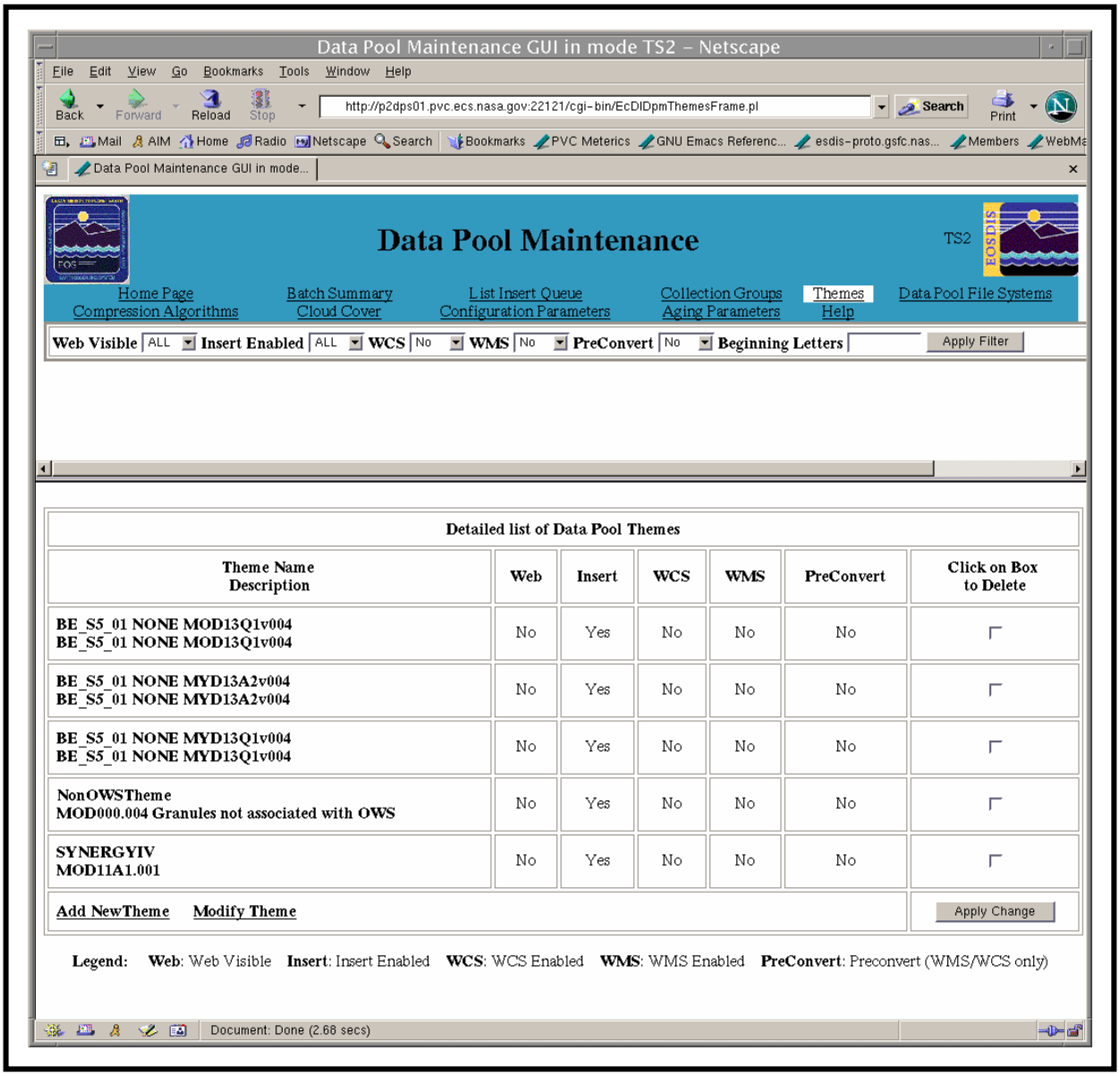


Figure 56. Detailed List of Data Pool Themes Page

If a full-capability operator clicks on the **Add New Theme** link of the **Detailed List of Data Pool Themes** page shown in Figure 56, the **Add New Theme** page (Figure 57) is displayed. To specify a theme, the operator types information in the fields provided for the purpose. **Theme Name** and **Description** are text entry fields. There are check boxes to specify whether the theme is valid for various options (i.e., **Web Visible**, **Insert Enabled**, **WCS**, **WMS**, and/or **PreConvert**) or not. A click on the **Apply Change** button commits the changes to the Data Pool database and updates the **Detailed List of Data Pool Themes** page shown in Figure 56.

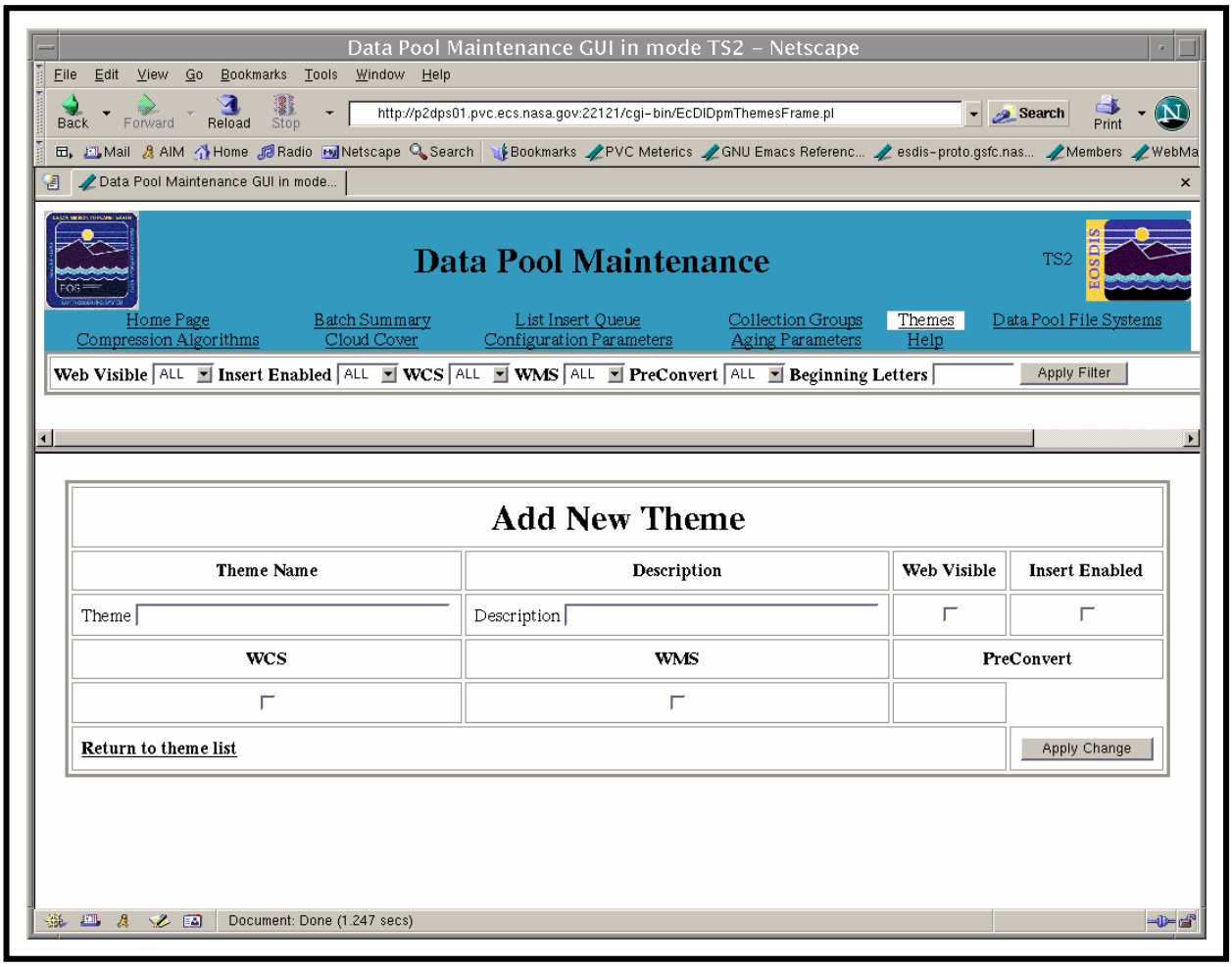


Figure 57. Add New Theme Page

If a full-capability operator clicks on the **Modify Theme** link of the **Detailed List of Data Pool Themes** page shown in Figure 56, the **Modify Theme** page (Figure 58) is displayed. **Theme Name** is the only field that is not editable. The operator can modify the description of a theme by simply retyping in the text area. The operator also can change the various options (i.e., **Web Visible**, **Insert Enabled**, **WCS**, **WMS**, and/or **PreConvert**) by selecting or deselecting the appropriate boxes. A click on the **Apply Change** button commits the changes to the Data Pool database and updates the **Detailed List of Data Pool Themes** page shown in Figure 56.

Help

Figure 59 illustrates the **Help** page that allows both full-capability and limited-capability operators to obtain information on using the **DPM GUI**. The **Help** page describes the features of the other pages of the **DPM GUI**.

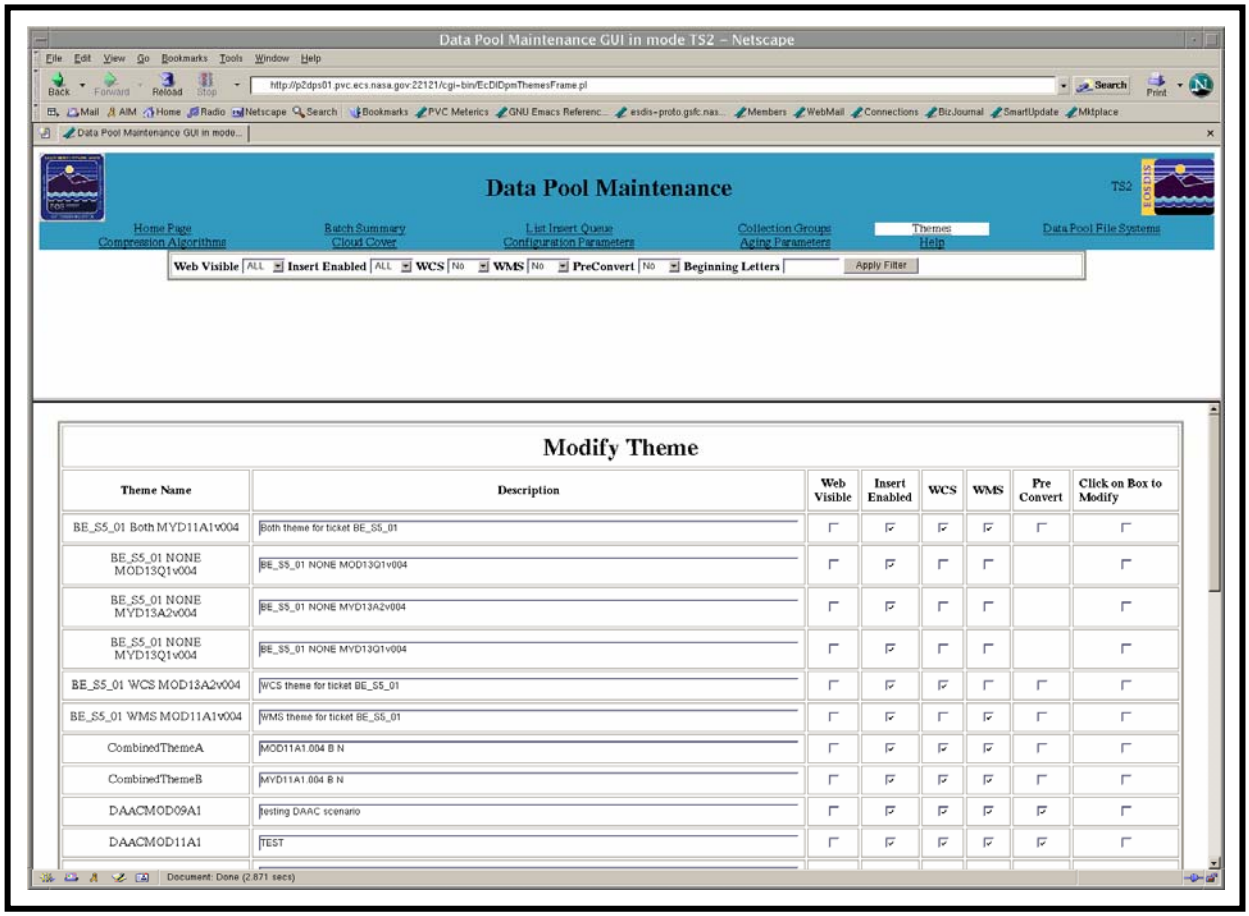
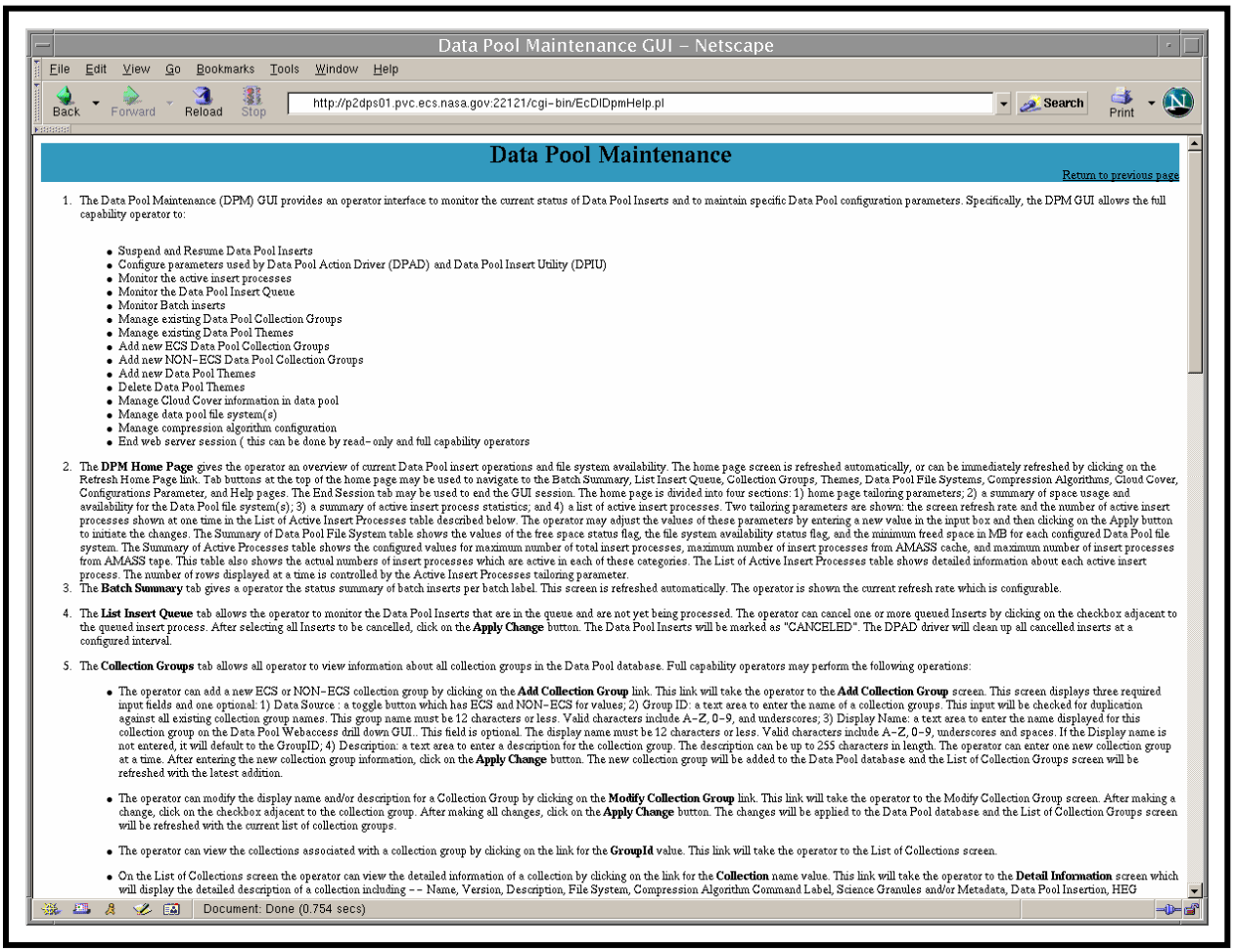


Figure 58. Modify Theme Page



Data Pool Maintenance GUI – Netscape

File Edit View Go Bookmarks Tools Window Help

Back Forward Reload Stop http://p2dps01.pvc.ecs.nasa.gov:22121/cgi-bin/EcDlDpmHelp.pl Search Print

Data Pool Maintenance

Return to previous page

1. The Data Pool Maintenance (DPM) GUI provides an operator interface to monitor the current status of Data Pool Inserts and to maintain specific Data Pool configuration parameters. Specifically, the DPM GUI allows the full capability operator to:

   - Suspend and Resume Data Pool Inserts
   - Configure parameters used by Data Pool Action Driver (DPAD) and Data Pool Insert Utility (DPIU)
   - Monitor the active insert processes
   - Monitor the Data Pool Insert Queue
   - Monitor Batch inserts
   - Manage existing Data Pool Collection Groups
   - Manage existing Data Pool Themes
   - Add new ECS Data Pool Collection Groups
   - Add new NON-ECS Data Pool Collection Groups
   - Add new Data Pool Themes
   - Delete Data Pool Themes
   - Manage Cloud Cover information in data pool
   - Manage data pool file system(s)
   - Manage compression algorithm configuration
   - End web server session (this can be done by read-only and full capability operators

2. The **DPM Home Page** gives the operator an overview of current Data Pool insert operations and file system availability. The home page screen is refreshed automatically, or can be immediately refreshed by clicking on the Refresh Home Page link. Tab buttons at the top of the home page may be used to navigate to the Batch Summary, List Insert Queue, Collection Groups, Themes, Data Pool File Systems, Compression Algorithms, Cloud Cover, Configurations Parameter, and Help pages. The End Session tab may be used to end the GUI session. The home page is divided into four sections: 1) home page tailoring parameters; 2) a summary of space usage and availability for the Data Pool file system(s); 3) a summary of active insert process statistics; and 4) a list of active insert processes. Two tailoring parameters are shown: the screen refresh rate and the number of active insert processes shown at one time in the List of Active Insert Processes table described below. The operator may adjust the values of these parameters by entering a new value in the input box and then clicking on the Apply button to initiate the changes. The Summary of Data Pool File System table shows the values of the free space status flag, the file system availability status flag, and the minimum freed space in MB for each configured Data Pool file system. The Summary of Active Processes table shows the configured values for maximum number of total insert processes, maximum number of insert processes from AMASS cache, and maximum number of insert processes from AMASS tape. This table also shows the actual numbers of insert processes which are active in each of these categories. The List of Active Insert Processes table shows detailed information about each active insert process. The number of rows displayed at a time is controlled by the Active Insert Processes tailoring parameter.
3. The **Batch Summary** tab gives a operator the status summary of batch inserts per batch label. This screen is refreshed automatically. The operator is shown the current refresh rate which is configurable.
4. The **List Insert Queue** tab allows the operator to monitor the Data Pool Inserts that are in the queue and are not yet being processed. The operator can cancel one or more queued Inserts by clicking on the checkbox adjacent to the queued insert process. After selecting all Inserts to be cancelled, click on the **Apply Change** button. The Data Pool Inserts will be marked as "CANCELED". The DPAD driver will clean up all cancelled inserts at a configured interval.
5. The **Collection Groups** tab allows all operator to view information about all collection groups in the Data Pool database. Full capability operators may perform the following operations:

   - The operator can add a new ECS or NON-ECS collection group by clicking on the **Add Collection Group** link. This link will take the operator to the **Add Collection Group** screen. This screen displays three required input fields and one optional: 1) Data Source : a toggle button which has ECS and NON-ECS for values; 2) Group ID: a text area to enter the name of a collection groups. This input will be checked for duplication against all existing collection group names. This group name must be 12 characters or less. Valid characters include A-Z, 0-9, and underscores; 3) Display Name: a text area to enter the name displayed for this collection group on the Data Pool Webaccess drill down GUI.. This field is optional. The display name must be 12 characters or less. Valid characters include A-Z, 0-9, underscores and spaces. If the Display name is not entered, it will default to the GroupID; 4) Description: a text area to enter a description for the collection group. The description can be up to 255 characters in length. The operator can enter one new collection group at a time. After entering the new collection group information, click on the **Apply Change** button. The new collection group will be added to the Data Pool database and the List of Collection Groups screen will be refreshed with the latest addition.
   - The operator can modify the display name and/or description for a Collection Group by clicking on the **Modify Collection Group** link. This link will take the operator to the Modify Collection Group screen. After making a change, click on the checkbox adjacent to the collection group. After making all changes, click on the **Apply Change** button. The changes will be applied to the Data Pool database and the List of Collection Groups screen will be refreshed with the current list of collection groups.
   - The operator can view the collections associated with a collection group by clicking on the link for the **GroupId** value. This link will take the operator to the List of Collections screen.
   - On the List of Collections screen the operator can view the detailed information of a collection by clicking on the link for the **Collection** name value. This link will take the operator to the **Detail Information** screen which will display the detailed description of a collection including -- Name, Version, Description, File System, Compression Algorithm Command Label, Science Granules and/or Metadata, Data Pool Insertion, HEG

Document: Done (0.754 secs)

Figure 59. Help Page

Procedures for Using the Data Pool Maintenance GUI

Launching and Shutting Down the DPM GUI

Let's examine how use 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.

Launch the DPM GUI

1 At the UNIX command shell prompt, type **setenv DISPLAY *clientname*:0.0** and then press the **Return/Enter** key.

- For ***clientname***, use either the local terminal/workstation IP address or its machine name.

2 Start the log-in to a Netscape host by typing **/tools/bin/ssh *hostname*** (e.g., e0ins01, g0ins01, l0ins01, or n0ins02) at the UNIX command shell prompt, and press the **Return/Enter** key.

- 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)?** type **yes** ("**y**" alone does 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, type your ***Passphrase*** and then press the **Return/Enter** key. Go to Step 5.

4 At the **<*user@remotehost*>'s password:** prompt, type your ***Password*** and then press the **Return/Enter** key.

- You are logged in and a UNIX command shell prompt is displayed.

5 Type **netscape &** then press **Return/Enter**.

- It may be necessary to type the path as well as the netscape command (e.g., **/tools/bin/netscape &**).
- 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 Netscape web browser (Figure 29) 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** button (or the **Communicator → Bookmarks** pull-down menu).

- The security login **Prompt** (Figure 30) is displayed.

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

- For example:

  http://x0dps01.daac.ecs.nasa.gov:54321/DataPool.html

- The security login **Prompt** (Figure 30) is displayed.

8 Type the appropriate user name in the **User Name** box of the security login **Prompt**.

9 Type the appropriate password in the **Password** box of the security login **Prompt**.

NOTE: If the security login prompt reappears after the first time the user name and password have been entered (and the **OK** button has been clicked), it may not be due to a data entry problem. Try again to log in using the same user name and password. Sometimes it is necessary to enter the user name and password for the GUI more than once.

10 Click on the appropriate button from the following selections:

- **OK -** to complete the log-in and dismiss the dialogue box.
  - The dialogue box is dismissed.
  - The **DPM GUI Home Page** (Figure 31) is displayed.
- **Cancel -** to dismiss the dialogue box without logging in.
  - The dialogue box is dismissed.
  - The Netscape web browser (Figure 29) is displayed.

At some point it becomes necessary to shut down the **DPM GUI** (end a **DPM GUI** session). The procedure that follows is recommended and is applicable to both full-capability and limited-capability operators.

Shut Down the DPM GUI (End a DPM GUI Session)

---

1 Click on the **Home Page** link at the top of the **DPM GUI**.

- The **DPM GUI Home Page** is displayed.

2 Click on the **End Session** link at the top of the **Home Page**.

- A log-out page containing the following message is displayed.

  Click on Button Below to End Session: NOTE: THIS WOULD ALSO SHUT DOWN THE BROWSER :

NOTE: To abort the log-out and return to the **Home Page**, click on the browser **Back** button.

3 Click on the **ShutDown** button.

- The Netscape browser is dismissed.

Monitoring Data Pool Active Insert Processes and Insert Actions

You may wish to keep an instance of the **DPM GUI** displayed to monitor Data Pool Active Insert Processes. The procedure for using the **DPM GUI** to monitor Data Pool active insert processes is applicable to both full-capability and limited-capability operators.

Monitor Data Pool Active Insert Processes Using the DPM GUI

---

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 displayed.

2 Observe information displayed on the **DPM GUI Home Page**.

- The **Home Page** has the following links for access to Data Pool maintenance function pages:
  - **Data Pool File Systems.**
  - **Compression Algorithms.**
  - **Cloud Cover.**
  - **List Insert Queue.**
  - **Batch Summary.**
  - **Collection Groups.**
  - **Themes.**
  - **Configuration Parameters.**
  - **End Session.**
- The **Home Page** has a summary of Data Pool file systems with the following columns:
  - **File System Label** (label representing an existing Data Pool file system).

- **Free Space Flag** (if set to "Y," free space is available for inserts; "N" means free space is not available).
- **Availability** (if set to "Y," the file system is currently available for Data Pool insert; "N" means the file system is not available for Data Pool insert).
- **Min Freed Space in MB** (value that represents the minimum amount of freed space in the file system in megabytes; it is an amount of space must remain free in order to make the file system available for insert).

- The **Home Page** has a summary of active processes with the following rows:
  - **Maximum allowed processes.**
  - **Maximum allowed processes from AMASS cache.**
  - **Maximum allowed processes from AMASS tape.**
  - **Total number of active insert processes running.**
  - **Number of active insert processes using AMASS cache.**
  - **Number of active insert processes using AMASS tape.**
- The **Home Page** has a table of active insert processes showing the following columns of detailed information for each process:
  - **Unix ProcessId** (UNIX process identifier).
  - **EcsID** (ECS identifier or Granule ID for the granule being processed).
  - **Collection** (to which the granule belongs).
  - **Version** (for the collection to which the granule belongs).
  - **StartTime** (time at which the insert processing started).
  - **StatusTime** (time at which the status listed in the **Status** column was achieved).
  - **Status** (current state of the insert process).
  - **AMASS Cache** [availability (**Y** or **N**) of the granule being processed].
  - **Retries** [number of attempts by the process to recover from retryable errors (e.g., Data Pool disk temporarily unavailable, Data Pool directory does not exist, or Data Pool database temporarily unavailable)].

NOTE: The system is designed for rapid insertion of data into the Data Pool by quickly processing data that are available in cache, such as data that are staged for archiving. If the insert processing is delayed and the data are removed from cache, the Data Pool insert is likely to fail.

3 To obtain an immediate screen refresh, click on the **Refresh Home Page** link near the upper right corner of the display.

- The displayed data are updated.

NOTE: The screen refreshes automatically at intervals determined by the number of seconds specified in the **Screen Refresh Rate** field.

4 To change the automatic screen refresh rate first type the desired number of seconds between refreshes in the **Screen Refresh Rate** text entry box.

5 To complete changing the automatic screen refresh rate click on the **Apply** button adjacent to the **Screen Refresh Rate** text entry box.

- The **Screen Refresh Rate** is changed to the new value.

6 To change the number of active insert processes displayed at a time in the **List of Active Insert Processes** table on the **Home Page** first type the desired number of rows to be displayed in the **Active Insert Processes** text entry box.

7 To complete changing the number of active insert processes displayed at a time in the **List of Active Insert Processes** table on the **Home Page** click on the **Apply** button adjacent to the **Active Insert Processes** text entry box.

- The number of active insert processes displayed at a time in the **List of Active Insert Processes** table is changed to the new value.

Managing Data Pool File Systems Using the DPM GUI

The **DPM GUI File System Information** page permits both full-capability and limited-capability operators users to view a list of Data Pool file systems and obtain information on the status of the free space flag, availability for insert, and minimum freed space for each file system. In addition it has links that allow full-capability operators to add new Data Pool file systems or modify existing file system information.

View a List of Data Pool File Systems Using the DPM GUI

---

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**, **Compression Algorithms**, **Cloud Cover**, **List Insert Queue**, **Batch Summary**, **Collection Groups**, **Themes, Configuration Parameters, Aging Parameters, and End Session**).

2 Click on the **Data Pool File Systems** link.

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

3 Observe data displayed on the **File System Information** page.

- The table on the **File System Information** page has columns containing the following types of Data Pool file system information:
  - **File System Label.**
  - **Absolute Path.**
  - **Free Space Flag.**
  - **Availability.**
  - **Min Free Space (in Megabytes).**
- The following links are available on the **File System Information** page:
  - **Add New File System.**
  - **Modify File System.**

The **DPM GUI** may be used to modify a Data Pool file system. This is useful if the Absolute Path, Free Space Flag, Availability (for Insert), and/or Min. Freed Space for a particular Data Pool file system need to be corrected or updated. Full-capability operators (only) can use the following procedure to modify a Data Pool file system:

Modify a Data Pool File System Using the DPM GUI

---

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**, **Compression Algorithms**, **Cloud Cover**, **List Insert Queue**, **Batch Summary**, **Collection Groups**, **Themes, Configuration Parameters, Aging Parameters, and End Session**).

2 Click on the **Data Pool File Systems** link.

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

3 Click on the **Modify File System** link at the bottom of the list of file systems (scrolling down if necessary).

- The **Modify File System Information** page is displayed, providing a table of Data Pool file system information showing six columns: **File System Label**, **Absolute Path**, **Free Space Flag**, **Availability**, **Min Free Space (in Megabytes)**, and **Click on box to modify** (containing a check box to mark the file system for change).
- There is an **Apply Change** button at the bottom of the page to implement changes.

4 To change the absolute path for a file system type the desired path in the **Absolute Path** field for the file system.

- The basic ftp root directory path is shown above the text entry box; data entered in the box will be appended to the base path shown.

5 To change a file system's free space flag setting click on the appropriate button in the **Free Space Flag** column.

- The following choices are available:
  - **ON.**
  - **OFF.**

6 To change the setting for a file system's availability for data insert click on the appropriate button in the **Availability** column.

- The following choices are available:
  - **YES.**
  - **NO.**

7 To change the minimum freed space for a file system type the desired value (in megabytes) in the appropriate **Min Free Space (in Megabytes)** field.

8 Click in the check box at the end of the row containing file system information to be modified.

- The selected file system information is marked for subsequent modification.

9 Repeat Steps 4 through 8 for any additional file systems to be modified.

10 Click on the **Apply Change** button.

- The revised file system information is entered in the Data Pool database.
- The **File System Information** page is displayed with the modified file system information.

Full-capability operators (only) can use the following procedure to add a Data Pool file system:

Add a Data Pool File System Using the DPM GUI

---

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**, **Compression Algorithms**, **Cloud Cover**, **List Insert Queue**, **Batch Summary**, **Collection Groups**, **Themes, Configuration Parameters, Aging Parameters, and End Session**).

2 Click on the **Data Pool File Systems** link.

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

3 Click on the **Add New File System** link at the bottom of the list of file systems (scrolling down if necessary).

- The **Add New File System Information** page is displayed, providing a table of Data Pool file system information showing five rows: **Label**, **Absolute Path**, **Free Space Flag**, **Availability**, and **Min Freed Space (in Megabytes)**.
- There is an **Apply Change** button at the bottom of the page to implement the new file system.

4 Type the desired file system label in the **Label** field.

- Enter a unique name with no more than 25 characters.

5 Type the desired path in the **Absolute Path** field.

- The basic ftp root directory path is shown adjacent to the text entry box; data entered in the box will be appended to the base path shown.

6 To display free space flag options click on the **Free Space Flag** option button.

- **Free Space Flag** options are displayed (i.e., **ON** and **OFF**).

7 To select a free space flag option click on the appropriate choice from the option list.

- **ON** should be selected if there is enough free space in the file system for inserts.
- **OFF** should be selected if there is not enough free space in the file system for inserts.

8 To display availability options click on the **Availability** option button.

- **Availability** options are displayed (i.e., **YES** and **NO**).

9 To select an availability option click on the appropriate choice from the option list.

- **YES** should be selected if the file system is currently available for inserts.
- **NO** should be selected if the file system is not currently available for inserts.

10 Type the desired value for minimum freed space (in megabytes) in the **Min Freed Space (in Megabytes)** field.

- **Min Freed Space** indicates how much space needs to be available to keep the file system available for insert.

11 Click on the **Apply Change** button.

- The file system information is entered in the Data Pool database.
- The **File System Information** page is displayed with the new file system information.

Enabling/Disabling Data Compression Using the DPM GUI

The data compression capability for Data Pool is implemented using DAAC-provided compression algorithms. When compression is turned on at the Data Pool subsystem level, all science granules will be compressed at the time of Data Pool insert, if there is a compression algorithm associated with the granule's collection. For each compressed granule, the name and size of the granule's compressed file(s), the size of the original uncompressed file(s), and the compression type will be stored in the file table in the Data Pool database. If checksumming is turned on, the checksum of the compressed file(s) is stored in the Data Pool database. Metadata and browse files are not compressed on insert.

The procedure for using the **DPM GUI** to enable or disable data compression can be performed by full-capability operators only. It involves changing the value for the **CompressOnInsert** configuration parameter in the Data Pool database. Changing the value from OFF to ON enables data compression. Changing the value from ON to OFF disables data compression. Limited-capability operators are not allowed to change the values assigned to configuration parameters.

Enable/Disable Data Compression Using the DPM GUI

---

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**, **Compression Algorithms**, **Cloud Cover**, **List Insert Queue**, **Batch Summary**, **Collection**

Groups, **Themes, Configuration Parameters, Aging Parameters, and End Session**).

2 Click on the **Configuration Parameters** link.

- The **List of Configuration Parameters** page is displayed, providing a table of parameters showing three columns: **Parameter Name**, **Parameter Value** (including an entry field with current value, followed by a brief description of the parameter), and **Click on Box to Modify Parm** (containing a check box to mark the parameter for change).
- There is an **Apply Change** button at the bottom of the page to implement any selected change(s).

3 To display **CompressOnInsert** options click on the option button in **Parameter Value** column of the row for the **CompressOnInsert** parameter.

- The following choices are available:
  - **ON.**
  - **OFF.**

4 To select a **CompressOnInsert** option click on the appropriate choice from the option list.

- Selected option is displayed in the field.

5 In the row for the **CompressOnInsert** parameter click in the check box in the **Click on Box to Modify Parm** column.

- The box is filled to indicate selection.

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

- The screen is refreshed, the check box is unfilled, and the displayed **Parameter Value** for **CompressOnInsert** reflects the change.

Managing Compression Algorithms Using the DPM GUI

The **DPM GUI Manage Compression Algorithms** link permits both full-capability operators and limited-capability operators to view the current compression algorithms. The full-capability operator only may add, modify, or deactivate compression algorithms. The procedures that follow are applicable.

View a List of Compression Algorithms Using the DPM GUI

---

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**, **Compression Algorithms**, **Cloud Cover**, **List Insert Queue**, **Batch Summary**, **Collection Groups**, **Themes, Configuration Parameters, Aging Parameters, and End Session**).

2 Click on the **Compression Algorithms** link.

- The **Compression Algorithms** page is displayed.

3 Observe data displayed on the **Compression Algorithms** page.

- The table on the **Compression Algorithms** page has columns containing the following types of compression algorithm information:
  - **Compression Label.**
  - **File Extension.**
  - **Compression Command.**
  - **Decompression Command.**
- The following links are available on the **Compression Algorithms** page:
  - **Add Compression Algorithm.**
  - **Modify Compression Algorithm.**
  - **Deactivate Compression Algorithm.**

The **DPM GUI** may be used to modify compression algorithms. This can be useful if the file extension, compression command, or decompression command needs to be modified or updated. A full-capability operator may use the procedure that follows to modify compression algorithms.

Modify Compression Algorithms Using the DPM GUI

---

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**, **Compression Algorithms**, **Cloud Cover**, **List Insert Queue**, **Batch Summary**, **Collection Groups**, **Themes, Configuration Parameters, Aging Parameters, and End Session**).

2 Click on the **Compression Algorithms** link.

- The **Compression Algorithms** page is displayed, providing a table listing all compression algorithms and showing algorithm-related information in four columns: **Compression Label**, **File Extension**, **Compression Command**, and **Decompression Command**. There are links (i.e., **Add Compression Algorithm**, **Modify Compression Algorithm** and **Deactivate Compression Algorithm**) that can be selected.

3 Click on the **Modify Compression Algorithm** link at the bottom of the list of **Compression Algorithms** (scrolling down if necessary).

- The **Modify Compression Algorithm** page is displayed.

4 To modify the default file extension for a compression algorithm type the desired file extension in the appropriate **File Extension** field.

- The file extension may have no more than ten characters.
- The typed entry is displayed in the field.

5 To modify the compression command for a compression algorithm type the desired compression command in the appropriate **Compression Command** field.

- The compression command may have no more than 255 characters.
- Include the full path and parameters.
- Use “%infile” to represent the file to be compressed.
  - For example:

    /usr/bin/gzip –1 %infile

- The typed entry is displayed in the field.

6 To modify the decompression command for a compression algorithm type the desired decompression command in the appropriate **Decompression Command** field.

- The decompression command may have no more than 255 characters.
- Include the full path and parameters.
- Use “%infile” to represent the file to be compressed.
  - For example:

    /usr/bin/gunzip –1 %infile
- The typed entry is displayed in the field.

7 Click in the check box at the end of the row containing modified compression algorithm information.

- The compression algorithm is marked for subsequent modification. (A check mark is displayed in the selected check box.)

8 Repeat Steps 4 through 7 as necessary to modify additional compression algorithms.

9 Click on the **Apply Change** button.

- The modified compression algorithm information is added to the Data Pool Database.
- The **Compression Algorithms** page is displayed with the revised compression algorithm information.

A full-capability operator may use the following procedure to add a compression algorithm.

Add a Compression Algorithm Using the DPM GUI

---

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**, **Compression Algorithms**, **Cloud Cover**, **List Insert Queue**, **Batch Summary**, **Collection Groups**, **Themes, Configuration Parameters, Aging Parameters, and End Session**).

2 Click on the **Compression Algorithms** link.

- The **Compression Algorithms** page is displayed, providing a table listing all compression algorithms and showing algorithm-related information in four columns:

Compression Label, **File Extension**, **Compression Command**, and **Decompression Command**. There are links (i.e., **Add Compression Algorithm**, **Modify Compression Algorithm** and **Deactivate Compression Algorithm**) that can be selected.

3 Click on the **Add Compression Algorithm** link at the bottom of the list of compression algorithms (scrolling down if necessary).

- The **Add Compression Algorithm** page is displayed, providing a table of compression algorithms showing four columns: **Compression Label**, **File Extension**, **Compression Command**, and **Decompression Command**.
- There is an **Add Algorithm** button at the bottom of the page to implement the new algorithm.

4 Type the desired compression label in the **Compression Label** field.

- The label may have no more than ten characters.
- The typed entry is displayed in the field.

5 If applicable, type the desired default file extension in the **File Extension** field.

- The file extension may have no more than ten characters.
- The typed entry is displayed in the field.

6 Type the compression command in the **Compression Command** field.

- The compression command may have no more than 255 characters.
- Include the full path and parameters.
- Use "%infile" to represent the file to be compressed.
  - For example:

    /usr/bin/gzip –1 %infile
- The typed entry is displayed in the field.

7 If applicable, type the decompression command in the **Decompression Command** field.

- The decompression command may have no more than 255 characters.
- Include the full path and parameters.
- Use "%infile" to represent the file to be compressed.
  - For example:

    /usr/bin/gunzip –1 %infile
- The typed entry is displayed in the field.

8 Click on the **Add Algorithm** button.

- The compression algorithm is added to the Data Pool Database.
- The **Compression Algorithms** page is displayed with the new compression algorithm.

A full-capability operator may use the following procedure to deactivate a compression algorithm:

Deactivate a Compression Algorithm Using the DPM GUI

NOTE: Deactivating a compression algorithm removes the algorithm from the list of compression algorithms and dissociates it from all collections. However, it will still be possible to decompress any granules compressed with the algorithm.

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**, **Compression Algorithms**, **Cloud Cover**, **List Insert Queue**, **Batch Summary**, **Collection Groups**, **Themes, Configuration Parameters, Aging Parameters, and End Session**).

2 Click on the **Compression Algorithms** link.

- The **Compression Algorithms** page is displayed, providing a table listing all compression algorithms and showing algorithm-related information in four columns: **Compression Label**, **File Extension**, **Compression Command**, and **Decompression Command**. There are links (i.e., **Add Compression Algorithm**, **Modify Compression Algorithm** and **Deactivate Compression Algorithm**) that can be selected.

3 Click on the **Deactivate Compression Algorithm** link at the bottom of the list of **Compression Algorithms** (scrolling down if necessary).

- The **Deactivate Compression Algorithm** page is displayed, providing a table of compression algorithms showing five columns: **Compression Label**, **File Extension**, **Compression Command**, **Decompression Command**, and **Check Box to Deactivate** (containing check boxes to mark algorithms for deactivation).
- There is a **Deactivate Selected** button at the bottom of the page to implement the deactivation(s).

4 To select compression algorithms to be marked for deactivation click in the appropriate check box(es) in the column on the far right of the **Deactivate Compression Algorithm** page.

- The compression algorithm(s) is (are) marked for subsequent deactivation. (A check mark is displayed in each selected check box.)

5 Click on the **Deactivate Selected** button.

- The selected compression algorithms are deactivated.

Managing Cloud Cover Information Using the DPM GUI

The **DPM GUI Manage Cloud Cover** link permits both full-capability and limited-capability operators to view all cloud cover information in the Data Pool database. In addition it allows full-capability operators to add new cloud cover information, modify cloud cover source descriptions, or delete cloud cover information. The procedures that follow are applicable to using the **DPM GUI** to manage cloud cover information.

View Cloud Cover Information Using the DPM GUI

---

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**, **Compression Algorithms**, **Cloud Cover**, **List Insert Queue**, **Batch Summary**, **Collection Groups**, **Themes, Configuration Parameters, Aging Parameters, and End Session**).

2 Click on the **Cloud Cover** link.

- The **Cloud Cover Information** page is displayed.

3 Observe data displayed on the **Cloud Cover Information** page.

- The table on the **Cloud Cover Information** page has columns containing the following types of cloud cover information:
  - **Source Type.**
  - **Source Name.**
  - **Source Description.**

- The following links are available on the **Cloud Cover Information** page:
  - **Add New Cloud Cover.**
  - **Modify Source Description.**
- An **Apply Change** button is available for deleting cloud cover information from the Data Pool database.

A full-capability operator may use the following procedure to add new cloud cover information:

Add New Cloud Cover Information Using the DPM GUI

---

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**, **Compression Algorithms**, **Cloud Cover**, **List Insert Queue**, **Batch Summary**, **Collection Groups**, **Themes, Configuration Parameters, Aging Parameters, and End Session**).

2 Click on the **Cloud Cover** link.

- The **Cloud Cover Information** page is displayed, providing a table listing all cloud cover information; i.e., **Source Type**, **Source Name**, and **Source Description**.
- The following links are available: **Add New Cloud Cover** and **Modify Source Description**.
- An **Apply Change** button is available for deleting cloud cover information from the Data Pool database.

3 Click on the **Add New Cloud Cover** link at the bottom of the **Cloud Cover Information** page (scrolling down if necessary).

- The **Add New Cloud Cover** page is displayed, providing a table of cloud cover information showing three rows: **Source Type**, **Source Name**, and **Source Description**.
- There is an **Apply Change** button at the bottom of the page to implement changes.

4 To display source type options click on the **Source Type** option button.

- Source type options are displayed (e.g., **Core Metadata** and **PSA**).

5 To select a source type click on the appropriate source type from the option list.

- If **Core Metadata** was selected, the **Source Name** field is automatically filled in.

6 To specify a source name type the desired name in the **Source Name** field.

- If **Core Metadata** was selected as the source type, the **Source Name** field is automatically filled in and cannot be edited.

7 Type a description of the cloud cover information in the **Source Description** field.

- The description may be up to 255 characters in length.

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

- The source name is validated against the Science Data Server database.
- The new cloud cover information is entered in the Data Pool database.
- The **Cloud Cover Information** page is displayed with the new cloud cover information.

Full-capability operators may use the following procedure to modify cloud cover source descriptions:

Modify Cloud Cover Source Descriptions Using the DPM GUI

---

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**, **Compression Algorithms**, **Cloud Cover**, **List Insert Queue**, **Batch Summary**, **Collection Groups**, **Themes, Configuration Parameters, Aging Parameters, and End Session**).

2 Click on the **Cloud Cover** link.

- The **Cloud Cover Information** page is displayed, providing a table listing all cloud cover information; i.e., **Source Type**, **Source Name**, and **Source Description**.
- The following links are available: **Add New Cloud Cover** and **Modify Source Description**.
- An **Apply Change** button is available for deleting cloud cover information from the Data Pool database.

3 Click on the **Modify Source Description** link at the bottom of the **Cloud Cover Information** page (scrolling down if necessary).

- The **Modify Source Description** page is displayed, providing a table of cloud cover information showing four columns: **Source Type**, **Source Name**, **Source Description**, and **Click on box to modify** (containing a check box to mark the source description for change).
- There is an **Apply Change** button at the bottom of the page to implement changes.

4 To start the process of changing a source description type the desired description in the appropriate **Source Description** field.

5 To continue the process of changing a source description click in the check box at the end of the row containing modified source description information.

- The source description is marked for subsequent modification. (A check mark is displayed in the selected check box.)

6 Repeat Steps 4 and 5 for any additional source descriptions to be modified.

7 Click on the **Apply Change** button.

- The revised source description information is entered in the Data Pool database.
- The **Cloud Cover Information** page is displayed with the modified cloud cover information.

A full-capability operator may use the following procedure to delete cloud cover information:

Delete Cloud Cover Information Using the DPM GUI

---

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**, **Compression Algorithms**, **Cloud Cover**, **List Insert Queue**, **Batch Summary**, **Collection Groups**, **Themes, Configuration Parameters, Aging Parameters, and End Session**).

2 Click on the **Cloud Cover** link.

- The **Cloud Cover Information** page is displayed, providing a table listing all cloud cover information; i.e., **Source Type**, **Source Name**, and **Source Description**.
- The following links are available: **Add New Cloud Cover** and **Modify Source Description**.
- An **Apply Change** button is available for deleting cloud cover information from the Data Pool database.

3 Click in the check box(es) at the end of the row(s) containing cloud cover information to be deleted.

- The selected source(s) is (are) marked for subsequent deletion.

4 Click on the **Apply Change** button.

- The selected source(s) is (are) deleted from the Data Pool database.
- If any cloud cover information is associated with any collection, it will not be deleted.
- The **Cloud Cover Information** page is displayed with the modified cloud cover information.

Checking the Status of Batch Inserts

The **DPM GUI** provides a page to display a summary of the status of batch Data Pool inserts made using the Synergy batch insert utility. The procedure that follows is applicable to both full-capability and limited-capability operators.

Check the Status of Batch Inserts Using the DPM GUI

---

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**, **Compression Algorithms**, **Cloud Cover**, **List Insert Queue**, **Batch Summary**, **Collection Groups**, **Themes, Configuration Parameters, Aging Parameters, and End Session**).

2 Click on the **Batch Summary** link.

- The **Batch Summary** page is displayed.

- The GUI displays the **Batch Summary** page, providing for each batch label the numbers of inserts for that label that are **New**, **Completed**, **Failed**, in **Retry**, and **Canceled**.
- The page also shows the screen refresh rate in minutes; the rate may be changed by clicking in the **Screen Refresh Rate** field, replacing the displayed value with the desired value, and clicking on the **Apply Refresh Rate** button.

3 Observe data displayed on the **Batch Summary** page.

- The table on the **Batch Summary** page has columns containing the following types of information:
  - **Batch Label.**
  - **New** (number of inserts for the label that are new).
  - **Completed** (number of inserts for the label that have been completed).
  - **Failed** (number of inserts for the label that have failed).
  - **Retry** (number of inserts for the label that have been retried).
  - **Canceled** (number of inserts for the label that have been canceled).

4 To change the automatic screen refresh rate first type the desired number of minutes between refreshes in the **Screen Refresh Rate** text entry box.

5 To complete changing the automatic screen refresh rate click on the **ApplyRefreshRate** button adjacent to the **Screen Refresh Rate** text entry box.

- The **Screen Refresh Rate** is changed to the new value.

6 Return to Step 3.

Checking the Data Pool Insert Queue and Canceling a Data Pool Insert Action

The **List Insert Queue** page of the **DPM GUI** provides a list of Data Pool inserts left to process that both full-capability and limited-capability operators can view. It also provides for each listed insert a check box permitting a full-capability operator to mark queued inserts for cancellation, and an **Apply Change** button to implement the cancellation.

Check the Data Pool Insert Queue and Cancel a Data Pool Insert Action

---

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**, **Compression Algorithms**, **Cloud Cover**, **List Insert Queue**, **Batch Summary**, **Collection Groups**, **Themes, Configuration Parameters, Aging Parameters, and End Session**).

2 Click on the **List Insert Queue** link.

- The **List Insert Queue** page is displayed.

3 Observe data displayed on the **List Insert Queue** page.

- The **List Insert Queue** page shows how many inserts are left to process as of the current date.
- The table on the **List Insert Queue** page has columns containing the following types of insert queue information:
  - **Data Source.**
  - **Batch Label.**
  - **Dispatch Priority.**
  - **RequestID.**
  - **SubID** (subscription identifier of the subscription selected by the software for processing).
  - **ECSID** (ECS identifier or Granule ID for the granule to be processed).
  - **Collection** (to which the granule belongs).
  - **Version** (for the collection to which the granule belongs).
  - **Science Granules and/or Metadata** (indication of whether the insert is to include science granules and metadata or just the metadata).
  - **Enqueue Time** (time when the insert was placed in the insert queue).
  - **Retries** [number of attempts by the process to recover from retryable errors (e.g., Data Pool disk temporarily unavailable, Data Pool directory does not exist, Data Pool database temporarily unavailable)].
  - **Status.**
  - **Click on Box to Cancel** (containing a check box to mark the insert for cancellation).

NOTE: There may be multiple subscriptions specifying insertion of specific data into the Data Pool, but only one insert is needed; therefore, only one of the subscriptions serves as the basis for the insert action. The **SubID** is of no particular significance to an operator and may safely be ignored.

- There is an **Apply Change** button at the bottom of the page for implementing cancellations.
- There is a **Continue** link at the bottom of the page; if there are more inserts than can be displayed in the space of one page, the **Continue** link displays the next page of the list.

4 To cancel an insert first click on the check box at the end of the row of information for the insert to be canceled.

- The insert is marked for subsequent cancellation.
- The check box for the selected insert is filled to indicate selection.

5 Repeat Step 4 for any additional insert to be canceled.

6 To implement the cancellation of insert(s) click on the **Apply Change** button.

- A confirmation message is displayed; it asks "Are you ready to cancel the insert for . . ." and there are links displayed for **Yes, cancel insert** and **No, return to previous page**.

7 To confirm cancellation, click on the **Yes, cancel insert** link.

- The **List Insert Queue** page is displayed with the canceled insert(s) removed and the count of inserts left to process reduced by the number of inserts canceled.

Managing Data Pool Configuration Parameters and Data Pool Tuning

The **List of Configuration Parameters** and **List of Aging Parameters** pages on the **DPM GUI** allow a full-capability operator to set or change values assigned to Data Pool Management configuration parameters. Limited-capability operators have read-only access to the pages.

The following parameters are examples of the types of parameters in the Data Pool database that the full-capability operator can modify:

- **ActionQueueCleanupFrequency -** frequency in seconds when the action queue is checked for completed actions and those older than the configured retention period are removed.
- **BatchSummaryAutoRefresh –** autorefresh rate for the **Batch Summary** page.
- **CompressOnInsert -** turns system-level compression **ON** or **OFF**.
- **DefaultRetentionPeriod -** default retention period in days for all Data Pool Insert Actions.
- **DefaultRetentionPriority -** default retention priority for all Data Pool Inserts actions. The range of valid values is 1 – 255.

- **DeleteCompletedActionsAfter -** time in minutes that operators let completed actions stay in the insert action queue before making them eligible for removal. The delay is intended to provide the operator with some ability to check on past actions. The time period should not be too long.
- **DisplayAIPChunkSize -** number of rows to return per chunk for the Active Insert Processes list.
- **HEGCleanupAge** – HDF-EOS to GeoTIF Converter (HEG) cleanup age in days.
- **IdleSleep -** number of seconds to sleep when there is nothing to do.
- **InCacheTimeLimit -** maximum time in minutes that operators are willing to wait for a Data Pool Insert Utility (DPIU) process to complete when its files are in cache. When the time limit is reached, the Data Pool Action Driver (DPAD) kills the process and retries the action.
- **InsertRetryWait -** number of seconds to wait before an insert that failed should be resubmitted (if it can be retried).
- **MAX_READ_DRIVES_*x*0*xx*g*nn*** – (multiple parameters as necessary) maximum number of simultaneous tape drives used for the specified archive.
- **MFSOnInsert** – specifies whether or not (**YES** or **NO**) DPAD should use the Multiple File System table.
- **MaxInsertRetries -** maximum number of times an insert should be tried again (-1 means forever).
- **MaxReadDrivesPerArchive -** maximum number of tape drives in use simultaneously.
- **MaxTapeMountPerRequest -** maximum number of tape mounts allowed per request.
- **NewActionCheckFrequency** – number of seconds before checking for new actions. DPAD always checks to determine whether we are out of actions that can be dispatched, so unless getting things queued up in memory is urgent, this could be a time interval of minutes.
- **NumOfAllowedCacheProcesses -** maximum number of insert processes that require access to cache.
- **NumOfAllowedInsertProcesses -** maximum number of insert processes running at any time.
- **NumOfAllowedNonCacheProcesses -** maximum number of insert processes that require access to tape.

- **OnTapeTimeLimit -** maximum time in hours operators are willing to wait for a DPIU process to complete when its files are not in cache. After the time limit, DPAD kills the process and retries the action.
- **OrderOnlyFSLabel -** order-only file system label.
- **RefreshRate - DPM Home Page** refresh rate in seconds.
- **RunAwayCheckFrequency** – number of seconds before checking again for runaway processes. It is recommended that **RunAwayCheckFrequency** not be much less than **InCacheTimeLimit**.
- **RunawayDuration -** maximum number of seconds to wait for an insert to complete before considering it a runaway.
- **SizeOfInsertQueueList -** number of Data Pool Insert Queue entries that the **DPM GUI** can display on a page at any one time.
- **StartUpWait -** number of seconds to delay start-up while trying to clean out leftover DPIU processes.
- **Aging Step -** aging rate by which the effective priority of a request increases for every hour it has been waiting.
- **Max Priority -** maximum priority a request can attain through the aging process.

A previous section of this lesson (**Monitoring Data Pool Active Insert Processes and Insert Actions**) addressed changing the **Screen Refresh Rate** parameter using an entry field on the **Home Page**. This parameter may also be changed using an entry field on the **Manage Configuration Parameters** page. In another section of the lesson (**Enable/Disable Data Compression Using the DPM GUI**) the procedure for changing the **CompressOnInsert** parameter was described.

Although most of the parameters managed on the **Manage Configuration Parameters** page are not likely to be changed frequently, the operator may want to change some of them for tuning the Data Pool. Data Pool tuning parameters can be used to help meter the flow of data into the Data Pool and to adjust retention priority and duration to maintain optimum usage of Data Pool storage. To determine the best settings, it is necessary to monitor Data Pool inserts and disk space and adjust the parameters based on experience and projected functioning.

There are four procedures in this section; i.e., **View DPM Configuration Parameter Values**, **Modify DPM Configuration Parameter Values**, **View DPM Aging Parameter Values** and **Modify DPM Aging Parameter Values**. Both full-capability operators and limited-capability operators can view DPM configuration parameter values. Full-capability operators only are allowed to modify DPM configuration parameter values.

View DPM Configuration Parameter Values

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**, **Compression Algorithms**, **Cloud Cover**, **List Insert Queue**, **Batch Summary**, **Collection Groups**, **Themes, Configuration Parameters, Aging Parameters, and End Session**).

2 Click on the **Configuration Parameters** link.

- The **List of Configuration Parameters** page is displayed.

3 Observe data displayed on the **List of Configuration Parameters** page.

- The table on the **List of Configuration Parameters** page has columns containing the following types of Data Pool configuration information:
  - **Parameter Name.**
  - **Parameter Value** (including an entry field with current value, followed by a brief description of the parameter).
  - **Click on Box to Modify Parm** (containing a check box to mark the parameter for change).
- The rows in the table indicate the current values and descriptions of the following types of parameters:
  - **ActionQueueCleanupFrequency.**
  - **BatchSummaryAutoRefresh.**
  - **CartMaxQueueSize.**
  - **CompressOnInsert.**
  - **DefaultRetentionPeriod.**
  - **DefaultRetentionPriority.**
  - **DeleteCompletedActionsAfter.**
  - **DisplayAIPChunkSize.**
  - **HEGCleanupAge.**
  - **HegActive.**

- **HegMaxProcesses.**
- **IdleSleep.**
- **InCacheTimeLimit.**
- **InsertRetryWait.**
- **MAX_READ_DRIVES_*x*0*xx*g*nn*.**
- **MFSOnInsert.**
- **MaxInsertRetries.**
- **MaxReadDrivesPerArchive.**
- **MaxTapeMountPerRequest.**
- **NewActionCheckFrequency.**
- **NumOfAllowedCacheProcesses.**
- **NumOfAllowedInsertProcesses.**
- **NumOfAllowedNonCacheProcesses.**
- **OnTapeTimeLimit.**
- **RefreshRate.**
- **RunAwayCheckFrequency.**
- **RunawayDuration.**
- **SizeOfInsertQueueList.**
- **StartUpWait.**

- There is an **Apply Change** button at the bottom of the page for implementing changes.

Modify DPM Configuration Parameter Values

---

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**, **Compression Algorithms**, **Cloud Cover**, **List Insert Queue**, **Batch Summary**, **Collection**

Groups, **Themes, Configuration Parameters, Aging Parameters, and End Session**).

2 Click on the **Configuration Parameters** link.

- The **List of Configuration Parameters** page is displayed, providing a table of DPM configuration parameters showing three columns: **Parameter Name**, **Parameter Value** (including an entry field with current value, followed by a brief description of the parameter), and **Click on Box to Modify Parm** (containing a check box to mark the parameter for change).
- There is an **Apply Change** button at the bottom of the page for implementing changes.

3 If there is an option list for the parameter value to be changed, first click on the corresponding option button.

- Options are displayed (e.g., **ON** and **OFF**).

4 If there is an option list for the parameter value to be changed, click on the appropriate choice (e.g., **ON**).

5 If there is no option list for the parameter value to be changed, type the desired value in the corresponding text entry box.

6 Click in the check box at the end of the row containing the parameter value to be modified.

- The selected configuration information is marked for modification.

7 Repeat Steps 3 through 6 for any additional parameter values to be modified.

8 To implement the modification of parameter value(s) click on the **Apply Change** button.

- The **List of Configuration Parameters** page is refreshed, the check box(es) is (are) unfilled, and the displayed **Parameter Value**(s) reflect(s) the change(s) implemented.

View DPM Aging Parameter Values

---

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**, **Compression Algorithms**, **Cloud Cover**, **List Insert Queue**, **Batch Summary**, **Collection Groups**, **Themes, Configuration Parameters, Aging Parameters, and End Session**).

2 Click on the **Aging Parameters** link.

- The **List of Aging Parameters** page is displayed.

3 Observe data displayed on the **List of Aging Parameters** page.

- The table on the **List of Aging Parameters** page has columns containing the following types of Data Pool configuration information:
  - **ECS Priority** (list of all ECS priorities).
  - **Starting Priority** (cannot be changed using the GUI).
  - **Aging Step** (includes an entry field with current value).
  - **Max Priority** (includes an entry field with current value).
  - **Click on Box to Modify Parm** (containing a check box to mark the parameter for change).
- The rows in the table indicate the current values and descriptions of the various ECS priorities, from LOW to EXPRESS.
- There is an **Apply Change** button at the bottom of the page for implementing changes.

Modify DPM Aging Parameter Values

---

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**, **Compression Algorithms**, **Cloud Cover**, **List Insert Queue**, **Batch Summary**, **Collection Groups**, **Themes, Configuration Parameters, Aging Parameters, and End Session**).

2 Click on the **Aging Parameters** link.

- The **List of Aging Parameters** page is displayed, providing a table of DPM aging parameters showing five columns: **ECS Priority**, **Starting Priority**, **Aging Step** (including an entry field with current value), **Max Priority** (including an entry field with current value), and **Click on Box to Modify Parm** (containing a check box to mark the parameter for change).
- There is an **Apply Change** button at the bottom of the page for implementing changes.

3 To change the value associated with **Aging Step** and/or **Max Priority** for a particular ECS priority first type the desired value(s) in the corresponding text entry box(s).

4 To continue the process of changing the value associated with **Aging Step** and/or **Max Priority** for a particular ECS priority click in the check box at the end of the row containing the parameter value(s) to be modified.

- The selected configuration information is marked for modification.

5 Repeat Steps 3 and 4 for any additional parameter values to be modified.

6 To implement the modification of parameter value(s) click on the **Apply Change** button.

- The **List of Aging Parameters** page is refreshed, the check box(es) is (are) unfilled, and the displayed **Aging Step** and **Max Priority** values reflect the change(s) implemented.

Number of Drivers to Run

It is recommended to start with three event drivers, three action drivers, one recovery driver, and one deletion driver, although larger DAACs may eventually need more event drivers and action drivers running. The numbers of drivers are determined by parameters specified in the command to execute the start script for NBSRV drivers. Although one recovery driver and one deletion driver will generally be sufficient, it may be desirable to run more event drivers and action drivers. For example, if increased throughput is needed, it may be appropriate to double the number of action drivers and event drivers to six each. The start script is **EcNbDriverStart**; it is located with other start scripts in the directory **/usr/ecs/<*MODE*>/CUSTOM/utilities**. To start NBSRV drivers, the command is:

EcNbDriverStart <*MODE*> d_e d_a d_r d_d

where **<*MODE*>** is the mode in which Data Pool is being run, d_e is the number of event drivers, d_a is the number of action drivers, d_r is the number of recovery drivers, and d_d is the number of deletion drivers. For example, to start the drivers in the OPS mode with the beginning recommendation, use the command **EcNbDriverStart OPS 3 3 1 1**.

To add drivers, the script can simply be run again. For example, to increase the number of event and action drivers to six each, execute the command **EcNbDriverStart OPS 3 3 0 0**.

A database query can give you an idea of the concurrency of processing by event and action drivers; i.e., how well action processing is keeping up with event processing. You can use **isql** and select **max(actionDateTime** from **EcNbActionQueueLog** where **actionStatus = 'Acquire'** or **actionStatus = 'ActionNotification'**. If the result is more than an hour earlier than the current time, this is an indication that action processing may be lagging behind event processing. In that event, it may be appropriate to try increasing the number of action drivers to one and one-half times the number of event drivers (e.g., six event drivers and nine action drivers). This can

result in greater concurrency in processing, but an increase in the number of waiting processes and/or deadlocks occurring on the Sybase server may impose a practical limit on the number of drivers that it is productive to run. Each site has to monitor Data Pool performance and make adjustments based on experience as necessary to achieve optimum performance.

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 Groups** 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.

View Collection Group and Collection Information Using the DPM GUI

---

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**, **Compression Algorithms**, **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.

3 Observe data displayed on the **Collection Groups** page.

- The table on the **Collection Groups** page has columns containing the following types of collection group information:
  - **Data Source** (i.e., ECS or NON-ECS).
  - **Group ID.**

- **Display Name.**
- **Description.**

- The following links are available on the **Collection Groups** page:
  - Each collection listed in the **Group ID** column links to a **Collection Group Detail** page.
  - **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** column.

- The **Collection Group Detail** page is displayed.

5 Observe data displayed on the **Collection Group Detail** page.

- Near the top of the **Collection Group Detail** page is the following basic collection group information:
  - **Data Source** (i.e., ECS or NON-ECS).
  - **Group ID.**
  - **Display Name.**
  - **Description.**
- There is a file system filter (and associated **Apply Filter** button) for displaying data on the **Collection Group Detail** page for all file systems or by individual file system.
- The table on the **Collection Group Detail** page has columns containing the following types of collection group information:
  - **Collection**.
  - **Version.**
  - **Compression Command Label.**
  - **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 **Collection Group Detail** 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 **Collection Group Detail** page first click on the **File System** filter option button.

- Options are displayed.

7 To 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 Group Detail** page click on the **Apply Filter** button.

- The **Collection Group Detail** page is displayed with the filtered collection group information.

9 If data displayed on the **Collection Group Detail** page were filtered, observe data displayed on the **Collection Group Detail** page.

- Refer to Step 5.

10 To obtain more information about one of the collections in the collection group, click on its link in the **Collection** column.

- The **Collection Detail** page is displayed.

11 Observe data displayed on the **Collection Detail** page.

- Near the top of the **Collection Detail** page is the following basic collection group information:
  - **Data Source** (i.e., ECS or NON-ECS).
  - **Group ID.**
  - **Display Name.**
  - **Description.**
- The table on the **Collection Detail** page has rows containing the following types of collection information:
  - **Collection**.

- **Version.**
- **Description.**
- **File System.**
- **Compression Command Label.**
- **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 Detail** page:
  - **Modify Collection.**
  - **Return to previous page.**

12 To view a description for another collection in the same group first click on the **Return to previous page** link.

- The **Collection Group Detail** page is displayed again.

13 To view a description for another collection in the same group return to Step 10.

14 To view a description for another collection in another group return to Step 2.

Rarely, 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:

Modify Collection Groups Using the DPM GUI

---

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**, **Compression Algorithms**, **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., **Source Type**, **Group ID**, **Display Name**, and **Description**.
- The following links are available: **Add Collection Group**, **Modify Collection Group**, and each collection listed in the **Group ID** column links to a **Collection Group Detail** page.

3 Click on the **Modify Collection Group** link at the bottom of the page.

- The **Modify Collection Group** page is displayed, providing a table of collection group information showing five columns: **Data Source**, **Group ID**, **Display Name**, **Description**, and **Check box to modify** (containing a check box to mark the collection group for change).
- 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 Groups** 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.

Add a Collection Group Using the DPM GUI

---

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**, **Compression Algorithms**, **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., **Source Type**, **Group ID**, **Display Name**, and **Description**.

- The following links are available: **Add Collection Group**, **Modify Collection Group**, and each collection listed in the **Group ID** column links to a **Collection Group Detail** page.

3 Click on the **Add Collection Group** link at the bottom of the page.

- The screen displays a page with four columns of text-entry fields: **Data Source**, **Group ID**, **Display Name**, and **Description**.

4 To display data source options click on the **Data Source** option button.

- **Data Source** options are displayed (i.e., **ECS** and **NON-ECS**).

5 To select a data source option click on the appropriate choice from the option list.

6 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.

7 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.

8 Type the description for the new collection group in the **Description** field.

- The **Description** may have no more than 255 characters.

9 Click on the **Apply Change** button.

- The new collection group information is entered in the Data Pool database.
- The **Collection Groups** page is displayed with the new collection group information.

Managing Data Pool Collections within Collection Groups

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. The procedure for adding ECS collections to a collection group is somewhat different from the procedure for adding a non-ECS

collection to a collection group. Full-capability operators (only) can use the following procedure to add an ECS collection to an existing collection group:

Add an ECS Collection to a Collection Group Using the DPM GUI

---

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**, **Compression Algorithms**, **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., **Source Type**, **Group ID**, **Display Name**, and **Description**.
- The following links are available: **Add Collection Group**, **Modify Collection Group**, and each collection listed in the **Group ID** column links to a **Collection Group Detail** page.

3 Click on the **Group ID** link for the ECS collection group to which the collection is to be added.

- The **Collection Group Detail (List of Collections)** page is displayed with the following basic collection group information near the top of the page: **Data Source** (i.e., ECS), **Group ID**, **Display Name**, and **Description.**
- There is a file system filter (and associated **Apply Filter** button) for displaying data on the **Collection Group Detail** page for all file systems or by individual file system.
- The table on the **Collection Group Detail** page has 13 columns containing the following types of collection group information: **Collection**, **Version**, **Compression Command Label**, **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**, and **Cloud Coverage.**
- The following links are available: **Add New Collection**, **Return to previous page**, and each collection listed in the **Collection** column links to a **Collection Detail** page.

4 Click on the **Add New Collection** link at the bottom of the **Collection Group Detail (List of Collections)** page.

- The **Collections Not in Data Pool** page is displayed with the following basic collection group information near the top of the page: **Data Source** (i.e., ECS), **Group ID**, **Display Name**, and **Description.**

- The table on the **Collections Not in Data Pool** page has three columns containing the following types of collection group information: **Collection**, **Version**, and **Description**.
- The following links are available: **Return to previous page** and each collection listed in the **Collection** column links to a **Collection Detail** page.

5 Click on the link (in the **Collection** column) of the collection to be added to the collection group.

- The **Add New Collection** page is displayed with the following basic collection group information near the top of the page: **Data Source** (i.e., ECS), **Group ID**, **Display Name**, and **Description.**
- The **Add New Collection** page has a table of collection information showing 13 rows: **Collection**, **Version**, **Description**, **File System**, **Compression Command Label**, **Science Granules and/or Metadata**, **Data Pool Insertion**, **Quality Summary Url**, **Spatial Search Type**, **Global Coverage**, **Day/Night Coverage**, **24 Hour Coverage**, and **Cloud Cover Type & Source**.
- There is an **Apply Change** button at the bottom of the page to implement the new collection in the collection group.

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 display file system options (if applicable) click on the **File System** option button.

- **File System** options are displayed (if there are multiple Data Pool file systems).

7 To select a file system option (if applicable) click on the appropriate choice from the **File System** option list.

8 To display compression command label options (if applicable) click on the **Compression Command Label** option button.

- **Compression Command Label** options are displayed.
- Selection of a compression command label is not required.

9 To select a compression command label option (if applicable) click on the appropriate choice from the **Compression Command Label** option list.

10 To display science granules and/or metadata options click on the **Science Granules and/or Metadata** option button.

- **Science Granules and/or Metadata** options (i.e., **Science and Metadata** and **Metadata Only**) are displayed.

11 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.

12 To display data pool insertion options click on the **Data Pool Insertion** option button.

- **Data Pool Insertion** options (i.e., **Invalid for Data Pool** and **Valid for Data Pool**) are displayed.

13 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.

14 To display ECHO export options click on the **Export Urls to ECHO** option button.

- ECHO export options (i.e., **Yes** and **No**) are displayed.

15 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.

16 If the collection is to be linked to a quality summary web site, enter the URL in the **Quality Summary** text entry field.

- Ensure that **http://** is included in the **Quality Summary** text entry field.

17 To display global coverage options click on the **Global Coverage** option button.

- **Global Coverage** options are displayed.

18 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.

19 To display day/night coverage options click on the **Day/Night Coverage** option button.

- **Day/Night Coverage** options are displayed.

20 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.

21 To display 24-hour coverage options click on the **24 Hour Coverage** option button.

- **24 Hour Coverage** options are displayed.

22 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.

23 To display cloud cover type and source options click on the **Cloud Cover Type & Source** option button.

- **Cloud Cover Type & Source** options are displayed.

24 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).

25 To view details of cloud cover type and source click on the **View Details** link adjacent to the **Cloud Cover Type & Source** option list.

26 Click on the **Apply Change** button.

- The new collection information is entered in the Data Pool database.
- The **Collection Group Detail** page is displayed with the new collection information.

Full-capability operators (only) can use the following procedure to add a non-ECS collection to an existing collection group:

Add a NON-ECS Collection to a Collection Group Using the DPM GUI

---

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**, **Compression Algorithms**, **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., **Source Type**, **Group ID**, **Display Name**, and **Description**.
- The following links are available: **Add Collection Group**, **Modify Collection Group**, and each collection listed in the **Group ID** column links to a **Collection Group Detail** page.

3 Click on the **Group ID** link for the non-ECS collection group to which the collection is to be added.

- The **Collection Group Detail (List of Collections)** page is displayed with the following basic collection group information near the top of the page: **Data Source** (i.e., NON-ECS), **Group ID**, **Display Name**, and **Description.**
- There is a file system filter (and associated **Apply Filter** button) for displaying data on the **Collection Group Detail** page for all file systems or by individual file system.
- The table on the **Collection Group Detail** page has 13 columns containing the following types of collection group information: **Collection**, **Version**, **Compression Command Label**, **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**, and **Cloud Coverage.**
- The following links are available: **Add New Collection**, **Return to previous page**, and each collection listed in the **Collection** column links to a **Collection Detail** page.

4 Click on the **Add New Collection** link at the bottom of the **Collection Group Detail (List of Collections)** page.

- The **Add New Collection** page is displayed with the following basic collection group information near the top of the page: **Data Source** (i.e., NON-ECS), **Group ID**, **Display Name**, and **Description.**
- The **Add New Collection** page has a table of collection information showing 13 rows: **Collection**, **Version**, **Description**, **File System**, **Compression Command Label**, **Science Granules and/or Metadata**, **Data Pool Insertion**, **Quality Summary Url**, **Spatial Search Type**, **Global Coverage**, **Day/Night Coverage**, **24 Hour Coverage**, and **Cloud Cover Type & Source**.
- There is an **Apply Change** button at the bottom of the page to implement the new collection in the collection group.

5 Type a name for the new collection in the **Collection** text entry field.

- The name for the new collection may have no more than eight characters.
  - Valid characters include A-Z, 0-9, and underscore.
- The name must start with a letter.

- The name will be compared with the existing collection names to ensure that it is not a duplicate of another one.

6 Type a version number for the new collection in the **Version** text entry field.

7 Type a description for the new collection in the **Description** text entry field.

- The description for the new collection may have no more than 255 characters.

8 To display file system options (if applicable) click on the **File System** option button.

- **File System** options are displayed (if there are multiple Data Pool file systems).

9 To select a file system option (if applicable) click on the appropriate choice from the **File System** option list.

10 To display compression command label options click on the **Compression Command Label** option button.

- **Compression Command Label** options are displayed.
- Selection of a compression command label is not required.

11 To select a compression command label option click on the appropriate choice from the **Compression Command Label** option list.

12 To display science granules and/or metadata options click on the **Science Granules and/or Metadata** option button.

- **Science Granules and/or Metadata** options (i.e., **Science and Metadata** and **Metadata Only**) are displayed.

13 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.

14 To display data pool insertion options click on the **Data Pool Insertion** option button.

- **Data Pool Insertion** options (i.e., **Invalid for Data Pool** and **Valid for Data Pool**) are displayed.

15 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.

16 To display ECHO export options click on the **Export Urls to ECHO** option button.

- ECHO export options (i.e., **Yes** and **No**) are displayed.

17 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.

18 If the collection is to be linked to a quality summary web site, enter the URL in the **Quality Summary** text entry field.

- Ensure that **http://** is included in the **Quality Summary** text entry field.

19 To display spatial search type options click on the **Spatial Search Type** option button.

- **Spatial Search Type** options (e.g., **Not Supported**, **Orbit**, **Rectangle**, and **GPolygon**) are displayed.

20 To select a spatial search type option click on the appropriate choice from the **Spatial Search Type** option list.

- **Not Supported** is the default spatial search type.

21 To display global coverage options click on the **Global Coverage** option button.

- **Global Coverage** options are displayed.

22 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.

23 To display day/night coverage options click on the **Day/Night Coverage** option button.

- **Day/Night Coverage** options are displayed.

24 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.

25 To display 24-hour coverage options click on the **24 Hour Coverage** option button.

- **24 Hour Coverage** options are displayed.

26 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.

27 To display cloud cover type and source options click on the **Cloud Cover Type and Source** option button.

- **Cloud Cover Type & Source** options are displayed.

28 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).

29 To view details of cloud cover type and source click on the **View Details** link adjacent to the **Cloud Cover Type & Source** option list.

30 Click on the **Apply Change** button.

- The new collection information is entered in the Data Pool database.
- The **Collection Group Detail** 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.

The procedure for modifying an ECS collection is somewhat different from the procedure for modifying a non-ECS collection. Full-capability operators (only) can use the following procedure to modify an ECS collection:

Modify an ECS Collection Using the DPM GUI

---

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**, **Compression Algorithms**, **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., **Source Type**, **Group ID**, **Display Name**, and **Description**.
- The following links are available: **Add Collection Group**, **Modify Collection Group**, and each collection listed in the **Group ID** column links to a **Collection Group Detail** page.

3 Click on the **Group ID** link for the collection group containing the collection to be modified.

- The **Collection Group Detail (List of Collections)** page is displayed with the following basic collection group information near the top of the page: **Data Source** (i.e., ECS), **Group ID**, **Display Name**, and **Description.**
- There is a file system filter (and associated **Apply Filter** button) for displaying data on the **Collection Group Detail** page for all file systems or by individual file system.
- The table on the **Collection Group Detail** page has 13 columns containing the following types of collection group information: **Collection**, **Version**, **Compression Command Label**, **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**, and **Cloud Coverage.**
- The following links are available: **Add New Collection**, **Return to previous page**, and each collection listed in the **Collection** column links to a **Collection Detail** page.

4 Click on the link (in the **Collection** column) of the collection to be modified.

- The **Collection Detail** page is displayed with the following basic collection group information near the top of the page: **Data Source** (i.e., ECS), **Group ID**, **Display Name**, and **Description.**
- The **Collection Detail** page has a table of collection information showing 17 rows: **Collection**, **Version**, **Description**, **File System**, **Compression Command Label**, **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**, and **Cloud Cover Description**.
- There is a **Modify Collection** link and a **Return to previous page** link at the bottom of the page.

5 Click on the **Modify Collection** link.

- The **Modify Collection** page is displayed with the following basic collection group information near the top of the page: **Data Source** (i.e., ECS), **Group ID**, **Display Name**, and **Description.**

- The **Modify Collection** page has a table of collection information showing 15 rows: **Collection**, **Version**, **Description**, **File System**, **Compression Command Label**, **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**, and **Cloud Cover Type & Source**.
- There is an **Apply Change** button at the bottom of the page to implement the new collection in the collection group.

NOTE: On the ECS collection version of the **Modify Collection** page the **Collection**, **Version**, **Description**, and **Spatial Search Type** fields cannot be edited.

6 To display file system options (if applicable) click on the **File System** option button.

- **File System** options are displayed (if there are multiple Data Pool file systems).

7 To select a file system option (if applicable) click on the appropriate choice from the **File System** option list.

8 To display compression command label options (if applicable) click on the **Compression Command Label** option button.

- **Compression Command Label** options are displayed.

9 To select a compression command label option (if applicable) click on the appropriate choice from the **Compression Command Label** option list.

10 To display science granules and/or metadata options (if applicable) click on the **Science Granules and/or Metadata** option button.

- **Science Granules and/or Metadata** options (i.e., **Science and Metadata** and **Metadata Only**) are displayed.

11 To select a science granules and/or metadata option (if applicable) click on the appropriate choice from the **Science Granules and/or Metadata** option list.

- **Science and Metadata** is the default option.

12 To display data pool insertion options (if applicable) click on the **Data Pool Insertion** option button.

- **Data Pool Insertion** options (i.e., **Invalid for Data Pool** and **Valid for Data Pool**) are displayed.

13 To select a data pool insertion option (if applicable) click on the appropriate choice from the **Data Pool Insertion** option list.

- **Valid for Data Pool** must be selected if the collection is to be eligible for insertion into the Data Pool.

14 To display ECHO export options click on the **Export Urls to ECHO** option button.

- ECHO export options (i.e., **Yes** and **No**) are displayed.

15 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.

16 If the collection is to be linked to a quality summary web site, enter the URL in the **Quality Summary** text entry field.

- Ensure that **http://** is included in the **Quality Summary** text entry field.

17 To display global coverage options (if applicable) click on the **Global Coverage** option button.

- **Global Coverage** options are displayed.

18 To select a global coverage option (if applicable) 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.

19 To display day/night coverage options (if applicable) click on the **Day/Night Coverage** option button.

- **Day/Night Coverage** options are displayed.

20 To select a day/night coverage option (if applicable) 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.

21 To display 24-hour coverage options (if applicable) click on the **24 Hour Coverage** option button.

- **24 Hour Coverage** options are displayed.

22 To select a 24-hour coverage option (if applicable) 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.

23 To display cloud cover type and source options (if applicable) click on the **Cloud Cover Type & Source** option button.

- **Cloud Cover Type & Source** options are displayed.

24 To select a cloud cover type and source option (if applicable) 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).

25 To view details of cloud cover type and source (if applicable) click on the **View Details** link adjacent to the **Cloud Cover Type & Source** option list.

26 Click on the **Apply Change** button.

- The modified collection information is entered in the Data Pool database.
- The **Collection Group Detail** page is displayed with the modified collection information.

Full-capability operators (only) can use the following procedure to modify a non-ECS collection:

Modify a NON-ECS Collection Using the DPM GUI

---

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**, **Compression Algorithms**, **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., **Source Type**, **Group ID**, **Display Name**, and **Description**.
- The following links are available: **Add Collection Group**, **Modify Collection Group**, and each collection listed in the **Group ID** column links to a **Collection Group Detail** page.

3 Click on the **Group ID** link for the collection group containing the collection to be modified.

- The **Collection Group Detail (List of Collections)** page is displayed with the following basic collection group information near the top of the page: **Data Source** (i.e., NON-ECS), **Group ID**, **Display Name**, and **Description.**
- There is a file system filter (and associated **Apply Filter** button) for displaying data on the **Collection Group Detail** page for all file systems or by individual file system.
- The table on the **Collection Group Detail** page has 13 columns containing the following types of collection group information: **Collection**, **Version**, **Compression Command Label**, **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**, and **Cloud Coverage.**
- The following links are available: **Add New Collection**, **Return to previous page**, and each collection listed in the **Collection** column links to a **Collection Detail** page.

4 Click on the link (in the **Collection** column) of the collection to be modified.

- The **Collection Detail** page is displayed with the following basic collection group information near the top of the page: **Data Source** (i.e., NON-ECS), **Group ID**, **Display Name**, and **Description.**
- The **Collection Detail** page has a table of collection information showing 17 rows: **Collection**, **Version**, **Description**, **File System**, **Compression Command Label**, **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**, and **Cloud Cover Description**.
- There is a **Modify Collection** link and a **Return to previous page** link at the bottom of the page.

5 Click on the **Modify Collection** link.

- The **Modify Collection** page is displayed with the following basic collection group information near the top of the page: **Data Source** (i.e., NON-ECS), **Group ID**, **Display Name**, and **Description.**
- The **Modify Collection** page has a table of collection information showing 15 rows: **Collection**, **Version**, **Description**, **File System**, **Compression Command Label**, **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**, and **Cloud Cover Type & Source**.
- There is an **Apply Change** button at the bottom of the page to implement the new collection in the collection group.

NOTE: On the NON-ECS collection version of the **Modify Collection** page the **Collection** and **Version** fields cannot be edited.

6 If applicable, type a description for the collection in the **Description** text entry field.

- The description for the collection may have no more than 255 characters.

7 To display file system options (if applicable) click on the **File System** option button.

- **File System** options are displayed (if there are multiple Data Pool file systems).

8 To select a file system option (if applicable) click on the appropriate choice from the **File System** option list.

9 To display compression command label options (if applicable) click on the **Compression Command Label** option button.

- **Compression Command Label** options are displayed.

10 To select a compression command label option (if applicable) click on the appropriate choice from the **Compression Command Label** option list.

11 To display science granules and/or metadata options (if applicable) click on the **Science Granules and/or Metadata** option button.

- **Science Granules and/or Metadata** options (i.e., **Science and Metadata** and **Metadata Only**) are displayed.

12 To select a science granules and/or metadata option (if applicable) click on the appropriate choice from the **Science Granules and/or Metadata** option list.

- **Science and Metadata** is the default option.

13 To display data pool insertion options (if applicable) click on the **Data Pool Insertion** option button.

- **Data Pool Insertion** options (i.e., **Invalid for Data Pool** and **Valid for Data Pool**) are displayed.

14 To select a data pool insertion option (if applicable) click on the appropriate choice from the **Data Pool Insertion** option list.

- **Valid for Data Pool** must be selected if the collection is to be eligible for insertion into the Data Pool.

15 To display ECHO export options click on the **Export Urls to ECHO** option button.

- ECHO export options (i.e., **Yes** and **No**) are displayed.

16 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.

17 If the collection is to be linked to a quality summary web site, enter the URL in the **Quality Summary** text entry field.

- Ensure that **http://** is included in the **Quality Summary** text entry field.

18 To display spatial search type options (if applicable) click on the **Spatial Search Type** option button.

- **Spatial Search Type** options (e.g., **Not Supported**, **Orbit**, **Rectangle**, and **GPolygon**) are displayed.
- **Spatial Search Type** can be changed only when the collection is not currently enabled for insert and the Data Pool contains no granules belonging to the collection.

19 To select a spatial search type option (if applicable) click on the appropriate choice from the **Spatial Search Type** option list.

20 To display global coverage options (if applicable) click on the **Global Coverage** option button.

- **Global Coverage** options are displayed.

21 To select a global coverage option (if applicable) 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.

22 To display day/night coverage options (if applicable) click on the **Day/Night Coverage** option button.

- **Day/Night Coverage** options are displayed.

23 To select a day/night coverage option (if applicable) 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.

24 To display 24-hour coverage options (if applicable) click on the **24 Hour Coverage** option button.

- **24 Hour Coverage** options are displayed.

25 To select a 24-hour coverage option (if applicable) 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.

26 To display cloud cover type and source options (if applicable) click on the **Cloud Cover Type and Source** option button.

- **Cloud Cover Type & Source** options are displayed.

27 To select a cloud cover type and source option (if applicable) 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).

28 To view details of cloud cover type and source (if applicable) click on the **View Details** link adjacent to the **Cloud Cover Type & Source** option list.

29 Click on the **Apply Change** button.

- The modified collection information is entered in the Data Pool database.
- The **Collection Group Detail** page is displayed with the modified collection information.

Managing Themes Using the DPM GUI

Users may search the Data Pool for data associated with themes. As data are inserted into the Data Pool, it is possible to associate the data with themes. The **DPM GUI Detailed List of Data Pool Themes** page permits both full-capability and limited-capability operators users to view a list of Data Pool themes. In addition it has links that allow full-capability operators to add new themes, modify existing themes, or delete themes.

Beginning with Synergy V the Deployment of Open Geospatial Consortium (OGC) Web Services (DOWS) allows OGC-compliant clients the ability to search, map, download, and subset (spatially and temporally) data residing in the EOS Data Pools. Consequently, a configurable subset of the EOS Data Pools' inventory can be accessed via the following OGC Web Services (OWS):

- Web Coverage Service (WCS).
- Web Mapping Service (WMS).

OWS (WCS/WMS) access to data in the Data Pool is established by defining WCS/WMS-enabled themes using the DPM GUI. Data are put in the Data Pool to populate the WCS/WMS-enabled themes via Spatial Subscription Server subscriptions for Data Pool inserts. In addition it is possible to enable "preconversion;" i.e., conversion of each granule data file from HDF-EOS to GeoTIFF format before the granule is inserted into the Data Pool (rather than leaving conversion until an end user tries to download the data from the Data Pool).

[Additional information concerning OWS is available in the section on the **Open Geospatial Consortium (OGC) Web Services (OWS)** (subsequent section of this lesson).]

View a List of Themes Using the DPM GUI

---

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**, **Compression Algorithms**, **Cloud Cover**, **List Insert Queue**, **Batch Summary**, **Collection Groups**, **Themes**, **Configuration Parameters**, **Aging Parameters**, and **End Session**).

2 Click on the **Themes** link.

- The **Detailed List of Data Pool Themes** page is displayed.

3 Observe data displayed on the **Detailed List of Data Pool Themes** page.

- The table on the **Detailed List of Data Pool Themes** page has columns containing the following types of Data Pool file system information:
  - **Theme Name/Description.**
  - **Web** [Visible].
  - **Insert** [Enabled].
  - **WCS** [Enabled].
  - **WMS** [Enabled].
  - **PreConvert** [Enabled].
  - **Click on Box to Delete** (containing a check box to mark the theme for deletion)**.**
- There are theme filters (and associated **Apply Filter** button) for displaying data on the **Detailed List of Data Pool Themes** page depending on whether or not the themes…
  - Are web visible.
  - Are insert enabled.
  - Are WCS enabled.
  - Are WMS enabled.
  - Are preconversion enabled.

  - Have certain letters at the beginning of the theme name.
- Filters can be applied individually or in any combination.
- The following links are available on the **Detailed List of Data Pool Themes** page:
  - **Add New Theme.**
  - **Modify Theme.**
- There is an **Apply Change** button at the bottom of the page to implement the deletion of selected themes.

4 To filter data displayed on the **Detailed List of Data Pool Themes** page use the **Filter a List of Themes** procedure (subsequent section of this lesson).

5 If data displayed on the **Detailed List of Data Pool Themes** page were filtered, return to Step 3.

The procedure to **Filter a List of Themes** is subordinate to other theme-related procedures (i.e., **View a List of Themes Using the DPM GUI**, **Modify a Theme Using the DPM GUI**, and **Delete a Theme Using the DPM GUI**). Both full-capability and limited-capability operators users may filter data displayed on the Themes pages to which they have access.

Filter a List of Themes

---

1 To filter data displayed on one of the **Themes** pages on the basis of whether or not the themes are enabled for web drill-down, insertion in to the Data Pool, WCS accessibility, WMS accessibility, file format preconversion first click on one of the option buttons (in the filter area of the page):

- **Web Visible.**
- **Insert Enabled.**
- **WCS.**
- **WMS.**
- **PreConvert.**

2 To continue the process of selecting a theme filter option click on the appropriate choice from the option list.

- **Yes** (View all themes enabled for the selected option).
- **No** (View all themes disabled for the selected option).
- **ALL** (View all themes regardless of whether the selected option is enabled or disabled).

3 Repeat Steps 1 and 2 as necessary to select additional filter options.

4 To select a theme filter option on the basis of the beginning letters of the theme (if applicable) type the beginning letter(s) of the theme in the **Beginning Letters** text entry field.

5 To implement the filtering of data displayed on one of the **Themes** pages click on the **Apply Filter** button.

- The page is displayed with the filtered theme information.

6 Return to the procedure that specified the **Filter a List of Themes** procedure.

Full-capability operators can use the **DPM GUI** to modify a theme. This can be useful if, for example, it is noted that access frequency for granules referencing a theme has declined to the point that the thematic collection should be removed from the Data Pool, but there are a few web users that still use it. In that case, it may be appropriate to change the description of the theme to alert users that the theme will be phased out soon.

Modify a Theme Using the DPM GUI

---

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**, **Compression Algorithms**, **Cloud Cover**, **List Insert Queue**, **Batch Summary**, **Collection Groups**, **Themes, Configuration Parameters, Aging Parameters, and End Session**).

2 Click on the **Themes** link.

- The **Detailed List of Data Pool Themes** page is displayed.

3 Click on the **Modify Theme** link.

- The **Modify Theme** page is displayed, providing a table of theme information showing eight columns: **Theme Name, Description, Web Visible, Insert Enabled, WCS, WMS, PreConvert,** and **Click on Box to Modify** (containing a check box to mark the theme for change).
- There are theme filters (and associated **Apply Filter** button) for displaying data on the **Modify Theme** page.
- There is an **Apply Change** button at the bottom of the page for implementing changes.

- The following links are available: **Return to previous page** and **Return to Main Theme Page**.

4 To filter data displayed on the **Modify Theme** page use the **Filter a List of Themes** procedure (previous section of this lesson).

5 To change the description of a theme (if applicable) type the desired description in the **Description** field for the theme name.

- The **Description** may have no more than 255 characters.

6 To change the theme from enabled to disabled (or vice versa) for one of the options (i.e., **Web Visible, Insert Enabled, WCS, WMS,** or **PreConvert**) (if applicable) click on the toggle button box in the corresponding column in the row for the theme.

- A check mark in the box indicates that the theme is enabled for the corresponding option.
- The absence of a check mark in the box indicates that the theme is not enabled for the corresponding option.

7 Click in the **Click on Box to Modify** check box at the end of the row containing the theme to be modified.

- The selected theme is marked for modification.

8 Repeat Steps 5 through 7 as necessary for any additional themes to be modified.

9 To implement the modification of theme(s) click on the **Apply Change** button.

- The theme information is entered in the Data Pool database.
- The **Detailed List of Data Pool Themes** page is displayed with the modified theme information.

Full-capability operators (only) can use the following procedure to add a theme:

Add a Theme Using the DPM GUI

---

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**, **Compression Algorithms**, **Cloud Cover**, **List Insert Queue**, **Batch Summary**, **Collection Groups**, **Themes, Configuration Parameters, Aging Parameters, and End Session**).

2 Click on the **Themes** link.

- The **Detailed List of Data Pool Themes** page is displayed.

3 Click on the **Add New Theme** link.

- The **Add New Theme** page is displayed, providing a table of theme information showing four columns: **Theme Name**, **Description**, **Web Visible**, and **Insert Enabled**.
- There are theme filters (and associated **Apply Filter** button) for displaying data on the **Modify Theme** page.
  - The filters serve no real function on this page (there is nothing to filter).
- There is an **Apply Change** button at the bottom of the page for implementing changes.
- The following link is available: **Return to theme list**.

4 Type a unique name for the theme in the **Theme Name** text entry field.

- The **Theme Name** may have no more than 40 characters.
- The **Theme Name** may not start with a number.
- The **Theme Name** may not duplicate the name of a collection, an ESDT, or another theme.

5 To enter a description of the theme type the desired description in the **Description** text entry field.

- The **Description** may have no more than 255 characters.

6 To enable the theme for one of the options (i.e., **Web Visible, Insert Enabled, WCS, WMS,** or **PreConvert**) (if applicable) click on the toggle button box in the corresponding area of the form.

- A check mark in the box indicates that the theme is enabled for the corresponding option.
- The absence of a check mark in the box indicates that the theme is not enabled for the corresponding option.

7 Repeat Step 6 as necessary to enable the theme for additional options.

8 Click on the **Apply Change** button.

- The new theme information is entered in the Data Pool database.
- The **Detailed List of Data Pool Themes** page is displayed with the new theme information.

Full-capability operators (only) can use the following procedure to delete a theme:

Delete a Theme Using the DPM GUI

---

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**, **Compression Algorithms**, **Cloud Cover**, **List Insert Queue**, **Batch Summary**, **Collection Groups**, **Themes, Configuration Parameters, Aging Parameters, and End Session**).

2 Click on the **Themes** link.

- The **Detailed List of Data Pool Themes** page is displayed, with columns containing the following types of Data Pool file system information:
  - **Theme Name/Description.**
  - **Web** [Visible].
  - **Insert** [Enabled].
  - **WCS** [Enabled].
  - **WMS** [Enabled].
  - **PreConvert** [Enabled].
  - **Click on Box to Delete** (containing a check box to mark the theme for deletion)**.**
- There are theme filters (and associated **Apply Filter** button) for displaying data on the **Detailed List of Data Pool Themes** page depending on whether or not the themes…
  - Are web visible.
  - Are insert enabled.
  - Are WCS enabled.
  - Are WMS enabled.
  - Are preconversion enabled.
  - Have certain letters at the beginning of the theme name.
- Filters can be applied individually or in any combination.

- The following links are available on the **Detailed List of Data Pool Themes** page:
  - **Add New Theme.**
  - **Modify Theme.**
- There is an **Apply Change** button at the bottom of the page to implement the deletion of selected themes.

3 To filter data displayed on the **Detailed List of Data Pool Themes** page use the **Filter a List of Themes** procedure (previous section of this lesson).

4 Click in the **Click on Box to Delete** check box at the end of the row containing the theme to be deleted.

- The selected theme is marked for deletion.

5 Repeat Step 4 as necessary for any additional themes to be deleted.

6 To implement the deletion of theme(s) click on the **Apply Change** button.

- The theme deletion information is entered in the Data Pool database.
- The **Detailed List of Data Pool Themes** page is displayed with the modified theme information.

Working with Data Pool Scripts

There are several Data Pool scripts that provide the operator with utilities or applications for managing Data Pool maintenance. These include:

- Update Granule Utility: a script to update granule expiration (extend the period of retention) and, optionally, retention priority, for selected science granules already in the Data Pool. For Synergy III, the utility allows operators to extend the expiration of all granules associated with a particular thematic collection or upgrade their expiration priority.
- Data Pool Cleanup Utility: a script to remove expired granules from Data Pool disks and inventory and free up space in the Data Pool. For Synergy III, it allows operators to clean up a thematic collection. It permits unlinking granules from a thematic collection, and permits data pool removal of granules that belong to a given thematic collection if they meet the cleanup criteria specified by the other command line parameters.
- Data Pool Access Statistics Utility (DPASU): scripts for processing available Data Pool access logs to extract and summarize statistics on FTP and web access to data in the Data Pool. The statistics are stored in the Data Pool database to be used for producing tabular reports that can be loaded into a spreadsheet program for sorting, graphing, or other manipulation.

- Data Pool Archive Access Statistics Data Utility: an operational support tool for archiving the granule access data from the Data Pool database into a tab-delimited ASCII file.
- Data Pool Delete Access Statistics Data Utility: an operational support tool for deleting granule access data in a specified time range from the Data Pool database.
- Data Pool Restore Access Statistics Data Utility: an operational support tool for restoring granule access data for a specific time range from an ASCII archive file to the Data Pool database.
- Batch Insert Utility: a command-line utility for inserting non-ECS data and ECS data that are already in the archive into the Data Pool.
- Most Recent Data Pool Inserts Utility: normally runs as a cron job that lists the most recent additions to the Data Pool. If necessary, the utility can be run from the command line.
- Data Pool Collection-to-Group Remapping Utility: a command-line utility interface that is used for reassigning a Data Pool collection to a collection group other than the one to which it was originally assigned.
- Data Pool QA Update Utility: a command-line interface for updating the QA data for granules in the Data Pool inventory. Both the inventory and the corresponding XML files on disk are updated.
- Data Pool Move Collections Utility: a command-line interface to move collections from one file system to another. The file system move is implemented as a copy operation to the new collection directory, followed by removal of the old collection directory and its contents.
- Data Pool Hidden Scrambler Utility: a command-line utility for making the transition to or renaming (with encrypted names) hidden directories for order-only granules in the Data Pool.

Extending the Period of Retention for Granules in the Data Pool

We have seen that a change in user interest in data from a particular location may arise because of unusual circumstances (e.g., weather, natural event) and that as a result it may be desirable to extend the period of retention in a Data Pool insert subscription. Such circumstances may also make it desirable to retain certain data already in the Data Pool for a longer period of time than originally specified. Data Pool maintenance personnel can run the Update Granule Utility to update the expiration date for selected science granules. This utility also permits modifying a granule's retention priority, which can affect how soon the Data Pool Cleanup Utility removes the granule from the Data Pool.

The Update Granule Utility permits updating granule information using a command-line interface. The following options may be used:

-noprompt: suppressing prompts and detailed information display.

-theme: specifies a valid theme name (i.e., a character string that matches an existing theme name in the Data Pool inventory).

A single granule may be updated using manual input. Multiple granule updates can be handled using an input file containing a list of granules to be updated, or by specifying a theme. The input file must be structured as a list of granules to be processed, one per line. Each line contains a granule ID (reflecting the Sybase entry in the Data Pool database), an expiration date, and (optionally) a new retention priority, the value of which may be null (i.e., left blank). The fields are separated by a single space. There should be no blank lines before the first or after the last granule in the list. The file contents should be similar to the following example.

```
GRANULE_ID_4832 EXP_DATE=2002/2/28 RETENTION=255
GRANULE_ID_4876 EXP_DATE=2002/2/28 RETENTION=200
GRANULE_ID_4883 EXP_DATE=2002/2/28 RETENTION=
GRANULE_ID_4937 EXP_DATE=2002/2/28
GRANULE_ID_4966 EXP_DATE=2002/2/28 RETENTION=255
```

When updating the granules associated with a theme, the utility updates the expiration date of a granule associated with that theme if and only if the new expiration date specified is later than the current expiration date of the granule. It updates the retention priority of a granule associated with that theme if and only if the new expiration priority specified is higher than the current retention priority of the granule.

The Update Granule Utility connects to the Data Pool database and calls Sybase stored procedures to perform the requested updates. Therefore, the utility runs only if the Data Pool database server is running and if the database is available. It also assumes the stored procedures are present. The Granule Update Utility may be run as a background process, with suppression of all warning/error messages and confirmation prompts if desired. When the utility is run, it writes information, any warnings, any errors, and messages to a log file about granules as they are updated.

Assume that a user contacts the DAAC with a request to update (extend) the expiration date to the end of February 2002 for selected granules in the Data Pool, and provides a list of granule IDs for the selected granules. The following procedure is applicable.

Extend the Retention for Selected Science Granules Using the Update Granule Utility

---

1 Log in at the machine on which the Update Granule Utility is installed (e.g., e0dps01, g0dps01, l0dps01, n0dps01).

2 To change to the directory for starting the Update Granule Utility, type **cd /usr/ecs/<MODE>/CUSTOM/utilities** and then press the **Return/Enter** key.

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

3 At the UNIX prompt, type the command to start the Update Granule Utility, in the form **EcDlUpdateGranule.pl <command line parameters>** then press the **Return/Enter** key.

- For this exercise, use the following command:

  EcDlUpdateGranule.pl <*MODE*> -file tr_list

- The first command-line parameter specified must be **<*MODE*>**, a valid, existing Data Pool mode (e.g., OPS, TS1, TS2).

- The following six permutations are valid command-line entries for initiating the Update Granule utility:

  – **EcDlUpdateGranule.pl <*MODE*> -file <*filename*>** (to update granules listed in an input file named **<*filename*>** while displaying all summary information to the operator, and asking confirmation of the update).

  – **Ec DlUpdateGranule.pl <*MODE*> -grnid <*granuleID*> -exp <*expiration date*> [-ret <*retention priority*>]** (to update a granule identified by its **<*granuleID*>** with a new expiration date and, optionally, a new retention priority while displaying all summary information to the operator, and asking confirmation of the update).

  – **EcDlUpdateGranule.pl <*MODE*> -noprompt -file <*filename*>** (to update granules listed in an input file named **<*filename*>** with no confirmation or information displayed to the operator).

  – **EcDlUpdateGranule.pl <*MODE*> -noprompt -grnid <*granuleID*> -exp <*expiration date*> [-ret <*retention priority*>]** (to update a granule identified by its **<*granuleID*>** with a new expiration date and, optionally, a new retention priority with no confirmation or information displayed to the operator).

  – **EcDlUpdateGranule.pl <*MODE*> -theme <*themename*> -exp <*expiration date*> [-ret <*retention priority*>]** (to update a granule identified by its **<*themename*>** with a new expiration date and, optionally, a new retention priority while displaying all summary information to the operator, and asking confirmation of the update).

  – **EcDlUpdateGranule.pl <*MODE*> -noprompt -theme <*themename*> -exp <*expiration date*> [-ret <*retention priority*>]** (to update a granule identified by its **<*themename*>** with a new expiration date and, optionally, a new retention priority with no confirmation or information displayed to the operator.

- The utility executes and displays a confirmation prompt similar to the following:

  You are about to start updating granules.

  Total number of granules: 11

Total size of granules: 8.61339673772454 MB
Do you wish to continue processing the update? [y/n]y

4 Type **y** and then press the **Return/Enter** key.

- The utility completes execution and displays output similar to the following:

  Update completed.
  Please check the database to ensure proper completion.

  Update took 2 seconds to complete

  Gracefully exiting...

- To check the database, have the Database Administrator use isql commands on the Data Pool database host to query the DlGranuleExpirationPriority table. It may also be useful to examine the Update Granule Utility log file to determine whether there were any problems with the execution. To examine that log file, go to Steps 5 and 6.

5 To change to the directory containing the Update Granule Utility log file and other log files, type **cd /usr/ecs/<MODE>/CUSTOM/logs** and then press the **Return/Enter** key.

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

6 To examine the Update Granule Utility log file, type **pg EcDlUpdateGranule.log** and then press the **Return/Enter** key.

- The first page of the log file is displayed; additional sequential pages can be displayed by pressing the **Return/Enter** key at the **:** prompt. It is also possible to search forward by typing **/<search item>**. For example, to search the log file for reference to one of the granules updated, type **/<granuleID>** and then press the **Return/Enter** key.

- Although this procedure is written for the **pg** command, any UNIX editor or visualizing command (e.g., **vi**, **view**, **more**, or **tail**) can be used to review the log.

- The log entries have a time and date stamp; about the time that the update was executed, the log should show entries similar to the following:

  2001/11/29 15:52:50.814:Update started...

  2001/11/29 15:52:50.964:Granule 4871 updated
  2001/11/29 15:52:51.083:Granule 4954 updated
  2001/11/29 15:52:51.212:Granule 4955 updated
  2001/11/29 15:52:51.346:Granule 4956 updated
  2001/11/29 15:52:51.409:Granule 4957 updated
  2001/11/29 15:52:51.688:Granule 4959 updated
  2001/11/29 15:52:51.778:Granule 4961 updated
  2001/11/29 15:52:51.998:Granule 4963 updated
  2001/11/29 15:52:52.107:Granule 4963 updated

2001/11/29 15:52:52.394:Granule 4964 updated
2001/11/29 15:52:52.569:Granule 4966 updated
2001/11/29 15:52:52.590:Update ended.

2001/11/29 15:52:52.608:This update took approximately 2 seconds

- If the log indicates errors or warnings, it may be necessary to correct the condition identified in the entry (e.g., edit the data in the granule list in the input file) and run the utility again. Specific error entries depend on the error that occurred; examples of error entries in the log may be similar to the following:

  4959 AST_04 1 0.03962299 Jul 30 2001 12:00AM Feb 2 1998 11:59PM 255 2
  Warning: The new expiration date for the above granule is less than or equal to today's date.

  DATABASE ERROR:Server message number=120001 severity=16 state=1 line=33 server=f2acg01_srvr procedure=ProcSelectGrExpiration text=ProcSelectGrExpiration: Requested granule id not in database.

  2001/11/29 15:50:36.647:Sybase Lookup ==> ERRORS WERE FOUND WITH GRANULE "4654". (It may not exist or contains the wrong format).

  2001/11/29 15:50:36.663:

  EcDlUpdateGranule_1.pl aborted due to insufficient processing data: All the granule triplets had errors.

Running the Data Pool Cleanup Utility

The Data Pool Cleanup Utility permits DAAC Operations Staff to remove expired granules from the Data Pool disks and inventory. In addition, the Data Pool Cleanup Utility reports (via an external utility) to the EOS Clearing House (ECHO) the granules that are to be (or that have been) deleted.

The Cleanup Utility may be executed using a **-noprompt** argument to suppress all confirmations and warnings normally displayed to standard output.

The Cleanup Utility must be executed on the machine where the granules are located. Qualification for cleanup is based on three criteria: expiration date/time, retention priority, and theme association. Figure 60 illustrates how the utility uses expiration date/time and retention priority in combination to select granules for removal from the Data Pool. The operator specifies values for these criteria and, if desired, themes and cross-references, to control cleanup operations. It is also possible to use the utility to validate the Data Pool inventory and disk cache by checking for and removing *orphans* (disk files not associated with any entry in the Data Pool inventory) and *phantoms* (entries in the Data Pool inventory that refer to files that no longer exist on the Data Pool disks).

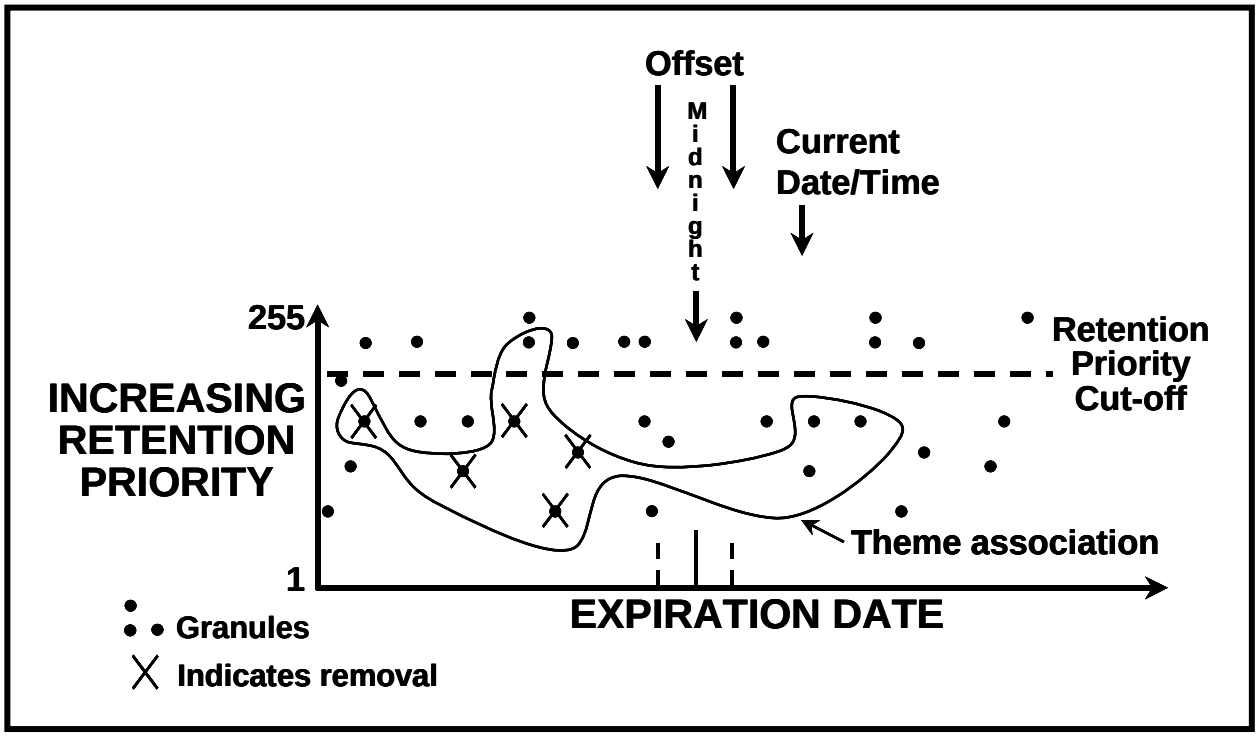


Figure 60. The Data Pool Cleanup Utility Uses Expiration Date, Retention Priority, and/or Theme Association to Qualify Granules, Marked with X, for Deletion

To determine whether a granule qualifies for deletion, the utility first compares the granule's expiration date (insert date plus retention period in days specified in the insert subscription) with a cut-off date/time. If a granule's expiration date is prior to the cut-off, the granule qualifies as expired.

- The default cut-off date/time is set to midnight of the previous day.

- The operator is permitted to specify an **offset** (from the previous midnight) in hours to add or subtract hours to determine a cut-off date/time for deletion.
  - For example, **-offset -5** would delete all granules that had expired as of 7:00 P.M. yesterday.

Next, the utility compares the granule's retention priority with any priority **limit** the operator has specified to identify those granules that should be retained in the Data Pool even though their expiration date has passed.

- Retention priority is an integer from 1 to 255.
  - For example, **-limit 150** would delete all granules with priority less than or equal to 150.
- Retention priority for granules already in the Data Pool may be modified using the granule expiration update script.

The Data Pool Cleanup Utility removes those granules with expiration date prior to the cut-off date/time and with retention priority less than or equal to the specified limit. If a priority limit is not specified in command-line input parameters at the time it is invoked, the Cleanup Utility reads the parameter 'DEFAULT_LIMIT' from its configuration file to get a priority limit. If the operator does not wish to use retention priority as a criterion for deletion, the default limit should be set to 255. If the operator specifies a theme name, the utility applies the removal criteria only to those granules associated with the theme.

The Cleanup Utility can alternatively take as input a file listing the granuleId for each granule to be deleted. The file can contain one granuleId or more than one granuleId per line separated by white space. The **–file** option may not be used with any other options other than the **-noprompt** option.

Another file option is **-geoidfile** (e.g., **-geoidfile geoid20040304**). It specifies the name of a file containing GeoIDs, which are a combination of science type, ESDT short name and version ID, and ECS Science Data Server database ID. Granules in the file whose ECS ID match those in the Data Pool are candidates for Data Pool cleanup if specified by this option. The **-geoidfile** option may not be used in conjunction with any other options other than the **-noprompt** option. Note that the GeoID file can contain science granules as well as non-science granules because the science data server may delete these types of granules. The input value for this parameter is logically defined to be the output of any Science Data Server Phase 1 (EcDsBulkDelete.pl) granule deletion run. This causes the Data Pool cleanup utility to clean up any Science Data Server granules found in the GeoID input file to be removed from the Data Pool database.

The Cleanup Utility cleans up non-ECS data just as it does ECS data. It can remove granule cross references associated with a given theme, and also remove the granules associated with the theme. The option **-themexref** specifies a theme for which all cross-references are to be removed from the Data Pool. The option **-theme** specifies a theme for which associated granules are to be removed. If a granule is referenced to more than one theme, the **-theme** option removes only the cross-reference to the specified theme, without removing the granule. The theme name must be enclosed in quotes (e.g., **-theme "Ocean Temperatures"** or **–themexref**

"Surface Reflectance"). The **-themexref** option cannot be used with any options other than the **-noprompt** option.

The **-ecsgrandel** option indicates that only granules removed in the ECS system from the Science Data Server inventory are to be removed from the Data Pool if they exist. The option may not be used in conjunction with any options other than the **-noprompt** option. No other cleanup occurs when the **-ecsgrandel** option is specified.

The **-echomode** parameter specifies the method by which the Cleanup Utility reports deletion candidates to ECHO. The **-echomode** parameter takes one of three values; i.e., **predelete**, **finishdelete** or **deleteall**.

When **predelete** is specified, the Cleanup Utility builds the list of items to clean up from the Data Pool and reports them to ECHO through the EcBmBulkURL utility. No data is actually cleaned up from the disks or database inventory using **predelete**.

When **finishdelete** is specified as the value for **-echomode**, the Cleanup Utility deletes all of the data that was last found during a run with the **predelete** parameter. The difference is that the EcBmBulkURL utility is not invoked because this run performs the actual cleanup of the database inventory and disks of what was presumably reported to ECHO during the previous run.

When **deleteall** is specified as the value for **-echomode**, the Cleanup Utility builds its list of items to clean up, actually cleans them up, and notifies ECHO via the EcBmBulkURL utility. The **deleteall** value does not allow for a time lag between the Cleanup Utility deleting the granules and ECHO performing its own clean up of URLs.

The normal sequence for cleanup is to run the Cleanup Utility twice: specifying **predelete** for the first run and **finishdelete** for the second run. Note that an **-echomode** parameter with a value of **finishdelete** can only be specified by itself because the list of items to delete will have already been determined by the previous (**predelete**) run.

If a **predelete** run is performed, the subsequent run *must* specify **finishdelete** in order to perform the actual deletions. The Cleanup Utility enforces the requirement to avoid operator error. The **predelete/finishdelete** run sequence can be viewed as a logical run done in two parts. The values of **predelete** and **deleteall** may be used with each of the other parameters specific to performing Data Pool Cleanup except **themexref**.

There are three types of runs that can be performed with the Cleanup Utility:

- **Cleanup only.**
- **Validation only.**
- **Cleanup followed by validation.**

When involved in "cleanup" processing, the Cleanup Utility performs the following actions:

- Removes from the Data Pool disks and inventory all Data Pool granules, associated browse files, and browse links that meet the specified cleanup criteria (provided that no other granules are cross-referenced to them – i.e., linked by a theme). This occurs

when the **-echomode** parameter has a value of **finishdelete** or **deleteall**. (No actual deletion occurs during **predelete**.)

- Removes all recent insert files (with names prefixed with DPRecentInsert) that are older than seven days. The relevant files are found in /datapool/<mode>/user/<fs1> and /datapool/<mode>/user/<fs1>/<group>/<esdt>.
- Exports a list of deleted granules for ECHO accessibility by invoking an external utility (i.e., EcBmBulkURLStart) when the Cleanup Utility **-echomode** parameter has a value of either **finishdelete** or **deleteall**.
  - If there are granules being deleted that qualify for ECHO export, the Cleanup Utility generates an XML file containing a list of those granules and stores it in the /datapool/<mode>/user/URLExport directory for files that are ftp pulled and ftp pushes files when Bulk URL is configured to ftp push the data to ECHO.
  - If the Data Pool Cleanup Utility is run in **-echomode finishdelete**, the EcBmBulkURLStart utility is not called.
- Removes all HEG conversion files associated with the HEG order IDs that have the status of "DONE" or "FAILED" and a timestamp older than a certain cleanup age.
  - The HEG order IDs are provided in the DlCartOrder table and the cleanup age is specified by the "HEGCleanupAge" parameter in the DlConfig table of the Data Pool database.
  - The HEG conversion files for each order ID are stored in the /datapool/<mode>/user/downloads/<orderID> directory. (HEG orders and conversion files are generated when end users request HEG-converted data using the Data Pool Web Access tool.)

When involved in “validation” processing, the Data Pool Cleanup Utility performs the following actions:

- Validates the Data Pool inventory and disk content by checking for the existence of orphans and/or phantoms and either removing them or just logging them depending on the command line options specified.

Validation requires either the **–orphan** parameter or the **–phantom** parameter or both. The **-orphan** parameter finds/removes data in the Data Pool that is not represented by entries in the Data Pool inventory. The **-phantom** parameter finds/removes entries in the Data Pool inventory that have one or more science or metadata files, or associated browse files, missing from the Data Pool. To specify just logging of the discrepancies, the operator uses the option **-nofix**.

The **-maxorphanage** validation option specifies the maximum orphan age in days (e.g., **-maxorphanage 5**). The value specified must be greater than or equal to three days. The Data Pool inventory validation function will consider only those files on disk as orphans whose age is equal to or larger than the maximum orphan age specified. If the parameter is omitted, the default value specified in the configuration file is used.

The **-collgroup** validation option limits the Data Pool validation to the specified collection group(s). Single or multiple collection groups can be specified on the command line. If multiple collection groups are specified, they must be separated by commas, with the string enclosed in double quotes (e.g., "MOAT, ASTT"). By default all collection groups in the Data Pool inventory are included in the validation if the **-collgroup** option is not specified.

A validation run can be time-consuming and should not be run as often as the cleanup runs, because it potentially involves the checking of all files in the entire Data Pool inventory against those on the Data Pool disk in order to find and remove the discrepancies. It is advised that the validation function be run using the **-collgroup** option whenever possible to limit the validation to a limited number of collection groups.

If the Cleanup Utility is interrupted during execution, upon restart it continues from the point of interruption. Furthermore, in the interest of low database contention, the Cleanup Utility allows only one instance of itself to execute concurrently.

Invoke the Data Pool Cleanup Utility Manually

---

1 Log in at the machine where the Data Pool Cleanup Utility is installed (e.g., e0dps01, g0dps01, l0dps01, n0dps01).

- The operator who is executing the script must have privileges for removing science, metadata, and browse files from the Data Pool disks.

2 To change to the directory for starting the Data Pool Cleanup Utility, type **cd /usr/ecs/<*MODE*>/CUSTOM/utilities** and then press the **Return/Enter** key.

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

NOTE: There are three types of runs that can be performed with the Cleanup Utility; i.e., "cleanup only," "validation only," or "cleanup followed by validation."

NOTE: The normal sequence for cleanup is to run the Cleanup Utility twice: specifying **predelete** for the first run and **finishdelete** for the second run. Note that an **-echomode** parameter with a value of **finishdelete** can only be specified by itself because the list of items to delete will have already been determined by the previous (**predelete**) run.

NOTE: If a **predelete** run is performed, the subsequent run *must* specify **finishdelete** in order to perform the actual deletions. The Cleanup Utility enforces that requirement to avoid operator error. The **predelete/finishdelete** run sequence can be viewed as a logical run done in two parts.

3 To perform a "cleanup only" run, at the UNIX prompt enter:

EcDlCleanupDataPool.pl <***MODE***> **-echomode** <***echomode***> **[-noprompt] [-offset** <***hours***>**] [-limit** <***priority***>**] [-theme** <***themeName***>**] [-nodmap]**

OR

EcDlCleanupDataPool.pl <***MODE***> **-echomode** <***echomode***> **[-noprompt] -file** <***fileName***>

OR

EcDlCleanupDataPool.pl <***MODE***> **-echomode** <***echomode***> **[-noprompt] -geoidfile** <***fileName***>

OR

EcDlCleanupDataPool.pl <***MODE***> **-echomode** <***echomode***> **[-noprompt] -ecsgrandel**

OR

EcDlCleanupDataPool.pl <***MODE***> **[-noprompt] -themexref** <***themeName***>

- <***echomode***> is the value specified for the ECHO mode. The value is either **predelete**, **finishdelete**, or **deleteall** (e.g., **-echomode predelete**).
- <***hours***> is the value of the offset. It can be either a positive number (e.g., **-offset 2**) or a negative number (e.g., **-offset -5**). If the **–offset** option is not specified, the Cleanup Utility uses the default value of 0 (zero).
- <***priority***> is the value of the priority limit. It is a number from 1 through 255 (e.g., **-limit 200**). If the **–limit** option is not specified, the Cleanup Utility uses the default value specified in the configuration file.
- <***themeName***> is the name of a theme to be associated with either the **–theme** option or the **–themexref** option. The name of the theme must be in quotes (e.g., **-theme "Ocean Temperatures"** or **–themexref "Surface Reflectance"**).
- <***fileName***> is the name of a file to be associated with either the **–file** option or the **-geoidfile** option (e.g., **-file clean20040404** or **-geoidfile geoid20040304**). The file will be read by the Cleanup Utility to determine what granules to clean up.
- The **-ecsgrandel** option indicates that only granules removed in the ECS system from the Science Data Server inventory will be removed from the Data Pool if they exist. No other cleanup occurs.
- The **-nodmap** option is used to avoid updating the current density map for each collection having granuleIds in DlFilesToBeDeleted.
- The Cleanup Utility runs and the Cleanup Utility log file **EcDlCleanup.log** records errors, warnings, and information about utility events.

4 To perform a “validation only” run, at the UNIX prompt enter:

EcDlCleanupDataPool.pl <*MODE*> -orphan | -phantom [-noprompt] [-collgroup <*groupList*>] [-maxorphanage <*days*>] [-nofix]

- For validation either the **–orphan** parameter or the **–phantom** parameter or both must be specified.
- <***groupList***> is the name of the collection group(s) to be validated (e.g., “MOAT, ASTT”). The collection group(s) must be enclosed in quotes and if there are multiple groups, they must be separated by commas. If the **–collgroup** option is not specified, all collection groups in the Data Pool inventory are included in the validation.
- <***days***> is the number of days (at least 3) after which files on the Data Pool disks are considered orphans if they do not have corresponding entries in the Data Pool inventory. The default value in the configuration file (e.g., 3) is used if the **-maxorphanage** option is not specified.
- The **–nofix** option prevents reconciling any discrepancies found during validation. The validation results are logged.

5 To perform a “cleanup followed by validation” run, at the UNIX prompt enter a command line with valid options from Steps 3 and 4 plus the **–cleanvalidate** parameter.

- For example:

  EcDlCleanupDataPool.pl OPS -echomode predelete -offset 5 -limit 200 -orphan -phantom -cleanvalidate

6 If **predelete** was specified as the value for the **–echomode** parameter in Step 3 or Step 5, after the Cleanup Utility has run to completion repeat Step 3 (or Step 5) to perform a cleanup using **finishdelete** as the value for the **–echomode** parameter.

In normal operations, the Cleanup Utility is run once a day as a cron job as a "cleanup only" run executing in echo mode of **predelete**. This builds the list of cleanup candidates (based on the expiration date and retention priority) that are reported to ECHO as those that will be deleted in the next run of cleanup. Also, the granules that have been entered in the Science Data Server’s deleted granules table will be reported. On a subsequent run within the same 24-hour period, the cleanup utility is run in **finishdelete** mode to perform the actual cleanup processing that was reported to ECHO in the **predelete** mode.

Establish Data Pool Cleanup to Run with *cron*

---

1 Log in at a system platform using an account with privileges to remove science, metadata, and browse files from Data Pool disks.

2 To ensure that the **crontab** command launches the vi editor, type **setenv EDITOR vi** and then press the **Return/Enter** key.

- It may be desirable to include this command in the operator's **.cshrc** file to set the **crontab** editor to **vi** as part of the environmental settings normally used routinely.

3 Type **crontab -e** and then press the **Return/Enter** key.

- The contents of the file are displayed, and the cursor is displayed on the first character at the upper left corner of the file.
- If the operator has no **crontab** file on the current platform, this command opens a new one for editing.

4 If necessary, use the down arrow key on the keyboard to move the cursor down to a blank line.

- The cursor is displayed at the beginning of the selected line.

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

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

6 Type the crontab entry, including the appropriate Cleanup Utility command (as described in Section 17.10.32, **Invoke the Data Pool Cleanup Utility Manually**).

- For example:

  0 1 * * * /usr/ecs/OPS/CUSTOM/utilities/EcDlDataPoolCleanup.pl OPS -echomode predelete -noprompt

- The example would start a **predelete** cleanup run at 1:00 A.M. every day.

7 Repeat Step 6 as necessary to enter additional crontab entries, including the appropriate Cleanup Utility command (e.g., to run a **finishdelete** cleanup run at 4:00 A.M. every day).

8 Press the **Esc** key.

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

9 Type **:wq** and then press the **Return/Enter** key.

- UNIX displays a message identifying the number of lines and characters in the **crontab** file (stored in the directory **/var/spool/cron/crontabs**) and then displays the UNIX prompt.

Running the Data Pool Access Statistics Utility

The Data Pool Access Statistics Utility (DPASU) parses logs of the Data Pool Web Access service and the FTP access service and stores the results in tables in the Data Pool database. The DPASU is a command-line utility that permits an option of entering input parameters. It is intended to be run with ***cron*** to cover an arbitrary 24-hour period starting at a time specified as a configuration parameter in a configuration file. However, an operator may run the utility from the command line specifying a start date as an input parameter to cover a period other than the normal 24-hour period addressed by ***cron*** or to cover that normal period if ***cron*** failed to process the logs for that period.

There are two versions of the DPASU, one for each type of log processed. The script named **EcDlRollupWebLogs.pl** runs on the Data Pool Web Access server and processes its log; its configuration file is **EcDlRollupWebLogs.CFG**. The script named **EcDlRollupFtpLogs.pl** runs on a server with access to SYSLOG with FTP access entries; its configuration file is **EcDlRollupFtpLogs.CFG**. These scripts capture data on downloads from the Data Pool, including date and time of access, path and file name of the file, and size of the file. The captured data are written to a temporary "flat file" -- a tab-delimited text file -- stored in the directory /<ECS_HOME>/<*MODE*>/CUSTOM/data/DPL/. The flat file is then exported to Sybase and stored in a table. The DPASU calls Sybase stored procedures to generate a separate rollup table, removes the flat file, and enters a record in a separate table identifying which periods have been rolled up in order to prevent inadvertent reprocessing of that period.

To prevent potential table locking, ***cron*** runs of the DPASU scripts should be separated so that they are not both running concurrently (e.g., separate their start times by at least 20 minutes). Use the following procedure to specify a 1:00 a.m. start time for the rollup and add a line to the ***crontab*** files to run the DPASU for the OPS mode beginning at 2:00 a.m. every day with a 20-minute separation between the scripts.

Specify Data Pool Access Statistics Rollup Start Time and DPASU Execution with *cron*

1 Log in at the host for EcDlRollupWebLogs.pl and its configuration file (e.g., e0dps01, g0dps01, l0dps01, or n0dps01).

2 To change to the directory containing the configuration file, type the command **cd /usr/ecs/OPS/CUSTOM/cfg** and then press the **Return/Enter** key.

- The working directory is changed to **/usr/ecs/OPS/CUSTOM/cfg**.

3 To look at the Rollup Start Time specified in the configuration file, type **vi EcDlRollupWebLogs.CFG** and then press the **Return/Enter** key.

- The contents of the file are displayed, and the last line of the file indicates the start time in format similar to the following:

  ROLLUP_START_TIME=3:00

  and the cursor is displayed on the first character at the upper left corner of the file.

- If the start time is correct, exit **vi** by typing **:q!** and pressing the **Return/Enter** key; then go to Step 10. Otherwise, to change the time, execute Steps 4 - 9.

4 Use the arrow keys on the keyboard to move the cursor down to the line specifying the **ROLLUP_START_TIME** and to move it to the right until it is located over the first character in the time value.

- The cursor is moved to the start time location; the line should look similar to the following:

  ROLLUP_START_TIME=3:00

5 Type **x** to delete the number under the cursor.

- The number is deleted; the line should look similar to the following.

  ROLLUP_START_TIME=:00

- If more characters in the time value are to be changed, you can type **x** repeatedly to delete additional characters. For this exercise, you need only delete one character.

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 Type **1**.

- The typed entry appears to the left of the cursor.

8 Press the **Esc** key.

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

9 Type **ZZ** (be sure to use upper case).

- The file is saved and the UNIX prompt is displayed.

10 To ensure that the **crontab** command launches the **vi** editor, type **setenv EDITOR vi** and then press the **Return/Enter** key.

- It may be desirable to include this command in the operator's **.cshrc** file to set the **crontab** editor to **vi** as part of the environmental settings normally used routinely.

11 Type **crontab -e** and then press the **Return/Enter** key.

- The contents of the file are displayed, and the cursor is displayed on the first character at the upper left corner of the file.
- If the operator has no **crontab** file on the current platform, this command opens a new one for editing.

12 If necessary, use the down arrow key on the keyboard to move the cursor down to a blank line.

- The cursor is displayed at the beginning of the selected line.

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

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

14 Type **0 2 * * * /usr/ecs/OPS/CUSTOM/utilities/EcDlRollupWebLogs.pl OPS -noprompt**.

- The typed entry appears to the left of the cursor.

15 Press the **Esc** key.

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

16 Type **:wq** and then press the **Return/Enter** key.

- UNIX displays a message identifying the number of lines and characters in the **crontab** file (stored in the directory **/var/spool/cron/crontabs**) and then displays the UNIX prompt.

17 Log in at the host for EcDlRollupFtpLogs.pl and its configuration file (e.g., e0dps01, g0dps01, l0dps01, n0dps01).

18 To change to the directory containing the configuration file, type the command **cd /usr/ecs/OPS/CUSTOM/cfg** and then press the **Return/Enter** key.

- The working directory is changed to **/usr/ecs/OPS/CUSTOM/cfg**.

19 To look at the Rollup Start Time specified in the configuration file, type **vi EcDlRollupFtpLogs.CFG** and then press the **Return/Enter** key.

- The contents of the file are displayed, and the last line of the file indicates the start time in format similar to the following:

  ROLLUP_START_TIME=3:00

  and the cursor is displayed on the first character at the upper left corner of the file.
- If the start time is correct, exit **vi** by typing **:q!** and pressing the **Return/Enter** key; then go to Step 21. Otherwise, to change the time, execute Step 20.

20 Repeat Steps 4-9 to change the time in **EcDlRollupFtpLogs.CFG**.

21 To ensure that the **crontab** command launches the **vi** editor, type **setenv EDITOR vi** and then press the **Return/Enter** key.

- It may be desirable to include this command in the operator's **.cshrc** file to set the **crontab** editor to **vi** as part of the environmental settings normally used routinely.

22 Type **crontab -e** and then press the **Return/Enter** key.

- The contents of the file are displayed, and the cursor is displayed on the first character at the upper left corner of the file.
- If the operator has no **crontab** file on the current platform, this command opens a new one for editing.

23 If necessary, use the down arrow key on the keyboard to move the cursor down to a blank line.

- The cursor is displayed at the beginning of the selected line.

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

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

25 Type **20 2 * * * /usr/ecs/OPS/CUSTOM/utilities/EcDlRollupFtpLogs.pl OPS -noprompt**.

- The typed entry appears to the left of the cursor.

26 Press the **Esc** key.

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

27 Type **:wq** and then press the **Return/Enter** key.

- UNIX displays a message identifying the number of lines and characters in the **crontab** file (stored in the directory **/var/spool/cron/crontabs**) and then displays the UNIX prompt.

Although the Data Pool Access Statistics Utility scripts are intended to be run with **cron**, if it is necessary to run them from the command line, it is possible to do so. For example, if **cron** fails to complete successfully for any reason, no entry is made into the record table to indicate that a period was processed. In that event, the statistics can be captured for the missing interval by running the utility manually.

There are seven command-line parameters for use with the utility scripts (see 609 document information):

- The **<*MODE*>** parameter indicates the mode (must specify a valid directory path) in which the script is to run; it is mandatory, unlabeled, and must be the first parameter following the command.

- The **-noprompt** parameter optionally specifies suppression of output to the screen.
- The **-nodelete** parameter optionally prevents the flat file from being deleted upon completion of the run.
- The **-flatfile <*path/file*>** parameter optionally provides an alternative path/file name for the flat file produced by the parser (useful only with the **-nodelete** option).
- The **-ftp <*path/file*>** parameter optionally indicates an alternative ftp log path/file(s) to be used instead of the configured default path/file (for the **EcDlRollupFtpLogs.pl** script only). Wildcards may be used, but must be escaped (i.e., preceded with a \).
- The **-web <*path/file*>** parameter optionally indicates an alternative web log path/file(s) to be used instead of the configured default path/file (for the **EcDlRollupWebLogs.pl** script only). Wildcards may be used, but must be escaped (i.e., preceded with a \).
- The **-start <*date*>** parameter optionally indicates an alternative start date for the rollup period, using the format MM/DD, and may be used to process a previously uncovered period.

With the exception of the mandatory **<*MODE*>** parameter, which must appear first after the command, the other parameters may be used in various orders and combinations. For example, to run without screen prompts or information, starting from December 22, and to retain the flat file, the command for accumulating statistics on web access should be entered as follows:

EcDlRollupWebLogs.pl OPS -noprompt -nodelete -start 12/22.

To run with normal screen information display, starting from February 15, but using an alternative file with wildcards for the web log, the command should be similar to the following:

EcDlRollupWebLogs.pl OPS - start 2/15 -web /usr/var/*.log.

Use the following procedure to run the Data Pool Access Statistics Utility scripts from the command line, with normal screen information display.

Specify Data Pool Access Statistics Utility Execution from the Command Line

---

1 Log in at the host for EcDlRollupWebLogs.pl and its configuration file (e.g., e0dps01, g0dps01, l0dps01, or n0dps01).

2 To change directory to the directory containing the script, type the command **cd /usr/ecs/<*MODE*>/CUSTOM/utilities** and then press the **Return/Enter** key.

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

3 Type **EcDlRollupWebLogs.pl <*MODE*>** and then press the **Return/Enter** key.

- The utility runs and displays information to the screen as it executes, in form similar to the following:

```
A Synergy II/Data Pool product
 ___  __   __   __  _    _
|    \| _\/   | /  _\| |  | |
|  0  |   / 0 |\  \  | |  | |
|     |  | /|  |_\   \ | \_/ |
|___/ |_|/  |_ |__  /   \___/
Data Pool Access Statistics Utility

Connecting to database...

The DPASU will examine the logs for access entries between the following times:
            Month   Day    Hour   Minute
---------------------------------------------
START:  11          26      03      00
   END:  11          27      02      59

Checking for already covered rollup periods...


File list:
/usr/ecs/OPS/COTS/www/ns-home/www/logs/access
Processing Web logs...
No access entries found in any of the Web logs

Cleaning up table "DlWebAccessLog"...OK

Exporting flat file to Sybase...OK

No access data was available to roll up.
DPASU will skip this step.

Rollup successful!
Removing flat file...OK
Gracefully exiting...
```

4 Log in at the host for EcDlRollupFtpLogs.pl and its configuration file (e.g., e0dps01, g0dps01, l0dps01, n0dps01).

5 To change directory to the directory containing the script, type the command **cd /usr/ecs/<*MODE*>/CUSTOM/utilities** and then press the **Return/Enter** key.

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

6 Type **EcDlRollupFtpLogs.pl <*MODE*>** and then press the **Return/Enter** key.

- The utility runs and displays information to the screen as it executes, in form similar to the following:

```
A Synergy II/Data Pool product

 ___  __   __   __  _    _
|    \| _\/   |/  _\| |  | |
|  0  |   / 0 |\  \  | |  | |
|     | | /|  |_\   \ | \_/ |
|___/ |_|/  |_ |__  /   \___/
Data Pool Access Statistics Utility

Connecting to database...

The DPASU will examine the logs for access entries between the following times:
            Month   Day    Hour   Minute
----------------------------------------------
START:  11          26      03      00
   END:  11         27      02      59

Checking for already covered rollup periods...


File list:
/var/adm/SYSLOG
Processing FTP logs...
No access entries found in any of the FTP logs

Cleaning up table "DlFtpAccessLog"...OK

Exporting flat file to Sybase...OK

No access data was available to roll up.
DPASU will skip this step.

Rollup successful!
Removing flat file...OK
Gracefully exiting...
```

The three remaining utilities are shell scripts for archiving, deleting, and restoring information in database tables populated by the DPASU. The **Data Pool Archive Access Statistics Data Utility** is run from the command line as needed or desirable to connect to the Data Pool database and write granule access data for a specified time range from the DlGranuleAccess, DlGranuleSubscription, and DlAccessRollup tables to an ASCII file. Once this is done, the operator can run the **Data Pool Delete Access Statistics Data Utility** from the command line to delete the archived data from the Data Pool database. If it is desirable to restore deleted data to

the database, the **Data Pool Restore Access Statistics Data Utility** can be run from the command line to restore the data. The following procedures are applicable.

Archive Access Statistics Using the Data Pool Archive Access Statistics Data Utility

---

1 Log in at the host for the Data Pool database (e.g., e0acg11, g0acg01, l0acg02, n0acg01).

2 To change directory to the directory containing the Data Pool Archive Access Statistics Data Utility, type **cd /usr/ecs/<*MODE*>/CUSTOM/dbms/DPL** and then press the **Return/Enter** key.

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

3 Type **DlDbArchiveAccessStat <*MODE*> <*STARTDATE*> <*STOPDATE*> <*ARCHIVEDIR*> <*USERNAME*> <*SERVER*> <*DBNAME*>** and then press the **Return/Enter** key.

- **<*MODE*>** is the mode in which the utility is being executed (e.g., OPS, TS1, TS2).
- **<*STARTDATE*>** is the start date time range, in format ***yyyymmdd***, for the data to be archived.
- **<*STOPDATE*>** is the stop date time range, in format ***yyyymmdd***, for the data to be archived.
- **<*ARCHIVEDIR*>** is the absolute path where the generated ASCII files are to be stored.
- **<*USERNAME*>** is the Sybase login name.
- **<*SERVER*>** is the Sybase Server for the Data Pool database (e.g., e0acg11_srvr, g0acg01_srvr, l0acg02_srvr, n0acg01_srvr).
- **<*DBNAME*>** is the name of the Data Pool database (e.g., DataPool_OPS).
- The script displays a prompt for entry of the password for the Sybase login.

NOTE: The step that follows may require input from the Database Administrator.

4 Type **<*password*>** and then press the **Return/Enter** key.

- The script runs and the Archive Access Statistics Utility log file **DlDbArchiveAccessStat.log** records errors, warnings, and information about utility events. The log is written to the directory **/usr/ecs/<*MODE*>/CUSTOM/logs**.

To run the Data Pool Delete Access Statistics Data Utility, use the following procedure.

Delete Access Statistics Using the Data Pool Delete Access Statistics Data Utility

---

1 Log in at the host for the Data Pool database (e.g., e0acg11, g0acg01, l0acg02, n0acg01).

2 To change directory to the directory containing the Data Pool Delete Access Statistics Data Utility, type **cd /usr/ecs/<*MODE*>/CUSTOM/dbms/DPL** and then press the **Return/Enter** key.

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

3 Type **DlDbDeleteAccessStat <*MODE*> <*STARTDATE*> <*STOPDATE*> <*USERNAME*> <*SERVER*> <*DBNAME*>** and then press the **Return/Enter** key.

- ***<MODE>*** is the mode in which the utility is being executed (e.g., OPS, TS1, TS2).
- ***<STARTDATE>*** is the start date time range, in format ***yyyymmdd***, for the data to be deleted.
- ***<STOPDATE>*** is the stop date time range, in format ***yyyymmdd***, for the data to be deleted.
- ***<USERNAME>*** is the Sybase login name.
- ***<SERVER>*** is the Sybase Server for the Data Pool database (e.g., e0acg11_srvr, g0acg01_srvr, l0acg02_srvr, n0acg01_srvr).
- ***<DBNAME>*** is the name of the Data Pool database (e.g., DataPool_OPS).
- The script displays a prompt for entry of the password for the Sybase login.

NOTE: The step that follows may require input from the Database Administrator.

4 Type ***<password>*** and then press the **Return/Enter** key.

- The script runs and the Delete Access Statistics Utility log file **DlDbDeleteAccessStat.log** records errors, warnings, and information about utility events. The log is written to the directory **/usr/ecs/<*MODE*>/CUSTOM/logs**.

To run the Data Pool Restore Access Statistics Data Utility, use the following procedure.

Restore Access Statistics Using the Data Pool Restore Access Statistics Data Utility

---

1 Log in at the host for the Data Pool database (e.g., e0acg11, g0acg01, l0acg02, n0acg01).

2 To change directory to the directory containing the Data Pool Restore Access Statistics Data Utility, type **cd /usr/ecs/<*MODE*>/CUSTOM/dbms/DPL** and then press the **Return/Enter** key.

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

3 Type **DlDbRestoreAccessStat <*MODE*> <*STARTDATE*> <*STOPDATE*> <*ARCHIVEDIR*> <*USERNAME*> <*SERVER*> <*DBNAME*>** and then press the **Return/Enter** key.

- **<*MODE*>** is the mode in which the utility is being executed (e.g., OPS, TS1, TS2).
- **<*STARTDATE*>** is the start date time range, in format ***yyyymmdd***, for the data to be restored.
- **<*STOPDATE*>** is the stop date time range, in format ***yyyymmdd***, for the data to be restored.
- **<*ARCHIVEDIR*>** is the absolute path of the storage location for the ASCII files containing the data to be restored.
- **<*USERNAME*>** is the Sybase login name.
- **<*SERVER*>** is the Sybase Server for the Data Pool database (e.g., e0acg11_srvr, g0acg01_srvr, l0acg02_srvr, n0acg01_srvr).
- **<*DBNAME*>** is the name of the Data Pool database (e.g., DataPool_OPS).
- The script displays a prompt for entry of the password for the Sybase login.

NOTE: The step that follows may require input from the Database Administrator.

4 Type **<*password*>** and then press the **Return/Enter** key.

- The script runs and the Archive Access Statistics Utility log file **DlDbRestoreAccessStat.log** records errors, warnings, and information about utility events. The log is written to the directory **/usr/ecs/<*MODE*>/CUSTOM/logs**.

Using the Batch Insert Utility

The Batch Insert Utility allows operators to specify Data Pool insert for granules residing in the ECS archive, as well as data from outside ECS (non-ECS granules). The utility queues the granules up for dispatch by the Data Pool Action Dispatcher (DPAD) for insertion by the Data Pool Insert Utility (DPIU). It accepts either a list of ECS granule identifiers or a list of non-ECS names; the list can be provided either as an input file or as standard input. A label identifying a batch of granules is specified as a command-line parameter, using the option **-label**, so that operators can monitor a batch with the DPM GUI.

Granules to be inserted can also be linked to a theme, using the option **-theme**. In fact, the Batch Insert Utility can also be used with that option to link granules already present in the Data Pool to a theme, or to additional themes. However, it is important to note that if the granules were

originally inserted into the Data Pool using the Batch Insert Utility, you must use a different batch label when linking the granules to the theme than was used for the original insert. This is necessary because the Batch Insert Utility is designed to reject inserts that are in a batch with a label identical to one for which granules are already being processed. So, even if the batch has been inserted, if the inserts are still in the queue (e.g., with a status of **Completed**), you cannot run another batch with the same label to link them to a theme.

The following procedure is applicable.

Perform Batch Insert of Data into the Data Pool

---

1 Log in at the machine where the Data Pool Batch Insert Utility is installed (e.g., e0dps01, g0dps01, l0dps01, or n0dps01).

NOTE: The login must be as either **cmshared** or **allmode** to ensure correct permissions.

2 To change to the directory for starting the Batch Insert Utility, type **cd /usr/ecs/<*MODE*>/CUSTOM/utilities** and then press the **Return/Enter** key.

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

3 At the UNIX prompt, type the command to start the Batch Insert Utility, in the form **EcDlBatchInsert.pl <*MODE*> -ecs | -nonecs [-file <*pathname*>] [*options*]** then press the **Return/Enter** key.

- The following examples show valid command-line entries for initiating the Batch Insert Utility:
  - **EcDlBatchInsert.pl <*MODE*> -ecs -file /home/cmshared/<*filename*>** (to add actions to the action insert queue for all ECS granules specified by granule IDs in the specified file. Because the command does not specify a **-label** parameter, the label is formed from the first 16 characters of the input file name).
  - **EcDlBatchInsert.pl <*MODE*> -nonecs -file /home/cmshared/<*filename*> -label Chig_volcano -theme "Chiginagak Volcano 2002"** (to add actions to the insert action queue for all non-ECS granules specified by XML pathnames in the specified input file, with all granules linked with the theme name **"Chiginagak Volcano 2002"** in the Data Pool database).
    - The theme name must already be in the Data Pool database in the DlThemes table; if necessary, use the DPM GUI **Manage Themes** tab to define the theme before running the batch insert.
    - You can use Batch Insert with the **-theme** option to link granules already in the Data Pool to a theme, but if the granules were originally inserted using the Batch Insert Utility, you must use a different batch label than was used for the original insert; otherwise, the insert of the theme links may be rejected.

- **EcDlBatchInsert.pl <*MODE*> -ecs -file /home/cmshared/<*filename*> -mdonly** (to add actions to the action insert queue for all ECS granules specified by granule IDs in the specified file, but insert metadata only. Because the command does not specify a **-label** parameter, the label is formed from the first 16 characters of the input file name).
- **EcDlBatchInsert.pl <*MODE*> -ecs -file /home/cmshared/<*filename*> -rpriority 255** (to add actions to the action insert queue for all ECS granules specified by granule IDs in the specified file, and to set their retention priority to 255. Because the command does not specify a **-label** parameter, the label is formed from the first 16 characters of the input file name).
- **EcDlBatchInsert.pl <*MODE*> -ecs -file /home/cmshared/<*filename*> -rpriority 255 -rperiod 10 -dpriority 5** to add actions to the action insert queue for all ECS granules specified by granule IDs in the specified file, and to set their retention priority to 255 and their retention period to 10 days, with dispatch priority set to 5. Because the command does not specify a **-label** parameter, the label is formed from the first 16 characters of the input file name).

- The Batch Insert Utility runs and events and errors are recorded in the Batch Insert Utility log file **EcDlBatchInsert.log**.

Running the Bulk Metadata Generation Tool (BMGT)

In order to support the development of value-added providers (e.g., IIMS, ESIPs, RESACs, and InfoMarts), ECS sites make an external representation of their metadata holdings available and provide a capability for bulk distribution of browse data through normal ECS distribution methods.

Several ECS data collections are created maintained at each site to store the bulk metadata and bulk browse data. Bulk Metadata Generator and Bulk Browse Generator Tools are run daily at each site to populate the data collections. One metadata product is created per ESDT group per day. Each product contains an external representation of the metadata for each new, updated, or deleted granule that is a member of the ESDT group. The format used for the external representation of the metadata is XML. One bulk browse product is produced per day that contains references to all new, updated, or deleted browse granules. Value-added providers may use any of the standard ECS search, order, and subscription capabilities to find and order these bulk metadata and browse products.

Each of the BMGT-specific ESDTs stores products that contain an XML representation of the following kinds of data or metadata:

- ECSMETC - ECS collection-level metadata and the packaging options that may be used when ordering products from each collection.
- ECSMETG - ECS granule-level metadata.

- ECSBBR - Browse images.
- ECSMETV - ECS collection and granule valid values.
- ECSMETU - ECS granule-level QA Updates.

The ECSMETC and ECSMETG data collections store products that contain metadata for multiple collections and multiple granules. The metadata is grouped by instrument and mission except for metadata related to the MODIS instrument, which is grouped, by mission and major discipline (ocean, atmosphere, land, and snow & ice). Each product in these collections has a group identifier Product Specific Attribute (PSA) called GroupId. The mapping of specific ESDTs to groups is provided as a configuration file with the BMGT.

The ECSBBR collection stores products that contain browse product references. The ESDT has a custom acquire service that converts the browse product references into actual browse products during distribution.

The ECSMETV collection stores products that contain the entire set of valids contained within the SDSRV database for a particular instance in time.

For each of the archivable XML products, BMGT creates a .met (metadata) file. In addition, it creates a Product Delivery Record (PDR) for each type of product. BMGT places XML products, .met and PDR files in a polling directory for Ingest pick up via the Science Investigator-Led Processing System (SIPS) interface. After successful archiving, the products are pushed to ECHO via SIPS interface.

The distribution of bulk browse products is not a function of BMGT. It is managed by OMS as follows:

- If bulk browse (ECSBBR collection) is set up in OMS for Synergy III processing, when a user requests an ECSBBR granule, the request is submitted to the SDSRV.
  - The SDSRV retrieves the ECSBBR granule from the archive; i.e., it obtains the file which contains the browse cross-reference.
  - SDSRV requests the distribution of the browse files referenced in the cross-reference file plus the original cross-reference file from the DDIST CI.
  - DDIST stages the browse file (via Storage Management) from the Archive. From a DDIST perspective, the ECSBBR granule looks like a multi-file granule.
- With Release 7.11 bulk browse could be reconfigured in OMS to remove the Synergy III exception for ECSBBR. In this case the OMS submits an insert request for the ECSBBR granule to the DPL CI (rather than SDSRV).
  - DPL stages the cross-reference file in the Data Pool hidden directory structure.
  - OMS executes the OMS Bulk Browse Utility, which extracts the browse cross-reference and copies into the Data Pool Storage Area Network (SAN) any relevant browse granule files that don’t reside there already.

- OMS Bulk Browse Utility updates the file list for the granule in OMS to include the new files.
- OMS performs the remainder of the distribution as usual. To OMS, the granule looks no different than any other multi-file granule. (To DPL the granule looks like a single-file granule of the ECSBBR collection. DPL does not need to provide any special handling for the granules.)

Normal operation of the BGMT includes the following events, which occur on a daily basis:

- The BMGT is invoked at the specified time each day via a cron job.
- The BGMT determines the date of the previous day and then executes a series of Sybase stored procedures against the SDSRV inventory database to extract metadata for all collections, granules, browse and valids that were inserted, updated, or deleted during the target day.
- For each target ESDT version that had collection level metadata inserted, updated, and/or deleted, perform the following steps:
  - If this collection is the first collection in a group, then create a new file and append an XML representation of the packaging options to the file.
  - Append an XML representation of the collection level metadata to the file.
- Insert each file, as a product, into the ECSMETC data collection setting the value for the GroupId PSA along with the starting date and ending date of the insert, update, and/or delete activity covered by this file. In this case, the starting and ending dates are the same since the period covered is a single day.
- For each target ESDT version that had granule level metadata inserted, update, and/or deleted, perform the following steps:
  - If this granule is the first granule in a group, then create a new file.
  - Append an XML representation of the metadata for each active granule to the file.
- Insert each file, as a product, into the ECSMETG data collection setting the value for the GroupId PSA along with the starting date and ending date of the insert, update, and/or delete activity covered by this file. In this case, the starting and ending dates are the same since the period covered is a single day.
- For all browse images that were inserted or deleted within a specified time period, extract the browse identifiers and associated browse file names for each browse product and insert an XML file, called the Browse Reference File (BRF) file, as a product into the ECSBBR data collection. Set the value for start and end date of the insert, update, and/or delete activity covered by this file, allowing it to be an ECS product that can be ordered and distributed via normal ECS search and order mechanisms.

- If any collections were inserted, updated, and/or deleted during the period then create a new file, append an XML representation of the valids information, and insert the file, as a product, into the ECSMETV data collection setting the starting date and ending date of the insert, update, and/or delete activity covered by this file. In this case, the starting and ending dates are the same since the period covered is a single day.

Occasionally, the cron job that automatically executes the BGMT fails to operate on a daily basis. This may happen due to a variety of reasons including software failure, hardware failure, or changes in DAAC operational priorities. When this happens, it is necessary for DAAC operations to manually invoke the tool to generate ECSMET products for all days that were missed. The sequence of operations is described in the preceding list of events.

Running the Bulk Metadata Generation Tool (BMGT)

1 Log in at the machine where the Bulk Metadata Generation Tool (BMGT) is installed (e.g., e0acs11, g0acs11, l0acs03, or n0acs04).

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

- Change directory to the directory containing the Bulk Metadata Generation Tool startup script (e.g., EcBmBMGTStart).
- <***MODE***> will most likely be one of the following operating modes:
  - OPS (for normal operation).
  - TS1 (for SSI&T).
  - TS2 (new version checkout).
- Note that the separate subdirectories under /usr/ecs apply to different operating modes.

3 To run the Bulk Metadata Generation Tool, at the UNIX prompt enter (as applicable):

EcBmBMGTStart

OR

EcBmBMGTStart -P <PRODUCT>

OR

EcBmBMGTStart -P <PRODUCT> -I <INPUTSOURCE> -S<InputSOURCEPRODUCT>

OR

EcBmBMGTStart -P <PRODUCT> -I <INPUTSOURCE> -L <LIST_ONLY> -S<InputSOURCEPRODUCT>

OR

EcBmBMGTStart -P <PRODUCT> -I <INPUTSOURCE> -L <LIST_ONLY> -S<InputSOURCEPRODUCT> -A <ARCHIVE>

OR

EcBmBMGTStart -P <PRODUCT> -I <INPUTSOURCE> -F <FILE_NAME> -S<InputSOURCEPRODUCT>

OR

EcBmBMGTStart -P <PRODUCT> -I <INPUTSOURCE> -L <LIST_ONLY> -S<InputSOURCEPRODUCT> -A <ARCHIVE> -F <FILE_NAME>

OR

EcBmBMGTStart -I <INPUTSOURCE> -S<InputSOURCEPRODUCT>

OR

EcBmBMGTStart -I <INPUTSOURCE> -L <LIST_ONLY> -S<InputSOURCEPRODUCT>

OR

EcBmBMGTStart -I <INPUTSOURCE> -L <LIST_ONLY> -A <ARCHIVE> -S<InputSOURCEPRODUCT>

OR

EcBmBMGTStart -I <INPUTSOURCE> -L <LIST_ONLY> -A <ARCHIVE> -F <FILE_NAME> -S<InputSOURCEPRODUCT>

OR

EcBmBMGTStart -I <INPUTSOURCE> -L <LIST_ONLY> -F <FILE_NAME> -S<InputSOURCEPRODUCT>

OR

EcBmBMGTStart -I <INPUTSOURCE> -F <FILE_NAME> -S<InputSOURCEPRODUCT>

- **<PRODUCT>** specifies what product the BMGT is to generate. The value is either **Valids**, **Collection**, **Granule**, or **Browse**.
- **<INPUTSOURCE>** is the input source for the processing of granules. The value is either **FILE**, or **GRANULES**.
  - Input source can be a file (absolute path) with GranuleID's/GeoId's or both.
  - Input source can also be granuleId/GeoID as string as the command line option.

- Multiple GranId's/GeoId's be placed inside " " as one string with spaces separating each dbID.

- **<InputSOURCEPRODUCT**> is the input source product for the processing of granules. The value is either **Granule, Browse, or Both**.
  - This is the product that is generated for the list provided by INPUTSOURCE.
  - Option “Both” will generate Granule and Browse products for the list of DbID’s specified by <INPUTSOURCE>.
- **<PRODUCT>** and **<INPUTSOURCE>** both are being provided as arguments. The value is either **Granule, Browse, or Both and** either **Valids**, **Collection**, **Granule**, or **Browse**.
  - BMGT will generate the Product specified by <PRODUCT> along with the product specified by <InputSOURCEPRODUCT> for all ESDTs and list of dbIDs in the INPUTSOURCE.
- **<LIST_ONLY**> is the output BMGT will generate. The value is either **[Y/N]**
  - If LIST_ONLY is Y, BMGT will generate XML, MET and PDR for only the list of granules in input source based on <InputSOURCEPRODUCT>.
  - If LIST_ONLY is Y and PRODUCT is Valids it generates XML, MET, PDR for Valids and the LIST in the input source.
  - If LIST_ONLY is N, BMGT will generate all the products based on other parameters including the list of granules from the Input source.
- **<FILE_NAME**> is a file name for renaming files when processing granules and Browse XML files (only).
  - This is used to rename the Granule and Browse XML files (only) generated for the list of dbIDs from the input source.
  - The file should have .XML/.xml extension.
  - If absolute path is not provided the renamed file is placed in the Output directory for BMGT.
  - If the renamed file cannot be placed in the directory specified (OS constraints) by the absolute path the renamed file will be placed in the Output directory instead.
- **<ARCHIVE**> is the output BMGT will generate This value can be either **[Y/N].**
  - If ARCHIVE is Y .MET and .PDR will be generated along with XML for the List of granules in Input source.
  - If Archive is N (this can only be when LIST_ONLY is set to Y not otherwise) .MET and .PDR will not be generated for granules in Input source.

- **<InputSOURCEPRODUCT**> is ignored if <INPUTSOURCE> is not provided.
- **<InputSOURCEPRODUCT**> is defaulted to "Both" if <INPUTSOURCE> is provided and <InputSOURCEPRODUCT> is not provided as argument.
- If <INPUTSOURCE> is not provided BMGT ignores <InputSOURCEPRODUCT>, <LIST_ONLY>, <FILE_NAME>, and <ARCHIVE> flags.
- If <LIST_ONLY> is N, <ARCHIVE> Flag is ignored.
- The BMGT produces ECSMETC, ECSMETG, ECSBBR and ECSMETV products, which consist of XML files.
- The BMGT runs and the BMGT ALOG file, **EcBmBMGT.ALOG**, and debug log, **EcBmBMGTDebug.log**, record errors, warnings, and information about utility events.

Running the Bulk URL Utility

The **Bulk URL Utility (EcBmBulkURL)** exports to ECHO the metadata content of products in the Data Pool. It can be run with either an “Insert” option or a “Delete” option.

With the “Insert” option, the Bulk URL utility is run on a daily cron job or from the command line, and should be run after the daily BMGT run has completed (this maximizes the probability that ECHO has already received information about the granules from the BMGT and has added the granules to its inventory).

With the “Delete” option the Bulk URL utility is run as an integral part of the Data Pool Cleanup utility. When the Cleanup utility is run, the operator chooses whether to run the cleanup as a one pass (setting the echo mode command line parameter to “deleteall”) or two-pass (setting the echo mode command line parameter to “predelete’ on the first pass, and setting the echomode command line parameter to “finishdelete” on the second pass) process.

The procedure for running the Bulk URL Utility is based on the following assumptions:

- Data Pool database server is running.
- Data Pool database is available.
- Stored procedures are present.

Running the Bulk URL Utility

---

1 Log in at the host for the Bulk URL Utility (e.g., e0dps01, g0dps01, l0dps01, n0dps01).

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

- The working directory is changed to cd /usr/ecs/<*MODE*>/CUSTOM/ utilities directory which contains the Bulk URL Utility start script (.g., EcBmBulkURLStart).

3 Type **EcBmBulkURLStart <*MODE*> [Insert | Delete]** then press the **Return/Enter** key.

- **<*MODE*>** is the mode in which the utility is being executed (e.g., OPS, TS1, or TS2).
- **Insert** option - the Bulk URL utility will export an xml representation of the ftp URL information for science files, metadata files, and browse files associated with Data Pool granules, which meet all of the following conditions:
  - The granule is in a collection that is eligible for export (from Data Pool) to ECHO.
    - This is determined by whether the bulkExportedFlag = "Y" for that collection in the DlCollections table in the Data Pool database. (NOTE: The BMGT determines which collections are eligible for export (from SDSRV) using the EcBmBMGTGroup.xml file in the /usr/ecs/<mode>/CUSTOM/cfg directory on x0acs0n, but the Bulk URL utility uses the bulkExportedFlag in DlCollections instead.) DAAC operations staff should set the bulkExportedFlag for each collection that is also exported by the BMGT, according to DAAC export policy. The bulkExportedFlag should then only be changed if the export policy changes.
  - The granule was inserted into the Data Pool during the time period specified by the run. This time period is either:
    - A time period ending when the Bulk URL utility begins its run, if the EcBmBulkURLConfigParams.xml file has doPreviousFlag = true; the time period can be in days (duration = day and count = <number of days>), or hours (duration = hour, and count = <number of hours>); or
    - The time period between a specified start and end date/time, if doPreviousFlag = false and <start Date> and <end Date> are specified in the EcBmBulkURLConfigParams.xml file.
  - The granule is not an "order only" granule, i.e., it was not placed in the Data Pool for the sole purpose of fulfilling a specific OMS order. If the granule is an "order only" granule, the isOrderOnly flag in the DlGranules table will have been turned on when the granule was inserted in the Data Pool.
    - This constraint was introduced with Release 7.0, when the Data Pool began to be used as a staging area for OMS orders. The intent of the Bulk URL utility is only to export Data Pool URLs for granules placed in the Data Pool by subscription or batch insert – these granules are intended to be available to the general Data Pool user community.

- **Delete** option - the Bulk URL utility exports an XML representation of a list of granuleURs for Data Pool granules, which meet all of the following conditions:
  - The granule is in a collection that is eligible for export (from Data Pool) to ECHO.
    - This is determined by whether the bulkExportedFlag = “Y” for that collection in the DlCollections table in the Data Pool database.
  - The granule will be deleted from the Data Pool in this (possibly two-part) run of the Cleanup utility (i.e., the granule is represented in the DlFilesToBeDeleted table in the Data Pool database).
    - In the one-pass process (echomode = “deleteall”), Cleanup deletes expired granules from the Data Pool and invokes the Bulk URL utility to notify ECHO of the deletions. If the one pass Cleanup process is used, it is possible for a granule to be deleted from the Data Pool before ECHO removes the granule from its inventory, thus making it possible for an ECHO user to receive an error when attempting to access an already-deleted granule.
    - In the two-pass process, during the first pass (echomode = “predelete”), Cleanup determines which granules are expired, invokes the Bulk URL utility to notify ECHO of their pending deletion, but does not actually delete the granules. After a time lag during which ECHO removes the to-be-deleted granules from its inventory, Cleanup is run in the second pass (echomode = “finishdelete”) and all of the to-be-deleted granules are actually removed from the Data Pool. The Bulk URL utility is not invoked during the second pass of the Cleanup utility.
  - The granule is not “order only.”
- For further information concerning the Data Pool Cleanup Utility refer to the procedure for **Running the Data Pool Cleanup Utility** (previous section of this lesson).
- The Bulk URL Utility runs and the ALOG file, **EcBmBulkURL.ALOG**, and debug log, **EcBmBulkURLDebug.log**, record errors, warnings, and information about utility events.
- The Bulk URL Utility does not perform automatic recovery operations. If the log file indicates that the utility failed to run to successful completion, the utility should be reinvoked.

Running the Most Recent Data Pool Inserts Utility

The **Most Recent Data Pool Insert Utility (EcDlMostRecentInsert)** lists the most recent additions to the Data Pool. The output of the utility is a set of files that a user could download and quickly inspect to identify recent additions to the Data Pool.

The utility takes in a date command-line parameter indicating the day of interest to the user. Files inserted into Data pool on the specified day are subsequently listed in the output files. If no date is provided, the utility uses the preceding day as a default with a time range of midnight to midnight.

The Most Recent Data Pool Insert Utility normally runs as a cron job. However, if it is necessary to run the utility from the command line it is possible to do so.

The procedure for running the Most Recent Data Pool Insert Utility is based on the following assumptions:

- Data Pool database server is running.
- Data Pool database is available.
- Stored procedures are present.

Running the Most Recent Data Pool Inserts Utility

---

1 Log in at the host where the Most Recent Data Pool Inserts Utility is installed (e.g., e0dps01, g0dps01, l0dps01, or n0dps01).

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

3 Type **EcDlMostRecentInsert.pl <*MODE*> [-insertDate <*YYYY/MM/DD*>]** and then press the **Return/Enter** key.

- **<*MODE*>** is the mode in which the utility is being executed (e.g., OPS, TS1, or TS2).
- **-insertDate** is an optional parameter specifying date of user's interest . If the date parameter is not specified, the preceding day's date is used as the default value.
  - For example, if today were July 11, 2005, the following command:

    EcDlMostRecentInsert.pl OPS

    would generate files concerning additions to the Data Pool between midnight July 9, 2005 and midnight July 10, 2005.

- The Most Recent Data Pool Inserts Utility runs and generates a set of files:
  - One file, named DPRecentInserts_<*YYYYMMDD*>, is located at the top-level Data Pool directory. It contains distinct GroupID, ShortName, and VersionID. For example, the file DPRecentInserts_20051102 in directory /datapool/OPS/user might contain the following types of entries:

    START_FILE: Entries:: 7
    GROUP_ID SHORT_NAME VERSION_ID
    MOAT MOD02QKM 077
    MOAT MOD15A2 077
    MOAT MYD01 077
    MOAT MYD02OBC 077
    MOAT MYD02QKM 077
    MOAT MYD35_L2 077
    TEST3 MOD35_L2 077
    END_FILE: Written 7

  - There is a file in each of the collection-level directories named DPRecentInserts_<*ShortName*>_<*VersionID*>_<*YYYYMMDD*>. The files contain SHORTNAME, VERSION_ID, and RELATIVE_PATH. For example, the file DPRecentInserts_MYD35_L2_077_20051102 in directory /datapool/OPS/user/MOAT/MYD35_L2.077 might contain the following types of entries:

    START_FILE: FileEntries:: 616
    SHORTNAME VERSION_ID RELATIVE_PATH
    MYD35_L2 077 /MYD35_L2.077/2001.07.01/labtest_2019099138
    MYD35_L2 077 /MYD35_L2.077/2001.07.01/labtest_2019099140
    MYD35_L2 077 /MYD35_L2.077/2001.07.01/labtest_2019099237
    MYD35_L2 077 /MYD35_L2.077/2001.07.01/labtest_2019099247
    MYD35_L2 077 /MYD35_L2.077/2001.07.01/labtest_2019099263
    […]

  - If it is unable to create a file at the top-level Data Pool directory, the Most Recent Data Pool Inserts Utility shuts down and logs an error message.
  - If it is unable to create a file at a collection level directory, the Most Recent Data Pool Inserts Utility logs an error message and continues processing other valid directories.
- When the Most Recent Data Pool Inserts Utility runs, it creates a log file, **EcDlMostRecentInsert.log**, which records errors, warnings, and information about utility events.
- The Most Recent Data Pool Inserts Utility does not perform automatic recovery operations. If there is an execution failure as a result of database server or system

shut down, rerun the script. This will create a new set of files (overwriting previous ones) listing additions to Data Pool for the specified insert date.

Running the Data Pool Collection-to-Group Remapping Utility

The **Data Pool Collection-to-Group Remapping Utility (EcDIRemap)** is a command-line utility interface that is used for reassigning a Data Pool collection to a collection group other than the one to which it was originally assigned.

The procedure for running the Data Pool Collection-to-Group Remapping Utility is based on the following assumptions:

- "Insert Enabled Flag" for the source collection has been turned off using the Data Pool Maintenance GUI
- The group to which the user is mapping the collection already exists in the Data Pool database.
- The group to which the user is mapping the collection is not the BRWS (browse) group.
- The collection to be remapped is not the Browse (Browse.001) collection.
- Data Pool database server is running.
- Data Pool database is available.
- Stored procedures are present.

There are several assumptions expected of the Data Pool Collection-to-Group Remapping Utility. The utility expects the exists of the collection in the Data Pool to which the user is mapping; it assumes that the browse collection is always located in the group "BRWS"; it also assumes the stored procedures are present. The Group Mapping utility runs only if the Data Pool database server is running and if the database is available.

Running the Data Pool Collection-to-Group Remapping Utility

---

1 Ensure that the "Insert Enabled Flag" for the source collection has been turned off.

- For detailed instructions refer to the **Modify an ECS Collection Using the DPM GUI** procedure (previous section of this lesson).

2 Log in at the host where the Data Pool Collection-to-Group Remapping Utility is installed (e.g., e0dps01, g0dps01, l0dps01, or n0dps01).

3 Type **cd /usr/ecs/<*MODE*>/CUSTOM/utilities**, then press the **Return/Enter** key.

4 Type **EcDlRemap.pl <*MODE*> -esdt <*name*> -version <*version*> -oldgrp <*old group*> -newgrp <*new group* >** and then press the **Return/Enter** key.

- <***MODE***> is the mode in which the utility is being executed (e.g., OPS, TS1, or TS2).
- <***name***> is the name of the source collection being remapped.
- <***version***> is the version of the source collection version being remapped.
- <***old group***> is the name of the collection group name that currently contains the collection.
- <***new group***> is the name of the collection group to which the source collection is being remapped.
- For example:

  EcDlRemap.pl OPS –esdt MOD29 –version 4 –oldgrp MOST –newgrp MOSS

  would remap collection MOD29, Version 4, (i.e., MOD29.004) from collection group MOST to collection group MOSS. The Data Pool database inventory would be updated to reflect the new location of the files.
- The Data Pool Collection-to-Group Remapping Utility runs and the log file, **EcDlRemap.log**, records errors, warnings, and information about utility events.
- The Data Pool Collection-to-Group Remapping Utility is able to recover from aborted runs by using the DlRecoveryParameters table to checkpoint its progress. In the event of an aborted run, the utility reads the recovery parameters table to determine at which point the utility left off when it aborted. This ensures that remappings that were taking place prior to the abort finish correctly. After recovery processing takes place, the utility processes the current run by acting on the latest input parameters.

Running the Data Pool QA Update Utility

The **QA Update Utility (EcDlQaUpdate)** provides the DAAC Operations Staff with a command-line interface for updating the QA data for granules in the Data Pool inventory. Both the inventory and the corresponding XML files on disk are updated.

The QA Update Utility can operate in either of two modes, depending on the command-line parameters:

- Single file – the utility reads an input file specified on the command line.
- Multiple files – the utility reads multiple files from the QAMUT's undo directory (e.g., /usr/ecs/OPS/CUSTOM/data/DSS/QAMUT/QAMUTUndo).

Each input file contains a list of granule db ids and the QA updates to be performed for those granules. The files, called "undo" files, are generated by the QAMUT utility in its "undo" directory. Therefore, the QA Update must be run after QAMUT has been run.

The procedure for running the QA Update Utility is based on the following assumptions:

- Data Pool database server is running.
- Data Pool database is available.
- Stored procedures are present.

Running the QA Update Utility

NOTE: The QA Update Utility is typically run after the QAMUT utility has run because output from the QAMUT is used as input to the QA Update Utility.

1 Log in at the host where the **QA Update Utility** is installed (e.g., e0acg11, g0acg01, l0acg02, or n0acg01).

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

3 Type **EcDlQaUpdateStart -mode <*MODE*> [-file <*filename*>]** and then press the **Return/Enter** key**.**

- <***MODE***> is the mode in which the utility is being executed (e.g., OPS, TS1, or TS2).
- <***filename***> is the name of a specific file to be used as input to the utility. The file must reside in the QAMUT undo directory (e.g. /usr/ecs/OPS/CUSTOM/data/DSS/QAMUT/QAMUTUndo). The **-file** parameter is optional. If no **–file** parameter is provided, the utility uses all undo files in the QAMUT undo directory as input.
- The QA Update Utility runs and the log file, **EcDlQaUpdate.log**, records errors, warnings, and information about utility events.
- The QA Update Utility is able to recover from an execution failure caused by situations such as system faults or database errors leaving all or some of the QA updates unprocessed. The utility detects such failures during the next run and continues processing the QA udpates that were left unprocessed in the previous run. The operator is given no choice as to recovery: either recovery proceeds or the Data Pool inventory and disk files are corrupted.

Running the Data Pool Move Collections Utility

The **Data Pool Move Collections Utility (EcDIMoveCollection)** provides the DAAC Operations Staff with a command-line interface to move collections from one file system to another. The file system move is implemented as a copy operation to the new collection

directory (destination), followed by removal of the old collection directory (origination) and its contents. The utility then establishes a link to the new location in place of the old directory. As a result, existing URLs are not invalidated by the move and no updated URLs need to be exported to ECHO. However, any existing URLs and file pointers become invalid from the time when the utility starts deleting the existing directories until the time the link is established. During that time the following errors could be encountered with respect to the collection being moved:

- A Data Pool ftp user or an EDG user clicking on a URL might experience a temporary error when trying to access files and directories. File transfers already in progress at the beginning of the deletion should complete normally.
- FTP Pull users could experience similar temporary problems trying to access links in FTP Pull directories that were established by the OMS and that point to granules in the moving collection.
- The Data Pool Web GUI returns an error if a user tries to access the collection via a bookmark. It flags the collection and does not display it as an active link on the collection drill down web page, thus temporarily preventing drill down access to the collection.
- The Data Pool insert service looks up the collection path in the Data Pool database during the insert process. The collection path is updated once the copy step is complete. Any Data Pool insert processes that looked up the copy path before it was updated insert their granules into the old directory location. If those granules are not copied but are then removed, they become phantoms and could trigger additional errors downstream, e.g., in distribution. Alternatively, if they were not removed, they would cause the move process to fail, as the Data Pool insert service would re-create the deleted directories.
- The OMS looks up granule file locations immediately before performing an FTP Push operation. If the lookup occurs just before the collection information in the Data Pool database is updated, but the copy operation starts after the file was deleted, the FTP Push operation fails and causes an operator intervention. Since the time window between file location look up and ftp push start is small, the chances for that occurring are very small. In such cases the operator would need to resubmit the request, and since the directory entry would now have been updated, the ftp push operation would succeed.

If the preceding impacts are not acceptable, operators can suspend inserts and web access for the original file system by marking it as "unavailable" in the DPM GUI. This would also halt staging operations for that file system in OMS.

The procedure for running the Data Pool Move Collections Utility is based on the following assumptions:

- Data Pool database server is running.
- Data Pool database is available.

- Stored procedures are present.

Running the Data Pool Move Collections Utility

1 Log in at the host where the Data Pool Move Collections Utility is installed (e.g., e0dps01, g0dps01, l0dps01, or n0dps01).

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

3 Type **EcDlMoveCollection.pl <*MODE*> -shortname <*shortname*> -versionid <*versionID*> -targetfs <*path*> [-verbose]** and then press the **Return/Enter** key.

- **<*MODE*>** is the mode in which the utility is being executed (e.g., OPS, TS1, or TS2).

NOTE: When the utility is run in OPS mode an operator prompt is displayed to prevent any accidental loss of data.

- **<*shortname*>** is the shortname of the collection being moved. It is a mandatory parameter.
- **<*versionID*>** is the version identifier of the collection being moved. Leading zeros must **not** be included. **–versionid** is a mandatory parameter.
- **<*path*>** is the relative path to the target file system (the file system to which the collection is being moved). Note that all Data Pool file systems must be mounted under the Data Pool root (e.g., /datapool/OPS/user). This parameter is mandatory.
- **-verbose** is an optional parameter. When **–verbose** is specified, some information is displayed on the screen and detailed information is written to the utility's log. Nonverbose is the default.
- For example:

  EcDlMoveCollection.pl OPS -shortname MODVOLC -versionid 1 -targetfs fs1 -verbose

  would move collection MODVOLC, Version 1, (i.e., MODVOLC.001) from its current directory as specified in the database to the new filesystem fs1. The collection would be moved from /datapool/OPS/user/MOAT to /datapool/OPS/user/fs1/MOAT. The utility would run using the verbose option, which would display information to screen and to the log.
- The Data Pool Move Collections Utility runs and the log file, **EcDlMoveCollection.log**, records errors, warnings, and information about utility events.
- The Data Pool Move Collections Utility is from an execution failure caused by situations such as system faults or database errors leaving all or some of the file moves unprocessed.
  - At startup the utility determines whether or not an execution failure has occurred. If there has been an execution failure, the utility prompts the operator

as to whether or not to attempt recovery. If the operator chooses to recover, the utility completes the processing of file moves that were left unprocessed in the previous run. Upon completion of the recovery, the utility runs again with the current command-line parameters.

– An operator may not wish to recover (e.g. if the target filesystem has become corrupted or full). In this case, recovery is not attempted, and the utility runs with the current command-line parameters. The moveFlag is automatically reset to "N". Any files that were copied to the target file system that experienced the failure would have to be deleted manually.

Running the Data Pool Hidden Scrambler Utility in Rename Mode

In Synergy V a new capability was introduced whereby the Data Pool Insert Service stores granules that are staged to the Data Pool for ordering purposes only in separate directories whose contents are not visible to anonymous ftp users. The new capability addresses the following two issues that were not resolved in Synergy IV:

- Order-only granules were accessible by the general public.
- Billable and restricted granules had to be processed in Synergy III mode to prevent accessibility by the general public.

Prior to Synergy V order-only granules (granules that are referenced by orders that are processed in Synergy IV mode) were visible to anonymous ftp users. However, making order-only granules publicly visible via the Data Pool ftp interface had the potential for confusing end users because it was not consistent with the DAACs' established Data Pool population policies for inserting granules into or removing granules from the Data Pool. The Synergy V capability renders invisible to Data Pool ftp users those granules that are staged to the Data Pool for ordering purposes only.

In Synergy IV orders that included billable or restricted granules had to be processed in Synergy III mode; otherwise, they would become visible to anonymous ftp users, which would subvert billing/restriction policies. The Synergy V capability makes it unnecessary to have a separate processing mode for billable and restricted granules because they are hidden in directories with encrypted names that cannot be accessed by anonymous ftp users. Consequently billable and restricted granules can be processed by OMS like any other order.

In Synergy V the Data Pool Insert Service puts order-only granules into special Data Pool directories that are reserved for this purpose. The directories are hidden from ftp users, and the granules in the hidden directories are not visible via Data Pool web drill-down. However, the directories are accessible by anonymous ftp via links in user FTP Pull directories to support FTP Pull orders.

When an order-only granule is subsequently inserted via a "normal" Data Pool insert, it becomes a normal Data Pool granule, and the Data Pool Insert Service moves it from the hidden directory to the appropriate place in the public directory structure. Of course, such transfers are not

allowed for billable or restricted granules. DAACs should not and generally do not insert granules from billable collections into the Data Pool, and the Data Pool Insert Service performs various checks (including examination of the DFA flag) to prevent the insert of granules that are flagged as “Hidden” or “Restricted” in ECS.

To decouple Data Pool and OMS file references, ordered granules are always represented in the hidden directory structure, either by files (if the granule is not in the public Data Pool) or by links pointing to the public files (if the granule is in the public Data Pool). Whenever a granule is converted from an order-only to a public Data Pool granule, its files are moved and links are left behind in the order only directories. Although a “metadata only” granule would be considered public, its science file would remain inaccessible (i.e., it would never be in the public Data Pool).

The hidden directory structure is below the FTP root because of the need to support FTP Pull access to ordered data. To hide the directories, the FTP root contains a directory that serves as the root for the hidden directory structure. (The directory is directly below the file system level, i.e., at the level of the Data Pool collection groups.) While it is impossible to hide that directory from view, it and all directories below it are configured in such a manner that their contents cannot be listed via ftp, in effect hiding all lower-level directories as well as their contents from public view.

The hidden directory structure mimics the public Data Pool directory structure (i.e., it is organized by collection group and collection); however, the hidden directory structure uses encrypted directory names so the pathnames cannot be guessed, preventing anonymous ftp users from switching into a hidden directory via the **cd** command. The Data Pool Hidden Scrambler Utility (EcDlHiddenScrambler.pl) allows the DAAC to re-encrypt directory names during system maintenance periods (i.e., during Data Pool down time), either on a regular basis or when intrusion is suspected.

Other Data Pool utilities (e.g., Data Pool collection move and re-map utilities) have been changed to the extent that they rely on the pre-Synergy V directory naming conventions so they can deal with granules in the hidden directory structure.

In Synergy V, OMS takes responsibility for removing order-only data from the Data Pool when they are no longer needed. However, the responsibility for cleaning up the public Data Pool remains with the Data Pool Cleanup Utility.

It is essential that the names of the hidden directory names do not become public knowledge. An external user could use knowledge of directory names and clever guessing of file names to download from them via anonymous ftp. The Data Pool cannot prevent this because it is necessary to allow ftp pull download from the directories via anonymous ftp. However, the Data Pool log analyzer will detect any attempt to access the hidden directories directly and will send an e-mail message to a DAAC-configured address to report security breaches. When that occurs, the DAAC should shut down ftp access to the Data Pool as soon as possible and run the Data Pool Hidden Scrambler Utility, which generates a new set of hidden directory names and updates the existing data holdings.

The Data Pool Collection-to-Group Remapping Utility will move the hidden collection directories when it moves the public collection directories. The Data Pool Move Collections Utility may also move the hidden directories for a collection depending on whether the order-only files are in the same file system as the public collection or are in a designated file system of their own.

The Data Pool Hidden Scrambler Utility (EcDlHiddenScrambler.pl) can be run in either of the following two modes:

- Transition.
- Rename.

In transition mode the utility generates hidden directory names and corresponding database entries for every collection defined for Data Pool in the affected operating mode. The transition mode can be used while Data Pool is up. The utility should be run in transition mode only once; i.e., the first time the utility is run in any given operating mode. Because transition mode is not used during normal operation, it is not described in any detail in this section.

In rename mode the utility re-encrypts all of the hidden directory names. This involves updates to the directory in the file system and to the database. Links from the FtpPull area (and elsewhere) are preserved. Re-encryption must be done during DAAC downtime only.

If the Data Pool Hidden Scrambler Utility is interrupted during execution, upon restart it detects failures from the previous run and continues processing the directories and files that were left unprocessed in the previous run. The operator is given no choice as to recovery. Recovery proceeds so that the Data Pool inventory and disk files will not be left in a corrupted state.

The procedure for running the Data Pool Hidden Scrambler Utility in rename mode starts with the assumption that the Data Pool is down for maintenance (no orders being processed, no external access to the Data Pool for downloading data, etc.).

WARNING

The Data Pool Hidden Scrambler Utility should be run in **transition** mode only once; i.e., the first time the utility is run in any given operating mode. In normal operations, the Data Pool Hidden Scrambler Utility is run in **rename** mode.

Running the Data Pool Hidden Scrambler Utility in Rename Mode

NOTE: In normal operations, the Data Pool Hidden Scrambler Utility is run in rename mode whenever hidden directory intrusion is detected/suspected. In addition, it is

recommended that the Data Pool Hidden Scrambler Utility be run in rename mode on a scheduled basis (e.g., monthly) at the DAAC's discretion.

1 Log in at the machine where the Data Pool Hidden Scrambler Utility is installed (e.g., e0dps01, g0dps01, l0dps01, or n0dps01).

- The script must be run from a user account with privileges to rename directories on the Data Pool.

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

- Change directory to the directory containing the Data Pool Hidden Scrambler Utility script (e.g., EcDlHiddenScramblerDataPool.pl).
- <***MODE***> will most likely be one of the following operating modes:
  - OPS (for normal operation).
  - TS1 (for SSI&T).
  - TS2 (new version checkout).
- Note that the separate subdirectories under /usr/ecs apply to different operating modes.

3 To perform a "rename" run, at the UNIX prompt enter:

EcDlHiddenScramblerDataPool.pl <*MODE*>

OR

EcDlHiddenScramblerDataPool.pl <*MODE*> -collgroup <*collgroup*>

OR

EcDlHiddenScramblerDataPool.pl <*MODE*> -shortname <*shortname*> -versionid <*versionid*>

- **<*collgroup*>** is a particular Data Pool collection group with collection directories to be renamed using the Hidden Scrambler Utility. If the –collgroup parameter is specified, the -shortname and –versionid parameters may not be used. If not all directories for collections within a collection group are to be renamed, run the Hidden Scrambler Utility using the –shortname and –versionid parameters to rename the directory for each affected collection.
- **<*shortname*>** is the name of a particular Data Pool collection, the directory for which is to be renamed using the Hidden Scrambler Utility. If the –shortname parameter is specified, the -versionid parameter must be specified too. If the –shortname parameter is specified, the -collgroup parameter may not be used.
- **<*versionid*>** is the version ID of a particular Data Pool collection, the directory for which is to be renamed using the Hidden Scrambler Utility. If the -versionid

parameter is specified, the –shortname parameter must be specified too. If the -versionid parameter is specified, the -collgroup parameter may not be used.

- The following examples show valid command-line entries for a "rename" run of the Hidden Scrambler Utility:

  EcDlHiddenScramblerDataPool.pl OPS

  – The Hidden Scrambler Utility performs rename processing (re-encrypts the hidden directory names) for all collection groups and all collections in the Data Pool in OPS mode.

  EcDlHiddenScramblerDataPool.pl OPS -collgroup MOAT

  – The Hidden Scrambler Utility performs rename processing (re-encrypts the hidden directory names) in OPS mode for the MOAT collection group, including all collections in the MOAT collection group.

  EcDlHiddenScramblerDataPool.pl OPS -shortname AST_L1B -versionid 3

  – The Hidden Scrambler Utility performs rename processing (re-encrypts the hidden directory names) in OPS mode for the AST_L1B.003 collection. [Note that the hidden directory name of the corresponding collection group (ASTT) would not be re-encrypted.]

- If applicable, usage errors (e.g., failure to specify a mode) are displayed on the terminal screen.

- The Data Pool Hidden Scrambler Utility records events and errors in the **EcDlHiddenScrambler.log** (in the /usr/ecs/<*MODE*>/CUSTOM/logs directory). If the log file exists already, the new information is automatically appended. If there is no existing log file named EcDlHiddenScrambler.log, the utility creates a new log file with that name.

- The Data Pool Hidden Scrambler Utility provides a capability to recover from interruptions caused by situations such as the system faults or database errors that leave all or some of the directories unprocessed. The utility detects such failure upon the next run and continues processing the directories and files that were left unprocessed in the previous run. The operator is given no choice as to recovery. Recovery proceeds so that the Data Pool inventory and disk files will not be left in a corrupted state.

This page intentionally left blank.

Tuning Data Pool and Order Manager Subsystem Configuration Parameters

When operating in Synergy IV mode the OMS has responsibility for most orders, i.e., for staging data from the archive into the Data Pool in preparation for distribution and for completing ftp pull and ftp push orders. The OMS handles orders for distribution on physical media and the data required for completing the physical media orders get staged to the Data Pool.

The following rules apply to the staging of data:

- If an ordered granule is in the Data Pool, it is considered staged.
- If an ordered granule is in AMASS cache, a request for inserting it into the Data Pool is sent immediately to the DPL insert service and the granule is copied from the cache into the Data Pool as quickly as possible via a separate Data Pool in-cache queue.
- If a distribution request references any additional granules that are not in either the Data Pool or AMASS cache, the request is eventually promoted into the "Staging" state and its granules are queued for Data Pool insert. The DPL insert service places the granules in the from-tape queue.

The following rules apply to order/request completion:

- An ftp pull order is considered complete when all its granules are in the Data Pool; however, the granules are retained in the Data Pool until the ftp pull order expires (according to the DAAC configured ftp pull retention time).
- For an ftp push order, the OMS queues an ftp push operation for each granule as soon as the granule is available in the Data Pool, provided the order has either reached the Staging state or requires no staging. The ftp push order is considered complete when all granules have been pushed.
- The OMS submits each physical media distribution request to the Production Module after all the relevant granules are in the Data Pool.

Provisions exist to account for failed granules (which no longer need to be staged), and for orders that are canceled.

To support orders/requests that cannot be processed in Synergy IV mode each DAAC configures DDIST and STMGT thread pools to regulate how the available archive bandwidth is made available to the orders that are being queued up (as described in the section on **Tuning Data Server Subsystem Parameters** in the **Data Distribution** lesson). In Synergy IV mode it is irrelevant whether a granule is in the read-only cache, the AMASS cache, or needs to be fetched from tape because OMS schedules archive operations in a different manner. As previously mentioned, granules that do not require actual tape access are processed in an expedited fashion.

The OMS and the DPL insert service allocate the archive resources (i.e., tape mounts) to the remaining granules using a set of staging policies.

Staging Policies

The tape archive is one of the critical limited resources for data distribution. Each archive has a limited number of tape drives and not all of them can be used for data distribution because some of them have to be reserved for other purposes (e.g., ingest). The most time-consuming operations in the archive include the mounting and dismounting of a tape and advancing the tape to the beginning of the requested file. Consequently, the OMS and Data Pool services implement a number of policies to optimize the use of archive resources. In some cases, a DAAC can influence the OMS behavior via tuning parameters; in other cases, the behavior is fully automatic and requires no tuning.

The staging policies have the following goals:

- Do not let a small set of distribution requests occupy a large number of archive tape drives for extended time periods.
- Adjust the pace of staging for a given device (or ftp connection) to slow down if the device or connection cannot keep up with the staging rate. (Do not let the staging rate significantly outstrip the pace of request completion.)
- Throttle the archive staging activity for output devices and ftp connections to prevent consuming a lot of disk space for orders that cannot be completed.
- Ensure that even low-priority requests move through the staging state at a reasonable pace.
- Ensure that high-priority requests are processed in an expedited fashion.
- Ensure that enough archive tape drives remain available for non-OMS/non-DPL-related activities.
- Manage Data Pool in-cache and from-tape insert processes efficiently.
  - Ensure that throughput keeps up with demand so granules are retrieved before AMASS removes them from its cache.
  - Ensure that additional insert processes can be dispatched in the eventuality that an archive tape that is mounted contains several requested granules.
- Optimize the use of tape archive resources as much as possible:
  - Use a single tape mount to read all granules that are currently on order and that reside on the same tape.
  - Do not let the drives of an archive become idle as long as there are granules from the archive on order that still need to be staged, regardless of where the corresponding order is in the distribution queue.

– If an order includes granules that are currently in an archive cache, copy the granules to the Data Pool as soon as possible to preclude the eventual need for a tape mount.

The paragraphs that follow describe the goals of the staging policies and the tuning parameters that can be modified to achieve those goals. Some of the tuning parameters are OMS parameters; others are Data Pool parameters. It is important to coordinate the settings for OMS and DPL parameters to facilitate achieving the goals of staging and to avoid conflicts between OMS and DPL.

Preventing a Set of Distribution Requests from Monopolizing Archive Tape Drives

A small set of distribution requests should not be allowed to occupy a large number of archive tape drives for extended time periods. The effect would be to block other distribution requests that also require tape mounts. Although it is unlikely that a single request would cause a very long delay (because of the limits the DAAC can configure for the maximum size of a distribution request before it is partitioned), if several large requests are dispatched simultaneously, they could keep the archive tape drives busy for many hours. This could happen, for example, if within a short time frame a single user submits a large number of orders for thousands of granules. When these requests work their way to the top of the distribution queue, they could block other orders.

The following tuning parameters can be used to prevent a set of distribution requests from occupying a large number of archive tape drives for extended time periods:

- **MaxTapeMountPerRequest** parameter on the **Data Pool Maintenance GUI** (**DPM GUI**).
- **Max Cheap Requests** parameter on the **OM GUI**.
- **Max Moderate Requests** parameter on the **OM GUI**.
- **Max Expensive Requests** parameter on the **OM GUI**.

MaxTapeMountPerRequest is a Data Pool configuration parameter that is set on the **List of Configuration Parameters** page of the **DPM GUI**. It is used to configure the maximum number of tape mounts that will be made for a single request.

Occasionally, a DAAC receives orders containing requests that require many, many tape mounts. The **MaxTapeMountPerRequest** parameter prevents such requests from occupying too many tape drives, which could keep some other requests from being serviced at all. However, while it is important to keep such requests from blocking all drives, it is also important to keep the requests moving through staging at a reasonable pace. For example, assuming three (3) minutes per tape mount, a request that requires 1,000 tape mounts and is allowed five (5) concurrent mounts at a time, would require at least 3*1,000/5 = 600 minutes (10 hours) to get through staging. If DAAC policy required even large requests to complete staging within about three (3) hours, it would be necessary to either permit up to 16 or 17 tape mounts for a single request (3*1,000/16 = 187.5 minutes or three hours, seven and a half minutes), or reduce the maximum number of granules in a request to a few hundred (e.g., 3*300/5 = 180 minutes).

Use the following guidelines to determine the appropriate value for the parameter:

- Initially set the value of the **MaxTapeMountPerRequest** parameter on the **DPM GUI** at five (5).
  - For detailed instructions on how to modify DPL parameter values refer to the **Modify DPM Configuration Parameter Values** procedure (previous section of this lesson).
- Based on DAAC experience and policy, adjust the setting (as necessary) to accommodate the number of tape mounts that are required for the largest requests that do not require partitioning.
  - Determine the average amount of time for a tape mount. If not known, assume three (3) minutes per tape mount.
  - Determine the maximum number of tape mounts for completing the largest request, assuming one tape mount per granule. The maximum number of granules in a request can be viewed as the value assigned to the **Max Request Granules** parameter on the **OM GUI** (e.g., 5,000).
  - Determine the maximum allowable time (in minutes) for completing the largest request [e.g., 720 minutes (12 hours)].
  - Calculate the maximum number of tape mounts as follows:
    - Multiply the average amount of time for a tape mount by the maximum number of tape mounts for completing the largest request and divide the product of that operation by the maximum allowable time for completing the largest request.
    - For example, three minutes for a tape mount times 5,000 tape mounts for the largest request equals 15,000 minutes, which is divided by 720 minutes (maximum allowable time for a request) to give an answer of 20.8 (i.e., 21) tape mounts.
  - If the maximum number of tape mounts is too large, consider getting permission to change some combination of the following values:
    - Maximum allowable time for completing the largest request.
    - Maximum number of granules in a request (**Max Request Granules** parameter on the **OM GUI**).

Max Cheap Requests, **Max Moderate Requests**, and **Max Expensive Requests** are OMS configuration parameters that are used for limiting the number of concurrent requests in a particular class that can be promoted to staging. The number of archive tape mounts the request needs in order to stage its granules is the basis for defining the classes. The OMS performs the classification when the request is initially validated. A request is reclassified if it is revalidated because it was resubmitted (e.g., following an operator intervention).

It is possible to configure three classes of request; i.e., Cheap, Moderate, and Expensive:

- Requests are classified as Expensive when the number of tapes they require exceeds the lower limit for Expensive requests as configured by DAAC operations (**Min Expensive Request** parameter on the **OM GUI**).
- Requests that are below this limit are classified Moderate if their tape mounts exceed the lower limit configured by DAAC operations for Moderate requests. (**Min Moderate Request** parameter on the **OM GUI**).
- All other requests are considered Cheap.

To limit the number of requests in each class that are in staging concurrently, as well as the total number of tape mounts these requests may require, the following three configuration parameters are configurable on the **OMS Server and Database Configuration** page of the **OM GUI**:

- **Max Cheap Requests -** maximum number of concurrent requests classified as Cheap that can be promoted to staging.
- **Max Moderate Requests -** maximum number of concurrent requests classified as Moderate that can be promoted to staging.
- **Max Expensive Requests -** maximum number of concurrent requests classified as Expensive that can be promoted to staging.

The parameters allow DAAC Operations to limit the number of requests in each class that are in staging concurrently, as well as the total number of tape mounts these requests may require. Thereby it is possible to ensure that expensive and moderate requests do not at times consume all available drives.

Use the following guidelines to determine the appropriate values for the parameters:

- Reserve the Expensive category for requests that will require several hours to get through staging because of their size even if they never had to yield tape mounts to lower-priority requests (e.g., several hundred tape mounts).
  - Set the request limit for this category (**Max Expensive Requests** parameter) to one (1).
    - For detailed instructions on how to modify OMS parameter values refer to the procedure for **Checking/Modifying OMS Server or Database Parameters** in the **Data Distribution** lesson.
- Do not limit the number of Cheap requests (**Max Cheap Requests** parameter).

- Reserve the Moderate category (**Max Moderate Requests** parameter) for requests that require more (but not significantly more) tape mounts than the configured limit per request.
  - Limit the number of these requests in staging such that they together with the Expensive request cannot consume more than 60 to 70 percent of the tape drives.
  - For example, if 60 drives are available for staging and the maximum number of concurrent tape mounts for a single request is set to five (5), limit the number of concurrent Moderate requests to seven (7) or fewer. Note that OMS will allocate the remaining tape drives to these requests (on a priority basis) if there is no competing workload from Cheap requests.
- If available, use a curve showing the distribution of requests by number of granules [excluding the requests known to get their granules from cache (e.g., subscriptions)].
  - Determine how many of these requests fall into the Moderate category.
  - Sum up the number of granules they require as the number of tape mounts they need (this assumes each granule requires a separate tape mount - where known otherwise, make appropriate adjustments).
  - Verify that the imposed limit will not make it impossible to achieve the average daily throughput.
  - For example, assume the Moderate requests require about 12,000 granules/tape mounts. A limit of seven (7) concurrent Moderate requests allows at most 35 concurrent tape mounts for Moderate requests [assuming the maximum number of concurrent tape mounts for a single request is set to five (5)], or at most 35*15 = 525 tape mounts/hour, or 12,600 granules in a 24-hour day. This is very close to the requirement and it may be advisable to either raise the number of concurrent Moderate requests permitted in staging, or raise the limit at which requests are classified as Moderate (**Min Moderate Request** parameter on the **OM GUI**).
- Adjust the parameters as necessary to accommodate the current user demand.

Adjusting the Pace of Staging

Simply servicing orders in the sequence in which they are submitted may result in poor utilization of media devices or ftp connections. For example, assume that the DAAC receives a large number of medium-sized orders to be distributed on 8mm tape and all of them need to have their data staged from the archive. Since the individual orders are not very large, the preceding limits would not prevent them from going into staging and subsequently keeping all the archive drives busy for some time. However, because 8mm tape drives are slow, most of the data would pile up in the Data Pool waiting for access to an 8mm tape drive. On the other hand, orders for other media types might get only sporadic service or no service and their output devices could eventually sit idle. The result is poor utilization of output devices. To prevent this type of

situation OMS has the following two types of tuning parameters for adjusting the pace of staging for a given device or ftp connection (to slow down if the device or connection cannot keep up with the staging rate):

- **RHWM** (Request High Water Mark) parameters on the **OM GUI**.
- **DHWM** (Data Volume High Water Mark) parameters on the **OM GUI**.

It is possible to set Request and Data Volume High Water Marks (RHWM, DHWM) for all physical media types and all ftp push destinations. The OMS uses the RHWM to limit the number of requests that are in progress (i.e., in staging or staged but not yet distributed) and DHWM to limit the amount of this data. The OMS can override the RHWM and DHWM settings for high-priority requests (see the description of low water marks in the subsequent section on **Ensuring That High-Priority Requests Are Expedited**). OMS ignores the RHWM/DHWM settings when there is no other work competing for archive resources.

The RHWM and DHWM settings have no effect on copying granules from the AMASS cache or retrieving granules that happen to reside on tapes mounted for other orders. However, when a request is promoted to the Staging state, all of its granules (including those fetched from AMASS cache) are counted when calculating the amount of work that is in progress. Note that the OMS continues to promote requests into staging until both the RHWM and the DHWM for that device/ftp destination have been reached.

RHWM and DHWM are OMS configuration parameters that are set using the **OM GUI**. RHWM and DHWM for physical media are configured on the **Media Configuration** page of the **OM GUI**. For ftp push destinations, the parameters are accessible on the **FTP Push Destination Details** page of the **OM GUI**.

The RHWM and DHWM parameters are used to prevent a peak in demand for some media type from temporarily blocking archive resources even though the data is staged faster than the device or ftp connection can distribute it. For example, assume that a group of users submitted several hundred orders for several thousand granules, all to be distributed via 8mm tape. Since the requests were submitted as a group, they would reach the head of the queue together and keep the archive busy for several hours with staging, even though it would be impossible to write the data to 8mm tape as fast as it was being staged. However, thanks to the RHWM and DHWM parameters this does not happen. Instead, the OMS limits the number of requests for 8mm tape that are in progress to the RHWM and DHWM settings. Once the assigned values are reached, OMS promotes a new request into staging only when a previous one completes.

More generally, the RHWM and DHWM are useful in preventing staging for an output device or ftp connection to occur much faster than the data can actually be distributed. The remaining archive bandwidth becomes available to service distribution requests for other media types or destinations.

RHWM and DHWM interfere with normal priority based on first-in-first-out dispatching of archive resources. It is necessary to determine when this is appropriate. For example, it may be desirable only if there is some abnormality in the distribution workload, such as a peak in order arrival for a particular device, or if an error occurs and some distribution channel operates at a degraded bandwidth. In such cases, RHWM and DHWM eventually limit the archive resources

to match the speed with which the staged data can be processed. In essence, RHWM and DHWM limit the amount of work the OMS buffers up for a given output channel. For a given media type or ftp destination, they must be larger than what needs to be kept "in work" to achieve the daily distribution targets. The recommendations that follow are based on the assumption that this is what is intended.

The DHWM for ftp pull requests is the point at which such requests are no longer eligible for promotion into the Staging state, effectively stopping the further influx of ftp pull data into the Data Pool. (However, granules staged for normal Data Pool use that also happen to be ordered for ftp pull do not count against the DHWM.)

The OMS removes ftp pull requests from its "books" during the day following their expiration. If ftp pull was stopped because the DHWM was reached, it resumes when enough requests have been removed. Of course, the actual removal of expired ftp pull granules from the Data Pool is done by the DPL cleanup utility; consequently, actual ftp pull usage at any particular moment can be higher or lower than the DHWM depending on when granule cleanup actually occurs.

The DHWM for ftp pull has no effect on copying granules for ftp pull from the AMASS cache. When an ftp pull request references granules that reside in the AMASS cache, the request continues to be serviced even if the ftp pull DHWM is reached. In addition, it is possible to configure a priority limit (**Min Pri to Preempt** parameter on the **OM GUI**) that causes the OMS to disregard the DHWM when processing requests with a priority that equals or exceeds the **Min Pri To Preempt** (see the description of **Min Pri to Preempt Parameter** in the subsequent section on **Ensuring That High-Priority Requests Are Expedited**).

Use the following guidelines to determine the appropriate values for the parameters:

- For a physical media type:
  - Set the RHWM to at least the number of available devices ("N") times two (2).
  - Set the DHWM to the media capacity times "N" times two (2).
  - If the distribution workload for the media type is much less than what the available devices can handle, use as "N" the number of devices actually needed to achieve the workload.
    - For detailed instructions on how to modify media parameter values refer to the procedure for **Checking/Modifying Media Parameters** in the **Data Distribution** lesson.

- For ftp push:
  - The archive will require a few minutes ("Ta") to stage data for a request from tape (assuming the request has sufficient priority and there are adequate archive resources). Add to this the average amount of time it takes to complete an ftp push request ("Tftp"). Set DHWM to at least the throughput (in MB) during Ta+Tftp that is required to achieve the workload target (alternatively, set it to several times that amount). For most ftp destinations, setting DHWM to the average 30-minute throughput (in MB) is adequate.
    - It ensures that DHWM does not throttle the distribution under normal circumstances, yet never queues up more than 30 minutes of ftp push work when there is either an unusual peak in the orders for that destination or a bandwidth problem.
    - For detailed instructions on how to modify ftp push parameter values refer to the procedure for **Checking/Modifying FTP Push Policy Configuration** in the **Data Distribution** lesson.
  - Set the RHWM to DHWM divided by the average request size (or set RHWM to the number of requests that typically need to be processed in a 30-minute interval).
  - Some ftp push destinations are connected to ECS via high-throughput networks. Unless such destinations receive large amounts of data from AMASS cache, they never have a significant ftp push queue because data is distributed as quickly as it can be staged. So DHWM and RHWM serve as throttles only when the connection experiences problems. Under normal circumstances, such destinations receive archive resources as quickly as their priority and the competing archive workload permit.
  - Ftp push requests that do not match any of the explicitly configured ftp push destinations are collected into a general ftp push group. The ftp hosts in the group vary from day to day, and the connection bandwidth to the hosts is generally unpredictable.
    - Set the RHWM to twice the number of ftp connections that the DAAC is willing to devote to these orders
    - Set the DHWM to the hourly amount of data that typically needs to be pushed for such orders. Ensure that the limit is several [e.g., five (5)] times larger than the configured maximum size for a single un-partitioned ftp push distribution request (so OMS can make full use of the configured number of connections).

- For ftp pull:
  - Start out by setting DHWM to the current size of the ftp pull area.
    - For detailed instructions on how to modify media parameter values refer to the procedure for **Checking/Modifying Media Parameters** in the **Data Distribution** lesson.
  - Adjust the ftp pull DHWM parameter as necessary to accommodate the current user demand for ftp pull.
    - For example, the DHWM could be set to the expected maximum daily ftp pull order volume times one day more than the number of days of ftp pull retention.
- Adjust the parameters based on experience.
  - For example, if it turns out that most of the orders are fairly large, the DHWM may need to be raised so it does not act as a throttle for normal distribution workloads.

Throttling Archive Staging for Output Devices and FTP Connections

Under normal circumstances the archive drives are the key distribution bottleneck. In many cases the output devices and ftp connections are able to distribute data as quickly as it can be staged. However, this can change if one of the output channels experiences problems; for example, if media drives fail or the throughput for some ftp connection suddenly deteriorates. If staging were to continue regardless of such problems, a lot of disk space might be consumed by orders that could not be completed and (consequently) could not have their data removed. At a minimum, it is desirable to throttle the archive staging activity for such devices or connections. The applicable tuning parameters are the same as those used in adjusting the pace of staging, specifically:

- **RHWM** (Request High Water Mark) parameters on the **OM GUI**.
- **DHWM** (Data Volume High Water Mark) parameters on the **OM GUI**.

Note that the OMS stops dispatching distribution requests that require resources that have been suspended. This behavior is automatic and there are no related tuning parameters apart from the retry behavior.

For additional information on RHWM/DHWM refer to the section on **Adjusting the Pace of Staging** (previous section of this lesson).

Ensuring the Staging of Low-Priority Requests at a Reasonable Pace

If the archive staging workload is close to the archive capacity for extended periods of time, requests that have a low priority could wait for a long time before being serviced. Once they are submitted to staging their tape-mount requests may be serviced infrequently and intermittently because higher-priority requests that get promoted into staging would be given preference. As a

result, low-priority requests may have to wait for a long time to get into the staging state and then stay in staging for a very long time. Eventually, a backlog of low-priority requests could build up and the response time would be very poor. Furthermore, once such a low-priority request got in staging, its data would accumulate in the Data Pool and could not be removed until the request completed. So it could end up blocking disk resources for an extended period of time.

To alleviate the problem of low-priority requests seeming to hang in Queued or Staging forever one can implement request aging, which is implemented through the following two types of aging parameters:

- OMS **Age Step** parameters on the **OM GUI**.
- OMS **Maximum Priority** parameters on the **OM GUI**.
- DPL **Aging Step** parameters on the **DPM GUI**.
- DPL **Max Priority** parameters on the **DPM GUI**.

OMS aging parameters (**Age Step** and **Maximum Priority**) cause OMS to increase the priority of a distribution request as it waits for promotion into the Staging state. This can help reduce the order completion time for low-priority requests.

DPL aging parameters raise the priority of requests that are in staging but have had long waits for tape mounts. By implementing DPL aging even low-priority requests can be made to move through the Staging state at a reasonable pace.

So OMS aging parameters are useful in increasing the priority of low-priority requests so the eventually get into the Staging state. DPL aging parameters raise the priority of requests so they get access to tape mounts and get out of the Staging state.

When request aging is in effect, OMS and DPL update request priorities regularly. The DAAC can enable or disable aging for each ECS priority level separately (e.g., aging may be in effect for LOW priority requests but not for any others); and the DAAC can set the hourly rate of priority increase and the maximum priority a request can achieve.

The OMS configuration parameters **Age Step** and **Maximum Priority** can be configured separately for each ECS priority level on the **Aging Parameters** page of the **OM GUI**. Similarly, the DPL **Aging Step** and **Max Priority** configuration parameters are configured separately for each ECS priority level on the **Aging Parameters** page of the **DPM GUI**.

Use the following guidelines to determine the appropriate values for the parameters:

- Settings are at the discretion of each DAAC; however, it is recommended that OMS request aging be turned off initially.
  - For detailed instructions on how to modify OMS aging parameter values refer to the procedure for **Checking/Modifying Aging Parameters** in the **Data Distribution** lesson.

- It is recommended that request aging be turned on for the DPL insert service for LOW and NORMAL priority requests to ensure that they complete staging promptly once they have started.
  - For example, a LOW or NORMAL priority request should attain the next higher ECS priority level after one or two hours of remaining in staging (but not go beyond that).
  - For detailed instructions on how to modify DPL aging parameter values refer to the **Modify DPM Aging Parameter Values** procedure (previous section of this lesson).
- Adjust the aging parameters as necessary to accommodate the current user demand.

Ensuring That High-Priority Requests Are Expedited

The mechanisms described in the preceding sections on **Adjusting the Pace of Staging** and **Throttling Archive Staging for Output Devices and FTP Connections** limit the number of requests that are submitted for staging from the archive. However, occasionally high-priority requests are received and should be processed in an expedited fashion. The following tuning parameters affect the expedited processing of high-priority requests:

- **RLWM** (Request Low Water Mark) parameters on the **OM GUI**.
- **DLWM** (Data Volume Low Water Mark) parameters on the **OM GUI**.
- **Min Pri to Preempt** parameter on the **OM GUI.**

It is possible to mark a point in the distribution queue where new requests that are queued below RLWM or DLWM values become eligible for preemptive dispatching, even if the requests/data in work are at RHWM/DHWM. An RLWM can be set for physical media only; a DLWM is available for physical media and each ftp push destination. Neither RLWM nor DLWM has much relevance to ftp pull, which has a special preemptive dispatching parameter, **Min Pri to Preempt**.

RLWM and DLWM are OMS configuration parameters that are set using the **OM GUI**. RLWM and DLWM for physical media are configured on the **Media Configuration** page of the **OM GUI**. For ftp push destinations, DLWM (only) can be configured on the **FTP Push Destination Details** page of the **OM GUI**.

The OMS and the DPL dispatch work in priority order and within the same priority on a first-in-first-out basis. This normally ensures that within a given output queue, higher-priority requests are serviced before lower-priority requests unless request aging is configured. However, if the output channel for some queue is fairly slow (e.g., 8mm tape or slow ftp connections), then the amount of work in progress usually will be at RHWM/DHWM. High-priority requests might be dispatched next as soon as other work completes, but this could be viewed as an unacceptable delay. It may be preferable for OMS to dispatch such requests in a preemptive manner, i.e., regardless of how much work is currently in progress (even if RHWM and DHWM have been reached or exceeded).

Use the following guidelines to determine the appropriate values for the parameters:

- No general recommendation is made; the settings for RLWM and DLWM are at the discretion of each DAAC.
  - A low setting for RLWM/DLWM requires most high-priority requests in the applicable output queue to wait for normal dispatching, i.e., until the work that is in progress drops below RHWM and DHWM.
    - That waiting time may be negligible if the output channel is fast.
  - If the amount of work in progress is hardly ever at RHWM and DHWM, configuring low watermarks is superfluous except to account for unusual circumstances.
    - For example, if some temporary device or connection problems caused a significant amount of data to be staged that is now waiting for ftp or transfer to a device (i.e., in a backlog situation).
  - For detailed instructions on how to modify ftp push parameter values refer to the procedure for **Checking/Modifying FTP Push Policy Configuration** in the **Data Distribution** lesson.
  - For detailed instructions on how to modify other media parameter values refer to the procedure for **Checking/Modifying Media Parameters** in the **Data Distribution** lesson.
- Adjust the RLWM/DLWM parameters as necessary to accommodate the current user demand.

Min Pri to Preempt is the preemptive dispatch priority for ftp pull requests. When an ftp pull request has the **Min Pri to Preempt** or a higher priority, it is dispatched even if the disk space currently consumed by unexpired ftp pull requests is at or above the HWM. So the **Min Pri to Preempt** parameter makes it possible to service high-priority ftp pull requests while lower-priority requests have to wait for disk space to become available.

Min Pri to Preempt is an OMS configuration parameter that is configured on the **OMS Server and Database Configuration** page of the **OM GUI**.

Use the following guidelines to determine the appropriate value for the parameter:

- No general recommendation is made; the setting for **Min Pri to Preempt** is at the discretion of each DAAC.
- Adjust the **Min Pri to Preempt** parameter as necessary to accommodate the current user demand for ftp pull.
  - For detailed instructions on how to modify OMS parameter values refer to the procedure for **Checking/Modifying OMS Server or Database Parameters** in the **Data Distribution** lesson.

Reserving Enough Tape Drives for Non-OMS/Non-DPL-Related Activities

The DAAC must be able to limit the number of tape drives that are made available for DPL staging in each archive. This ensures that there is a sufficient number of tape drives available for other, non-OMS/non-DPL-related activities. The following tuning parameters affect the number of available tape drives:

- MAX_READ_DRIVES_*x*0*xx*g*nn* (e.g., MAX_READ_DRIVES_e0acg11, MAX_READ_DRIVES_e0drg11, and MAX_READ_DRIVES_e0drg12) – parameters on the **List of Configuration Parameters** page of the **DPM GUI**.

Use the following guidelines to determine the appropriate values for the parameters:

- Based on historical data, estimate (for each archive) the total number of granules per day that are handled in the course of the following operations:
  - Ingest.
  - Synergy III-style (i.e., SDSRV-serviced) distribution (e.g., to handle billable granules, restricted granules, or secure ftp distributions) and that would not be expected to be in cache (e.g., EDG user orders).
  - Other archive accesses that are not managed by the OMS (such as ECS routine and on-demand processing and external subsetting).
- If the number of daily tape mounts a single drive can support is not known; assume that one tape drive is needed for every 250 granules ordered per day to support the workload.
- Use the estimated workload (total granules per day per archive) to estimate the number of archive tape drives that need to be set aside for the non-OMS/non-DPL-related activities.
- Allocate the remaining tape drives to DPL Staging by setting the **MAX_READ_DRIVES_*x*0*xx*g*nn*** parameter for each archive using the **DPM GUI**.
  - For detailed instructions on how to modify DPL configuration parameter values refer to the **Modify DPM Configuration Parameter Values** procedure (previous section of this lesson).
- Continue to monitor the ingest and non-OMS/DPL staging activities to verify that they receive neither too little nor too much archive bandwidth. Adjust the parameters as necessary.

Managing Data Pool In-Cache and From-Tape Insert Processes Efficiently

Managing Data Pool insert processes efficiently involves configuring the maximum number of concurrent Data Pool insert processes. The number of concurrent insert processes can be configured separately for data that is found in AMASS cache and data that needs to be read from archive tapes. The following tuning parameters affect the management of Data Pool insert processes:

- **NumOfAllowedCacheProcesses** parameter on the **DPM GUI**.
- **NumOfAllowedInsertProcesses** parameter on the **DPM GUI**.
- **NumOfAllowedNonCacheProcesses** parameter on the **DPM GUI**.

The **NumOfAllowedCacheProcesses** is the maximum number of insert processes that require access to cache (i.e., **in-cache** insert processes). The **NumOfAllowedNonCacheProcesses** is the maximum number of insert processes that require access to archive tapes (i.e., **from-tape** insert processes). The **NumOfAllowedInsertProcesses** is maximum number of insert processes running at any time (i.e., the total of **in-cache** insert processes and **from-tape** insert processes).

The number of **in-cache** insert processes must be limited to avoid overloading the Data Pool platform or the Data Pool database server. However, a sufficient number of in-cache insert processes must be available to ensure that throughput keeps up with demand so granules are retrieved before AMASS removes them from its cache (usually several hours).

A large enough number of **from-tape** insert processes must be provided so all the drives allocated for Data Pool staging can be kept busy (if there is sufficient distribution workload). Additional from-tape insert processes must be available so more insert processes can be dispatched in the eventuality that a tape that is mounted contains several requested granules. This is important in that it gives AMASS an opportunity to order the tape accesses by tape location.

NumOfAllowedCacheProcesses, **NumOfAllowedInsertProcesses**, and **NumOfAllowedNonCacheProcesses** are Data Pool configuration parameters that are configured on the **List of Configuration Parameters** page of the **DPM GUI**.

Use the following guidelines to determine the appropriate values for the parameters:

- Set the maximum number of **from-tape** insert processes (**NumOfAllowedNonCacheProcesses** parameter) to several (e.g., five) times the total number of tape drives allocated for Data Pool inserts.
  - The total number of tape drives allocated for Data Pool inserts is the sum of all the **MAX_READ_DRIVES_*x*0*xx*g*nn*** parameters for the various archives, as indicated on the **DPM GUI**.
  - For detailed instructions on how to modify DPL configuration parameter values refer to the **Modify DPM Configuration Parameter Values** procedure (previous section of this lesson).

- Set the number of **in-cache** insert processes (**NumOfAllowedCacheProcesses** parameter) as follows:
  - Determine the average number ("N") of granules per hour that an in-cache insert process can complete. This number depends on file sizes and checksum workload. (Testing during development indicated N = ~100.) A usable figure can be determined by dividing 3,600 by the average number of seconds needed to insert a granule into the Data Pool. The insert time for granules can be obtained from the Data Pool Action Driver (DPAD) log (EcDlActionDriver.ALOG on the Data Pool Server host, *x*0dps01).
  - Set the number of **in-cache** insert processes to the number of granules that need to be inserted per hour from AMASS cache divided by "N" and adjust the number upwards to leave room for peak demands (e.g., during catch-up situations).
    - The total number of granules expected to be inserted into the Data Pool from AMASS cache per hour equals the hourly average number of granules inserted into the Data Pool via subscriptions plus the hourly average number of granules ordered via subscriptions minus the estimated overlap between the two.
- Set the **NumOfAllowedInsertProcesses** to the sum of **NumOfAllowedCacheProcesses** and **NumOfAllowedNonCacheProcesses**.
- Adjust the settings based on experience.

Optimizing Tape Archive Resources

Since the tape archive is the most likely distribution bottleneck, it is desirable to optimize the use of this resource as much as possible. The following features optimize tape archive resources:

- Using a single tape mount to read all granules currently on order that reside on the same tape.
- Not letting the drives of an archive go idle as long as there are granules from the archive on order that still need to be staged, regardless of where the corresponding order is in the distribution queue.
- Copying ordered granules from archive cache to the Data Pool as soon as possible to preclude the eventual need for a tape mount.

There are no configuration parameters associated with the behaviors listed. They do not require tuning because they are automatic.

Summary

The OMS and the DPL offer a number of parameters that can be used to tune staging activity. With the exception of the number of tape drives available for staging, the parameters have no impact on archive throughput because the OMS and the DPL insert service always attempt to

maximize staging throughput. However, the parameters are useful for regulating archive activity across the different data distribution streams (i.e., by media type and ftp destination) and for preventing the kinds of situations mentioned in the preceding paragraphs. The parameters can be set such that they have no practical impact during normal operation, in which case archive tape capacity will be allocated strictly by priority and within the same priority on a first-come-first-served basis. Finally, the tuning parameters have no impact on distribution requests whose data is either already in the Data Pool or found in the AMASS cache. Such requests are processed in priority order as quickly as possible to reduce the chances of needing a tape access.

This page intentionally left blank.

Troubleshooting Data Pool Problems

Trouble Symptoms

Troubleshooting is a process of identifying the source of problems on the basis of observed trouble symptoms. Most Data Pool problems can be traced to some part of Data Pool. However, a common source of problems involves the reliance on messages or data from other subsystems. Like many other operational areas in the system, Data Pool has interfaces with other subsystems. Consequently, it is possible to trace some problems to another subsystem, including (but not necessarily limited to) those in the following list:

- Order Manager Subsystem (OMS).
- Spatial Subscription Server (SSS).
- Data Server Subsystem (DSS).
- Data Management Subsystem (DMS).

The general process of troubleshooting involves the following activities:

- Review the trouble symptoms.
- Check the status of relevant hosts/servers (as necessary).
- Check log files (as necessary).
- Take action to correct the problem(s).

The Data Pool maintenance modules use Perl scripts and Sybase database functions. Many of the scripts write to log files that can be checked if errors are detected when the scripts run. [Refer to the **Check Data Pool Log Files** procedure (subsequent section of this lesson).] Responses to many types of Data Pool problems are provided in Table 6, Responding to Data Pool Problems. If a tool cannot be launched, or does not function properly (e.g., cannot manage collection groups, does not update screens), you may need to ask the System Administrator to ensure that the appropriate scripts are functioning properly. It may be necessary to have the Database Administrator check to ensure that there are no problems with the database.

It is also possible to receive error messages when using the DPM GUI while it is apparently functioning normally. Error messages associated with the DPM GUI are listed in Appendix A of the *Operations Tools Manual* (Document 609-EMD-001, *Release 7.11 Operations Tools Manual for the EMD Project*). Table 7 is taken from the corresponding table in Document 609. If a problem cannot be identified and fixed without help within a reasonable period of time, the appropriate response is to call the help desk or submit a trouble ticket in accordance with site Problem Management policy.

Table 6. Responding to Data Pool Problems (1 of 11)

Symptom	Likely Cause(s)	Response
Compression: Compression of a granule fails even though compression is enabled for a given Shortname/VersionId and EcDlInsertUtility.log contains "DlInsertWorker.compress System level compression is disabled – compression will not occur for granule <*granule*>"	System level compression is not enabled.	On the **DPM GUI List of Configuration Parameters** page verify that system level compression is enabled (CompressOnInsert is ON) via the Maintenance GUI. [For detailed instructions refer to the **Enable/Disable Data Compression Using the DPM GUI** procedure (previous section of this lesson).]
Compression: EcDlInsertUtility.log contains "DlInsertWorker.compress: the expected compressed file <*compressedFileName*> does not exist"	Compression extension is not properly defined for the compression algorithm.	On the **DPM GUI OM Compression Algorithms** page determine whether or not the the compression extension is defined for the compression algorithm (e.g., /bin/compress has a compression extension of.Z and /bin/gzip has a compression extension of .gz). [For detailed instructions refer to the **View a List of Compression Algorithms Using the DPM GUI** procedure (previous section of this lesson).]
Compression: EcDlInsertUtility.log contains "DlInsertWorker.compress Compression not enabled for granule <*granule*>"	Not an error; indicates that granules of this ShortName/VersionId are not enabled for compression.	If granules of the specified type should be compressed, configure the collection for compresson. [For detailed instructions refer to the **Modify an ECS Collection Using the DPM GUI** procedure (previous section of this lesson).]

Table 6. Responding to Data Pool Problems (2 of 11)

Symptom	Likely Cause(s)	Response
Compression: EcDlInsertUtility.log contains "DlUtils.compressFile: could not compress file: <*compressCommand*> [OS STDERR:] <*stdErr*> [OS STDOUT:] <*stdOut*> "	Compression algorithm does not exist or is not executable.	1. On the command line type ***<compression command> <filename>*** (e.g., **/tools/bin/gzip -1 /home/cmshared/MISR_AM1_AS_AEROSOL_P015_O008407_F06_0013.hdf**) and then press the **Return/Enter** key. 2. Type **cd *path*** (e.g., **cd /home/cmshared**) and then press the **Return/Enter** key to change to the directory that should contain the compressed file. 3. Type **ls -al | grep *<compression extension>*** (e.g., **ls -al | grep gz**) and then press the **Return/Enter** key. 4. Observe the directory listing to identify whether there is a compressed file corresponding to the file that was to be compressed (e.g., **MISR_AM1_AS_AEROSOL_P015_O008407_F06_0013.hdf.gz**). 5. If there is no compressed file corresponding to the file that was to be compressed, call the help desk and submit a trouble ticket in accordance with site Problem Management policy.
Compression: EcDlInsertUtility.log contains "DlUtils.compressFile: could not compress file: <*compressCommand*> [OS STDERR:] <*stdErr*> [OS STDOUT:] <*stdOut*> "	File does not exist.	1. Type **cd *path*** (e.g., **cd /home/cmshared**) and then press the **Return/Enter** key to change to the directory that should contain the file to be compressed. 2. Type **ls -al | grep *<filename>*** (e.g., **ls -al | grep MISR_AM1_AS_AEROSOL_P015_O008407_F06_0013.hdf**) and then press the **Return/Enter** key. 3. Observe the directory listing to identify whether the file that was to be compressed is (e.g., **MISR_AM1_AS_AEROSOL_P015_O008407_F06_0013.hdf.gz**). 4. If there is no file corresponding to the file that was to be compressed, call the help desk and submit a trouble ticket in accordance with site Problem Management policy.

Table 6. Responding to Data Pool Problems (3 of 11)

Symptom	Likely Cause(s)	Response
Compression: EcDlInsertUtility.log contains "DlUtils.compressFile: could not compress file: <*compressCommand*> [OS STDERR:] <*stdErr*> [OS STDOUT:] <*stdOut*> "	Compression algorithm fails because the compressed file would be the same size or greater.	On the **DPM GUI OM Compression Algorithms** page define the algorithm to force compression (e.g., **/tools/bin/gzip -f %infile**) [For detailed instructions refer to the **View a List of Compression Algorithms Using the DPM GUI** procedure (previous section of this lesson).]
Deletions: ECS granule deletions are not propagating to the Data Pool	The format of entries in the GeoID file used as input to the Data Pool Cleanup Utility may be incorrect.	Check the format of entries in the GeoID file used as input to the Data Pool Cleanup Utility. [For detailed instructions refer to the **Check GeoID Files** procedure (subsequent section of this lesson).]

Table 6. Responding to Data Pool Problems (4 of 11)

Symptom	Likely Cause(s)	Response
Granule replacement does not occur in the Data Pool.	Granule replacement is not turned on.	**NOTE:** Ensure that granule replacement is turned on ONLY when needed because there are performance penalties associated with its use. 1. Check the value assigned to the GRANULE_REPLACEMENT parameter in the EcDlInsertUtlity.properties file (/usr/ecs/<*MODE*>/CUSTOM/cfg directory). 2. If the statement GRANULE_REPLACEMENT=Y is in the EcDlInsertUtlity.properties file, check for the following types of messages (indicating proper functioning) in the EcDlInsertUtility.log: **DlGranuleReplacementProcessor.replacementNeeded:** **DlGranuleReplacementProcesor.replacement needed: NO CANDIDATE FOUND** **DlGranuleReplacementProcessor.replacementNeeded: DlGranuleReplacementProcesor.replacement needed: REPLACEMENT CANDIDATE: 8365** **EcDlInsertUtility - exiting: CODE: 144 | STATUS: [exit code] 144 [message] RETRY [detail] GRANULE REPLACEMENT NEEDED** **[REPLACEMENT CANDIDATE DELETED: rangebeginningtime match 13:01:23]** **EcDlInsertUtility - exiting: CODE: 144 | STATUS: [exit code] 144 [message] RETRY [detail] GRANULE REPLACEMENT NEEDED** **[REPLACEMENT CANDIDATE DELETED: filename match l1bt.hdf]** [For detailed instructions refer to the **Check Data Pool Log Files** procedure (subsequent section of this lesson).] 3. If the statement GRANULE_REPLACEMENT=N is in the EcDlInsertUtlity.properties file, modify the statement to read “GRANULE_REPLACEMENT=Y”.

Table 6. Responding to Data Pool Problems (5 of 11)

Symptom	Likely Cause(s)	Response
Granule replacement does not occur in the Data Pool. (Cont.)	Granule replacement is not turned on. (Cont.)	4. After the Data Pool Insert Utility has run again check for the following types of messages (indicating proper functioning) in the EcDlInsertUtility.log: **DlGranuleReplacementProcessor.replacementNeeded:** **DlGranuleReplacementProcesor.replacement needed: NO CANDIDATE FOUND** **DlGranuleReplacementProcessor.replacementNeeded: DlGranuleReplacementProcesor.replacement needed: REPLACEMENT CANDIDATE: 8365** **EcDlInsertUtility - exiting: CODE: 144 | STATUS: [exit code] 144 [message] RETRY [detail] GRANULE REPLACEMENT NEEDED** **[REPLACEMENT CANDIDATE DELETED: rangebeginningtime match 13:01:23]** **EcDlInsertUtility - exiting: CODE: 144 | STATUS: [exit code] 144 [message] RETRY [detail] GRANULE REPLACEMENT NEEDED** **[REPLACEMENT CANDIDATE DELETED: filename match l1bt.hdf]** [For detailed instructions refer to the **Check Data Pool Log Files** procedure (subsequent section of this lesson).]

Table 6. Responding to Data Pool Problems (6 of 11)

Symptom	Likely Cause(s)	Response
Hanging request: Request is hanging in Staging status	Granule(s) of the request is (are) stuck in "Staging."	1. On the **Distribution Requests** page click on the request ID to bring up the **Distribution Request Detail** page. [For detailed instructions refer to the procedure for **Monitoring/Controlling Distribution Request Information on the OM GUI** in the Data Distribution lesson (625-EMD-009).] 2. Check the status of each individual granule in the request. (If one of them stays in "Staging," the whole request will remain in "Staging" until the granule finishes staging.) [For detailed instructions refer to the procedure for **Monitoring/Controlling Distribution Request Information on the OM GUI** in the Data Distribution lesson (625-EMD-009).] 3. If at least one of the granules in the request is still in "Staging," check the DPL DlActionDriver log and DlInsertUtility log to determine why the granule has not completed staging yet. [For detailed instructions refer to the **Check Data Pool Log Files** procedure (subsequent section of this lesson).]
Hanging request: Request is hanging in Staging status	Global Staging Status Parameter flag is suspended while the request is in the middle of staging.	Check the Global Staging Status parameter. (If the flag is suspended while the request is in the middle of staging, the request will stay in "Staging" until the suspension is lifted.) [For detailed instructions refer to the procedure for **Checking/Modifying Values Assigned to OMS Server or Database Parameters** in the Data Distribution lesson (625-EMD-009).]
Hanging request: Request is hanging in Staging status	Archive Sever queue is suspended while the request is in the middle of staging.	On the **OM Queue Status** page determine whether or not the Archive Server queue has been suspended. (If the archive is suspended while the request is in the middle of staging, the request will remain in that status until the suspension is lifted.) [For detailed instructions refer to the procedure for **Checking/Modifying OM Queue Status** in the Data Distribution lesson (625-EMD-009).]

Table 6. Responding to Data Pool Problems (7 of 11)

Symptom	Likely Cause(s)	Response
Hanging request: Request is hanging in Staging status	Request has not been handed off to DPL.	Check for the following types of messages in the Order Manager Debug log: **(04/09/04 11:35:02: Thread ID : 51 : ENTRY: OmSrDbInterface::SubmitToDPL2000663589** **exec OmInsDPLAction "3201075494","DVD",2000663589,150,"N","AST_L1B.001",NULL,NULL,NULL,NULL** **04/09/04 11:35:02: Thread ID : 51 : 0xfa81b0OmSrRequestGranule(RG:20000000006339-4)::SetState newState=1 oldState =0** **04/09/04 11:35:02: Thread ID : 51 : OmSrStagingPool(DVD)::CheckForPromotions candidates= 1** **04/09/04 11:35:02: Thread ID : 51 : OmSrStagingPool(DVD)::StageRequest(OmSrDR*) reqId=3201075494** **04/09/04 11:35:02: Thread ID : 51 : ENTRY: OmSrDbInterface::UpdateRequestCost RequestId=3201075494** **04/09/04 11:35:02: Thread ID : 51 : OmSrStagingPool(DVD):::STATS: UpdateWaterMarks Adding ReqId=3201075494 ReqSize=119.736 ReqPriority=NORMAL pending=0 active=1 RHWP=2 RHWM=100 DHWP=247.314 DHWM=1e+06** [For detailed instructions refer to the procedure for **Checking Log Files** in the Data Distribution lesson (625-EMD-009).]
Hanging request: Request is hanging in Staging status	DataPool Action Driver (DPAD) has not processed the request.	Check for the following types of messages in the EcDlActionDriver.log: **04/09/04 11:36:28: Thread ID : 21 : ENTRY: DlAdServer::TryDispatchQueuedActions** **04/09/04 11:36:28: Thread ID : 21 : Starting DPIU Process for dbId=2054841133 ShortName=AST_L1B RetryCount=0 basePrio=200 currPrio=200** [For detailed instructions refer to the **Check Data Pool Log Files** procedure (subsequent section of this lesson).]

Table 6. Responding to Data Pool Problems (8 of 11)

Symptom	Likely Cause(s)	Response
Hanging request: Request is hanging in Staging status	DPAD did not kick off the DlInsertUtility (DPIU) for the request.	Check for the following type of message in the EcDlInsertUtility.log: **04.09.2004 11:36:29.684][15257]PARSED COMMAND LINE: kwd: -shortname = AST_L1B | kwd: -mode = TS1 | kwd: -ppid = 29141 | kwd: -dbid = 2054841133 | kwd: -actionlist = (2054841133,1283) | kwd: -priority = 2 | kwd: -pid = 15257 | kwd: -retry = 0 | kwd: -incache = Y | kwd: -version = 3 | kwd: -mdonly = n | kwd: -period = 3 |** [For detailed instructions refer to the **Check Data Pool Log Files** procedure (subsequent section of this lesson).]
MISR Browse data is not being inserted in the Data Pool.	MISR special processing Is not turned on.	1. Check the value assigned to the MISR_SPECIAL_PROCESSING parameter in the EcDlInsertUtlity.properties file (/usr/ecs/<*MODE*>/CUSTOM/cfg directory). 2. If the statement MISR_SPECIAL_PROCESSING=Y is in the EcDlInsertUtlity.properties file, check for the following types of messages (indicating proper functioning) in the EcDlInsertUtility.log: **DlMisrProcessor.isDoubleInsertNeeded: returning MISBR.001 | true** **EcDlInsertUtility.exitInsert: *** STARTING THE DOUBLE INSERT/BROWSE PART ***** The VersionId (i.e., 001 in the preceding example) is the one that has been configured in the DlMisrProcessingCriteria database table (e.g., 1, 2, etc.) [For detailed instructions refer to the **Check Data Pool Log Files** procedure (subsequent section of this lesson).] 3. If the statement MISR_SPECIAL_PROCESSING=N is in the EcDlInsertUtlity.properties file, modify the statement to read “MISR_SPECIAL_PROCESSING=Y”. 4. After the Data Pool Insert Utility has run again check for the following types of messages (indicating proper functioning) in the EcDlInsertUtility.log: **DlMisrProcessor.isDoubleInsertNeeded: returning MISBR.001 | true** **EcDlInsertUtility.exitInsert: *** STARTING THE DOUBLE INSERT/BROWSE PART ***** [For detailed instructions refer to the **Check Data Pool Log Files** procedure (subsequent section of this lesson).]

Table 6. Responding to Data Pool Problems (9 of 11)

Symptom	Likely Cause(s)	Response
MISR Browse data is not being inserted in the Data Pool.	The Data Pool database table DlMisrProcessingCriteria has not been configured with the correct MISBR and level 1 and 2 science ESDTs that need special processing.	1. Check the entries in the DlMisrProcessingCriteria Data Pool database table to determine whether or not the table has been configured with the correct MISBR and level 1 and 2 science ESDTs that need special processing. [For detailed instructions refer to the procedure for **Using ISQL to Check Data Pool Database Tables** (subsequent section of this lesson).] 2. If the DlMisrProcessingCriteria Data Pool database table has not been configured with the correct MISBR and level 1 and 2 science ESDTs that need special processing, notify the database administrator to have the appropriate information added to the table.
MISR Browse data is not being inserted in the Data Pool.	The MISR collections have not been properly configured.	1. Use the DPM GUI to configure the MISR collections/collection groups. [For detailed instructions refer to the **Modify Collection Groups Using the DPM GUI**, **Modify an ECS Collection Using the DPM GUI**, and **Modify a NON-ECS Collection Using the DPM GUI** procedures (previous sections of this lesson).]
Most Recent Data Pool Inserts Utility does not execute properly.	Wrong date format was used when starting the script.	1. Check the log file (EcDlMostRecentInsert.log) for entries indicating problems with the date format. [For detailed instructions refer to the **Check Data Pool Log Files** procedure (subsequent section of this lesson).] 2. Try running the Most Recent Data Pool Inserts Utility (EcDlMostRecentInsert.pl) using the proper date format. [For detailed instructions refer to the **Running the Most Recent Data Pool Inserts Utility** procedure (previous section of this lesson).]
Most Recent Data Pool Inserts Utility does not execute properly.	User ID used to invoke the script did not have valid permissions.	Verify that the user ID used to start the script is appropriate for the task (e.g., user ID **allmode** or **cmshared** was used to start the script rather than a user ID like **jdoe**).
QA updates are not propagated to the Data Pool		Check for error messages in the EcDlQaUpdate.log.0. [For detailed instructions refer to the **Check Data Pool Log Files** procedure (subsequent section of this lesson).]

Table 6. Responding to Data Pool Problems (10 of 11)

Symptom	Likely Cause(s)	Response
Remapping: Data Pool Collection-to-Group Remapping Utility does not execute properly.	"Insert Enabled Flag" for source collection is on ("Insert Enabled Flag" for source collection must be turned off prior to execution).	1. On the **DPM GUI Modify Collection** page for the collection to be remapped verify that **Invalid for Data Pool** has been selected for the **Data Pool Insertion** parameter. [For detailed instructions refer to the **Modify an ECS Collection Using the DPM GUI** procedure (previous section of this lesson).] 2. Retry running the the Data Pool Collection-to-Group Remapping Utility. [For detailed instructions refer to the **Running the Data Pool Collection-to-Group Remapping Utility** procedure (previous section of this lesson).]
Remapping: Data Pool Collection-to-Group Remapping Utility does not execute properly.	Attempted to remap collection "Browse.001."	Collection "Browse.001" may not be remapped.
Remapping: Data Pool Collection-to-Group Remapping Utility does not execute properly.	Attempted to remap collection to BRWS group.	Collections may not be remapped to the BRWS group.
Remapping: Data Pool Collection-to-Group Remapping Utility does not execute properly.	Source collection does not exist in the source group.	1. On the **DPM GUI Collection Group Detail** page for the presumed collection group determine whether or not the collection to be remapped is in fact included in the collection group. [For detailed instructions refer to the **View Collection Group and Collection Information Using the DPM GUI** procedure (previous section of this lesson).] 2. If the collection to be remapped is in the specified collection group, check the log file (EcDlRemap.log) for entries indicating problems with the remapping. [For detailed instructions refer to the **Check Data Pool Log Files** procedure (subsequent section of this lesson).] 3. If the collection to be remapped is **not** in the presumed collection group, on the **DPM GUI Collection Groups** page determine which (if any) collection group includes the collection to be remapped (e.g., determine whether the collection has already been remapped). [For detailed instructions refer to the **View Collection Group and Collection Information Using the DPM GUI** procedure (previous section of this lesson).]

Table 6. Responding to Data Pool Problems (11 of 11)

Symptom	Likely Cause(s)	Response
Statistics: Data Pool Access Statistics Utility (DPASU) (EcDlRollupWebLogs .pl and EcDlRollupFtpLogs.pl) does not execute properly when running as a cron job.	The **-noprompt** argument was not used when the utility was run as a cron job.	1. Determine whether or not the **-noprompt** argument was used when the utility was run as a cron job. [For detailed instructions refer to the **Specify Data Pool Access Statistics Rollup Start Time and DPASU Execution with *cron*** procedure (previous section of this lesson).] 2. If **–noprompt** was not used, try running the Data Pool Access Statistics Utility (DPASU) (EcDlRollupWebLogs.pl or EcDlRollupFtpLogs.pl) manually. [For detailed instructions refer to the **Specify Data Pool Access Statistics Utility Execution from the Command Line** procedure (previous section of this lesson).] 3. If the utility executes properly when run manually, modify crontab entries to include the **–noprompt** argument. [For detailed instructions refer to the **Specify Data Pool Access Statistics Rollup Start Time and DPASU Execution with *cron*** procedure (previous section of this lesson).]
Statistics: Data Pool Access Statistics Utility (DPASU) (EcDlRollupWebLogs .pl and EcDlRollupFtpLogs.pl) does not execute properly when running from the command-line using the **-start** option.	Wrong date format was used when starting the script.	1. Check the appropriate log file (i.e.., EcDlRollupWebLogs.log or EcDlRollupFwFtpLogs.log) for entries indicating problems with the date format. [For detailed instructions refer to the **Check Data Pool Log Files** procedure (subsequent section of this lesson).] 2. Try running the Data Pool Access Statistics Utility (DPASU) (EcDlRollupWebLogs.pl or EcDlRollupFtpLogs.pl) using the proper date format. [For detailed instructions refer to the **Specify Data Pool Access Statistics Utility Execution from the Command Line** procedure (previous section of this lesson).]

Table 7. DPM GUI User Messages (1 of 2)

Message Text	Impact	Cause and Corrective Action
DB Error: You entered a duplicate collection group name that exists in the database. Please try again	Unable to add a new group id	Duplicate group name is entered. Check the list of group ids and enter a group name consisting of four letters, which is not on the list.
INPUT Error: You entered an invalid group name. Please see help page for more information. Please see section Add Collection Group	Unable to add a new group	Lower case letter is entered. Group id should be all Upper case letters.
INPUT Error: You entered an invalid name. Please see log for more details. Consult help tab and see section for NON-ECS add Collection Screen	Unable to add a new non-ECS collection	Special characters/small letters are entered. Non-ECS collection name should be in capital letters and without any special characters.
INPUT Error: You entered an invalid theme name. Please see help page for more information. Please see section: Add New Theme	Unable to add a new theme	Special characters/small letters are entered. Theme names should be in capital letters and without any special characters.
DB Error: Theme can not be null or empty	Unable to add a theme	A null or empty string is entered. Theme name should contain capital, small letters. Space is also allowed but no special characters.
DB Error: You entered either an existing theme name or a collection or a group name or an ESDT name. Check the log at /usr/ecs/<mode>/CUSTOM/log/EcDlDpmDataPoolGui.log for more details	Unable to add a theme	A name is entered, which is a duplicate name for a group, collection or an ESDT name.
DB Error: This collection is allowed for insertion therefore Spatial Search Type cannot be modified for this collection	Unable to modify Spatial Search type for a collection	Collection is not allowed for insertion. First make the collection allowed for insertion and then try to modify search type
DB Error: Error adding this collection. Collection entry <collection name> <version> already exist	Unable to add a collection.	Duplicate collection name entered. Verify the list of collection and then enter a name, which is unique.

Table 7. DPM GUI User Messages (2 of 2)

Message Text	Impact	Cause and Corrective Action
DB Error: Internal error occurred	A db transaction interrupted.	Database connection is lost for network error. No suggestion.
DB Error: delete failed because there are granules associated with this theme	Unable to delete a theme.	There are granules associated with this theme. Disassociate granules from this theme and then delete it.

Checking Data Pool Log Files

Log files can provide indications of the following types of problems:

- Communication problems.
- Database problems.
- Lack of disk space.

The procedure for checking Data Pool log files starts with the assumption that the operator has logged in to the system and the appropriate host.

Check Data Pool Log Files

1 Access a terminal window logged in to the appropriate host.

- APC Server (e.g., e0acg11, g0acg01, l0acg02, or n0acg01) host has the following types of Data Pool log files:
  - EcDlPopulateStatTables.log.
  - EcDlQuickServer.ALOG.
  - EcDlQuickServerDebug.log.
  - EcDlQaUpdate.log.0.
    - Log configuration is controlled in EcDlQaUpdate.properties file (/usr/ecs/<*MODE*>/CUSTOM/cfg directory).
    - Set LOG_SIZE = 10M (at least) or debugging gets confusing.
    - Set LOG_DETAIL_LEVEL= DETAILED for debugging only (generates lots of output) in cfg file (file). Normally leave set at FINER.

- FSMS Server (e.g., e0drg11, g0drg01, l0drg01, or n0drg01) host has the following types of Data Pool log files:
  - EcDlCheckSumServer.ALOG.
  - EcDlCheckSumServerDebug.log.
  - EcDlQuickServer.ALOG.
  - EcDlQuickServerDebug.log.
- Data Pool Server (e.g., e0dps01, g0dps01, l0dps01, or n0dps01) host has the following types of Data Pool log files:
  - EcDlActionDriver.ALOG.
  - EcDlActionDriverDebug.log.
  - EcDlBatchInsert.log.
  - EcDlCleanup.log.
  - EcDlDpmDataPoolGUI.log.
  - EcDlHiddenScrambler.log.
  - EcDlInsertUtility.log.
  - EcDlM2XT.ALOG.
  - EcDlM2XTDebug.log.
  - EcDlMostRecentInsert.log.
  - EcDlMoveCollection.log.
  - EcDlRollupFwFtpLogs.log.
  - EcDlWebaccess.log.
  - EcDlWebaccess.DEBUGLOG.
- For detailed instructions refer to the procedure for **Logging in to System Hosts** (previous section of this lesson).

2 Type **cd /usr/ecs/*MODE*/CUSTOM/logs** then press **Return/Enter**.

- Change directory to the directory containing the Data Pool log files (e.g., EcDlActionDriver.ALOG, EcDlInsertUtility.log).

3 Type **pg *filename*** then press **Return/Enter**.

- ***filename*** refers to the Data Pool log file to be reviewed (e.g., EcDlActionDriver.ALOG, EcDlInsertUtility.log).
- The first page of the log file is displayed.

- Although this procedure has been written for the **pg** command, any UNIX editor or visualizing command (e.g., **vi**, **view, more**) can be used to review the log file.

4 Review the log file to identify problems that have occurred.

5 Respond to problems as follows:

- Data Pool process-related problems.
  - For DPM GUI problems, refer to Table 7, DPM GUI User Messages.
  - For other types of Data Pool problems, refer to Table 6, Responding to Data Pool Problems.
- Communication problems.
  - Notify the Operations Controller/System Administrator of suspected communication problems.
- Database problems.
  - Verify that relevant database servers are running.
  - Check for lack of (or corruption of) data in the database using either a database browser or isql commands.
  - Notify the Database Administrator of suspected database problems.
- Lack of disk space.
  - Remove unnecessary files.
  - Notify the Operations Controller/System Administrator of recurring disk space problems.

Checking GeoID Files

GeoID files are input files for scripts used in deleting (or undeleting) ECS granules from the archive and/or Data Pool. A GeoID file consists of a list of GeoIDs for granules that are to be deleted (or undeleted).

The procedure for checking GeoID files starts with the assumption that the operator has logged in to the system and the appropriate host.

Check GeoID Files

1 Access a terminal window logged in to the Sun Consolidation Internal Server host (e.g., e0acs11, g0acs11, l0acs03, or n0acs04).

- For detailed instructions refer to the procedure for **Logging in to System Hosts** (previous section of this lesson).

2 Type **cd path** then press **Return/Enter**.

- For example:

  cd /usr/ecs/OPS/CUSTOM/utilities

- Change directory to the directory containing the GeoID files (e.g., SYN5_GranDel.geoid, Dec_23_AllDatatypes.geoid) used with the relevant script (e.g., EcDsBulkDelete.pl, EcDsBulkUndelete.pl, EcDsDeletionCleanup.pl).

3 Type **vi** ***filename*** then press **Return/Enter**.

- ***filename*** refers to the GeoID file to be reviewed (e.g., SYN5_GranDel.geoid, Dec_23_AllDatatypes.geoid).
- For example:

  x0acs11{cmops}[10]->vi SYN5_GranDel.geoid
  SC:AST_04.001:2020289624
  SC:AST_04.001:2020289642
  SC:AST_04.001:2020289660
  SC:AST_04.001:2020289678
  SC:AST_04.001:2020289696
  SC:AST_04.001:2020289714
  SC:AST_04.001:2020289732
  SC:AST_04.001:2020289750
  SC:AST_04.001:2020289770
  SC:AST_04.001:2020289786
  SC:AST_04.001:2020292170
  […]
  SC:AST_04.001:2020292219
  SC:AST_04.001:2020292263
  "SYN5_GranDel.geoid" 9740 lines, 255893 characters

  - Many blank lines have been deleted from the example.

- Although this procedure has been written for the **vi** editor, any UNIX editor can be used to create the file.

4 Review the file entries to identify problems that have occurred.

- The GeoID file must contain GeoIDs in the format ***<BaseType>*:*<ESDT_ShortName.VersionID>*:*<databaseID>***.
- For example:

  SC:AST_L1BT.001:5850

  - The GeoID in the example indicates BaseType SC (science granule), ShortName AST_L1BT (ASTER Level 1B thermal infrared data) VersionID 001, and databaseID 5850.
- There may be no spaces or blank lines in the file.

5 Use UNIX editor commands to fix problems detected in the file.

- The following vi editor commands are useful:
  - **h** (move cursor left).
  - **j** (move cursor down).
  - **k** (move cursor up).
  - **l** (move cursor right).
  - **a** (append text).
  - **i** (insert text).
  - **r** (replace single character).
  - **x** (delete a character).
  - **dw** (delete a word).
  - **dd** (delete a line).
  - ***n*dd** (delete *n* lines).
  - **u** (undo previous change).
  - **Esc** (switch to command mode).

6 Press the **Esc** key.

7 Type **ZZ**.

- **vi** exits and the new or edited file is saved.
  - To exit **vi** without saving the new entries in the file type **:q!** then press **Return/Enter**.

- UNIX command line prompt is displayed.

Using ISQL to Check Data Pool Database Tables

The DPL database is the repository of data concerning granules in the Data Pool. Depending on circumstances at a particular site it may be necessary to request that the Database Administrator check and/or modify parameters in the Data Pool database. The procedures that follow are provided to assist operators who can check entries in database tables themselves.

The data stored in databases can be checked using either a database browser or isql commands. The procedure in this section describes how to check the tables using isql commands.

The procedure for using isql to check database tables starts with the assumption that the operator has logged in to the system.

Using ISQL to Check Data Pool Database Tables

1 Access a terminal window logged in to the Access/Process Coordinators (APC) Server host.

- Examples of APC Server host names include **e0acg11**, **g0acg01**, **l0acg02,** and **n0acg01.**
- For detailed instructions refer to the procedure for **Logging in to System Hosts** (preceding section of this lesson).

2 Type **isql –U*userID* -S*DBServer*** then press **Return/Enter**.

- For example:

  isql –Udpl_role –Sx0acg01_srvr

3 At the **Password:** prompt type ***dbpassword*** then press **Return/Enter**.

- The ***dbpassword*** is the password for logging in to the database using the specified ***userID***.

4 Type **use *dbname*** at the **1>** prompt then press **Return/Enter**.

- The ***dbname*** is likely to be one of the following names:
  - **DataPool** [OPS mode].
  - **DataPool_TS1** [TS1 mode].
  - **DataPool_TS2** [TS2 mode].

5 Type **go** at the **2>** prompt then press **Return/Enter**.

6 Type **select * from *TableName*** at the **1>** prompt then press **Return/Enter**.

- For example:

  select * from DlMisrProcessingCriteria

- Alternatively, type **select *columnName* from *TableName*** at the **1>** prompt then press **Return/Enter**.
  - For example:

    select ShortName from DlMisrProcessingCriteria

- Another alternative is to type **select *columnName1,columnName2[,columnName3,…]* from *TableName*** at the **1>** prompt then press **Return/Enter**.
  - For example:

    select ShortName,VersionId from DlMisrProcessingCriteria

7 Type **go** at the **2>** prompt then press **Return/Enter**.

- Table contents are displayed.
  - If * (wildcard) was specified, all entries in the table are displayed.
  - If specific columnNames were entered, the data associated with those columns only are displayed.

- For example:

```
1> select * from DlMisrProcessingCriteria
2> go
 ShortName VersionId type
 --------- --------- ----
 MI1B2E          2  1
 MI1B2T          2  1
 MIL2ASAE        2  2
 MIL2ASLS        2  2
 MIL2TCAL        2  2
 MIL2TCCL        2  2
 MIL2TCCL        3  2
 MIL2TCST        2  2
 MIL3SAE         2  3
 MISBR           5  B

(10 rows affected)
```

8 To exit from isql type **quit** at the **1>** prompt then press **Return/Enter**.

Notes Concerning the Data Pool Database

The Data Pool database (DPL DB) contains a wealth of information, much of which can be seen on the GUIs. In those rare cases when useful information cannot be obtained from the GUIs it can be useful to use SQL queries.

Warning

Do not update the Data Pool database using SQL commands because it is very easy to get the database out of sync.

- The OmRequest table has CreationDate (time of order creation), MediaTypeId, Source (e.g., MTMGW), UserId (ECS User Profile) and RequestId (requestId from MSS EcAcRequest).

The Data Pool database has two tables that are of particular use in debugging:

- **DlInsertActionQueue** has all new insert actions by ecsId (SDSRV dbID). The status and statusDetail columns are of interest.
- **DlOMSGranules** has all the information on granules that DPL is staging for OMS. Look for granuleId (DPL granuleId) and owner (a value of "O" indicates that OMS still needs the granule for an order and a value of "C" means that the granule is being cleaned up).

When checking for system-wide problems (rather than a fault concerning a particular request) it may be useful to check the states of all requests from a particular Source (e.g., MTMGW or V0GW) in a specified time frame.

- For example:

  select count(1), a.requestStatus from mss_acct_db_TS1..EcAcRequest a where a.requestId in (select RequestId from EcOmDb_TS1..OmRequest where Source like "V0GW%" and CreationDate between "04/08/2004 10:00" and "04/08/2004 11:00" and RequestId = a.requestId) group by requestStatus

- Note that you can use Source = "MTMGW" for Machine to Machine Gateway orders. If you see a large number of requests in either the transferring state or failed state, there is probably a system-wide problem.
- The DB user who issues the command in the preceding example must have read permission in the MSS DB as well as access to the OMS and DPL DBs.

When checking the database for systemic issues involving the failure of a large number of orders (e.g., all AST_L1B.003 requests since 2:00 PM on 04/09/2004 are failing) , a query similar to the following one may be useful:

select status, statusDetail, enqueueTime from DataPool_TS1..DlInsertActionQueue a where a.ecsId in (select ECS_GranuleId from EcOmDb_TS1..OmGranule where ECS_GranuleId = a.ecsId and EsdtType = "AST_L1B.003") and enqueueTime > "04/09/2004 14:00" order by enqueueTime desc

- Successful inserts would have "COMPLETE" in the status column and "Science granule inserted" in the statusDetail column. If there were errors, the status would be "FAILED" and the statusDetail column would contain the error.
- If there were consistent failures of all AST_L1B.003 requests and the statusDetail column indicated that there was no collection mapped to type AST_L1B.003, the thing to do would be to set up AST_L1B.003 in the Data Pool (using the DPM GUI) and the requests could be processed.
- The DB user who issues the command in the preceding example must have read permission in the MSS DB as well as access to the OMS and DPL DBs.

Open Geospatial Consortium (OGC) Web Services (OWS) Support

Deployment of Open Geospatial Consortium (OGC) Web Services (DOWS)

For Synergy V the Deployment of Open Geospatial Consortium (OGC) Web Services (DOWS) allows OGC-compliant clients the ability to search, map, download, and subset (spatially and temporally) data residing in the EOS Data Pools. Consequently, a configurable subset of the EOS Data Pools' inventory can be accessed via the following OGC Web Services (OWS):

- Web Coverage Service (WCS).
- Web Mapping Service (WMS).

For Synergy V DOWS is implemented at the LP DAAC only.

The following activities support the access of WCS/WMS to data in the EOS Data Pools:

- Data enabling.
- Data conversion.
- Data exposure.
- Data publishing.

Data enabling involves the definition and population of OWS-enabled inventory in the Data Pool. The following mechanisms support OWS data enabling:

- The Data Pool Maintenance GUI "add theme" capability is used to define OWS-specific themes.
- Spatial Subscription Server subscriptions for Data Pool inserts associated with themes provide the method by which data are put in the Data Pool to populate the OWS themes.
  - Data inserted into the Data Pool inventory (via the Data Pool database) is tagged with the theme.
  - The tagging renders that piece of inventory OWS-enabled.

Using the Data Pool themes mechanism makes it possible to perform the following functions:

- Earmark arbitrary collections of granules for a particular instance of a web service in the form of an OWS theme.
- Easily create or eliminate OWS themes.

- Easily add or remove granules from OWS themes.
- Define OWS themes using bounding metadata. For example, defining an OWS theme by a certain spatial or temporal constraint.

Data conversion is necessary because the inventory of EOS data in each DAAC's Data Pool is stored in HDF-EOS format, a format that is not widely used in the OGC community. That community's preferred format is GeoTIFF.

Although it is possible to convert the data from HDF-EOS to GeoTIFF as it is accessed, the amount of time required to do so is not always compatible with a synchronous service. However, the performance of the service can be enhanced by pre-converting as much data to GeoTIFF format as possible. But there are resource constraints at operational DAACs; so it is necessary to strike a balance between storage and performance limits. The solution is to pre-convert a useful subset of the OWS-enabled inventory to GeoTIFF format. The remaining OWS-enabled data that are not pre-converted to GeoTIFF are subject to on-the-fly conversion by third-party WCS/WMS applications (when requested).

The pre-conversion of a single HDF-EOS file results in the generation of one or more GeoTIFF files. The GeoTIFF files are placed in the same directory as the parent HDF-EOS file within the Data Pool cache. Information concerning the physical location of all GeoTIFFs is stored in the Data Pool database.

When a piece of native (HDF-EOS) Data Pool inventory expires, both it and its child GeoTIFF file(s) are cleaned up when the Data Pool Cleanup utility runs.

Only a subset of the GeoTIFFs that can be created from a single HDF-EOS file are of relevance to the OWS applications. That subset will be the only GeoTIFF data stored in the Data Pool for a given HDF-EOS file.

The DAAC determines (for each relevant ESDT) what subset (if any) of the GeoTIFFs can be created from a single HDF-EOS file. That information is reflected in the configuration of the DOWS.

Pre-conversion is executed on a regular (e.g., daily) basis in conjunction with data exposure/synchronization.

Data exposure is handled primarily by third-party (e.g., IONIC) applications. In order to expose OWS-enabled data to the end user via WCS/WMS applications the data must be "indexed," which involves the replication of metadata pertaining to the OWS-enabled inventory from the Data Pool database to the WCS/WMS applications' database (via the IONIC Indexer API). Due to the volatility of the Data Pool inventory, it must be possible to add and remove OWS-enabled inventory from the WCS/WMS application's database.

The extent to which metadata is replicated depends on the number of searchable criteria to be exposed to the end user. The following examples illustrate the types of metadata to be conveyed:

- Physical location of data.
- Geospatial extent.
- Temporal extent.

Additional metadata should be indexed per piece of OWS-enabled inventory if necessary. At least the following metadata would be required:

- ECS Granule Id.
- Cloud Cover.

Data exposure involves maintaining a subset of the Data Pool database that is current. Any changes in the subset of OWS-enabled inventory within the Data Pool must be propagated to the indexer. This aspect of data exposure is called "synchronization."

Synchronization is executed on a regular (e.g., daily) basis due to the potential volatility of the Data Pool inventory. The process involves the synchronization of both WCS and WMS inventory between the Data Pool database and the indexer. Correctly exposed inventory is tagged in the Data Pool database so that it can be "published" via the data publishing process.

Data publishing means that all OWS-enabled data that has been successfully exposed (in the data exposure process) subsequently has its metadata exported to the ECHO system, which serves as the OGC catalog for Data Pool holdings. An OGC client (e.g., the Earth Science Gateway web client) could search for data offered by WMS and WCS via interrogation of an OGC catalogue service associated with ECHO.

The Data Pool database, the Bulk Metadata Generation Tool (BMGT) utility, OWS Binding Utility (OBU), and Bulk URL Utility are involved in ensuring that ECHO has up-to-date service information:

- The Data Pool database supports parameters that indicate what inventory is to be offered by the WMS and/or WCS and what inventory has been successfully exposed via the WMS/WCS services.
- To support the OWS-related data publishing capabilities of Synergy V a new BMGT utility configuration file has been created to make BMGT export relevant **collection-level** metadata about OWS-enabled granules to ECHO. (The impact of BMGT changes on the operator interface for Synergy V is limited to the configuration files.)
- The OBU detects which granules have been successfully indexed by the OWS Inventory Synchronization Utility for WMS/WCS and exports the relevant **granule-level** service metadata to ECHO.

- The Bulk URL Utility configuration file has been modified so the utility propagates the deletion of OWS-enabled granules from the Data Pool to ECHO. (As is the case with BMGT, the Bulk URL changes for Synergy V are in the configuration file only.)

Working with DOWS Scripts

There are several DOWS scripts that provide the operator with utilities or applications for managing DOWS maintenance. These include:

- DOWS Synchronization Utility: a utility for propagating changes in the OWS-enabled inventory in the Data Pool to the IONIC Indexer on a regular (e.g., daily) basis to make the data visible as WCS and/or WMS services to any OGC compliant web client [especially the Earth Science Gateway (ESG) OGC client]. The Synchronization Utility can be invoked from the command line if necessary, especially when operating on data from a time range prior to the regularly scheduled period.
- DOWS GeoTIFF Converter Utility: a utility for converting (i.e., "pre-converting") HDF-EOS files to GeoTIFF format on a regular (e.g., daily) basis. It can be invoked from the command line if necessary, especially when operating on data from a time range prior to the regularly scheduled period.
- OWS Binding Utility (OBU): a utility for exporting to ECHO granule-level metadata concerning OWS-enabled data in the Data Pool. The information indicates which granules from the OWS-enabled collections have been successfully indexed for WCS or WMS (or both) web service(s) by the DOWS Synchronization Utility. (The BMGT would have previously exported the collection-level metadata to ECHO.) The OBU can be invoked from the command line if necessary.

Manually Invoking the DOWS Synchronization Utility

The DOWS Synchronization Utility performs the following functions:

- Scans the Data Pool inventory for new and deleted OWS-enabled inventory.
- When inventory that is marked for pre-conversion is detected, creates a GeoTIFF Converter to convert HDF-EOS to GeoTIFF files and register the GeoTIFF files in the Data Pool database.
- If running in "insert" or "both" (i.e., both "insert" and "delete") synchronization mode, the new inventory (including generated GeoTIFF inventory) is added to the IONIC indexing system.
- If running in "delete" or "both" synchronization mode, the expired inventory is removed from the IONIC indexing system.

In addition to running in conjunction with another program, the Synchronizer may be invoked manually from the command line by a DAAC operator, especially when operating on data from a

time range prior to the regularly scheduled period. In such cases the operator specifies the ECS mode and the temporal extent of the synchronization session.

When specifying the temporal extent of the synchronization session, the operator specifies one of two mutually exclusive temporal modes:

- Previous mode, which specifies an amount of time before the time when the Synchronizer is executed (i.e., how long before the present time).
- Range mode, which specifies a period of time from a particular date/time to some other date/time.

In addition, the operator can specify a synchronization mode:

- Insert mode specifies that new inventory only (not deleted inventory) should be synchronized. The operator must specifically include this option on the command line to synchronize new inventory only.
- Delete mode specifies that deleted inventory only (not new inventory) should be synchronized. The operator must specifically include this option on the command line to synchronize deleted inventory only.
- Both mode specifies that both new inventory and deleted inventory should be synchronized. This is the default mode that the Synchronizer uses when the operator does not specifically include either the insert option or the delete option on the command line.

The procedure for manually invoking the Synchronization Utility (Synchronizer) starts with the assumption that the Data Pool and OGC database servers are currently running.

Manually Invoking the DOWS Synchronization Utility

---

1 Log in at the machine where the Synchronization Utility (Synchronizer) is installed (e.g., e0ogp01).

2 To change to the directory for starting the Synchronizer, type **cd /usr/ecs/<*MODE*>/CUSTOM/utilities** and then press the **Return/Enter** key.

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

NOTE: There are two mutually exclusive temporal modes in which the DOWS Synchronization Utility can be run; i.e., previous or range.

3 To run the Synchronizer in “range” mode, at the UNIX prompt enter:

EcOwSynchronizer -mode <*MODE*> -start <*STARTDATE*> [-stop <*STOPDATE*>] [-insert | -delete]

- **<*MODE*>** is the operating mode in which the utility is being executed (e.g., OPS, TS1, or TS2).

- ***<STARTDATE>*** is the start date/time range for the data to be synchronized. It has the format ***MM:DD:YY HH:MM***.
- ***<STOPDATE>*** is the ending date/time range, in format ***MM:DD:YY HH:MM***, for the data to be synchronized.
- The **-start** option takes a timestamp value in ***MM:DD:YY HH:MM*** format. It is mandatory for running the Synchronizer in the "range" mode.
- The **-stop** option takes a timestamp value in ***MM:DD:YY HH:MM*** format. It is optional (not mandatory) for running the Synchronizer in the "range" mode.
- The **-insert** option specifies that new inventory only (not deleted inventory) should be synchronized. It is optional (not mandatory).
- The **-delete** option specifies that deleted inventory only (not new inventory) should be synchronized. It is optional (not mandatory).
- For example:

  EcOwSynchronizer -mode OPS ·-start 03/06/01 22:30 -insert

  – The Synchronizer would synchronize new inventory inserted in the OPS mode in the time period between March 6, 2001 at 10:30 p.m. and the time when the command is executed.

  EcOwSynchronizer -mode OPS ·-start 03/06/01 22:30 -stop 03/07/01 01:20

  – The Synchronizer would synchronize inventory either inserted in or deleted from the OPS mode in the time period between on March 6, 2001 at 10:30 p.m. and March 7, 2001 at 1:20 a.m.
- The Synchronizer runs and records new operational-, debug-, and performance-related information in the **EcOwSynchronizer.ops.log**, **EcOwSynchronizer.debug.log**, and **EcOwSynchronizer.perf.log**.

4 To run the Synchronizer in "previous" mode, at the UNIX prompt enter:

EcOwSynchronizer -mode <*MODE*> -duration <*period*> [-count <*integer*>] [-insert | -delete]

- ***<period>*** specifies the type of units (i.e., "day", "hour", or "minute") that define the temporal constraint on the execution of the Synchronizer.
- ***<integer>*** is a positive integer value that describes how many of the "duration" units (i.e., days, hours, or minutes) are to be applied.
- The **-duration** option can take the values "day", "hour", or "minute". It is mandatory for running the Synchronizer in the "previous" mode.
- The **-count** option takes a positive integer value that describes how many of the duration units (i.e., days, hours, or minutes) are to be applied. It is optional (not

mandatory) for running the Synchronizer in the 'previous' mode. If no value for **–count** is specified, the assumed value is 1 (one).

- For example:

  EcOwSynchronizer -mode OPS -duration day -count 3

  – The Synchronizer would synchronize inventory either inserted in or deleted from the OPS mode in the time period between three days prior to the time when the command is executed and the time when the command is executed.

  EcOwSynchronizer -mode OPS -duration hour -insert

  – The Synchronizer would synchronize new inventory inserted in OPS mode in the time period between one hour prior to the time when the command is executed and the time when the command is executed.

- The Synchronizer runs and records new operational-, debug-, and performance-related information in the **EcOwSynchronizer.ops.log**, **EcOwSynchronizer.debug.log**, and **EcOwSynchronizer.perf.log**

Manually Invoking the DOWS GeoTIFF Converter Utility

The DOWS GeoTIFF Converter utility performs the following functions:

- Converts one or more HDF-EOS file(s) to GeoTIFF format.
- May register generated files in the Data Pool database (optional).
- May use a temporal constraint to determine a list of HDF-EOS files to be converted (optional).
- May use a supplied list of file names to determine which HDF-EOS files are to be converted (optional).

On a regular basis (e.g., daily) the Synchronizer (DOWS Synchronization Utility) supplies the GeoTIFF Converter with a list of science granules that the converter puts in a HEG request. The converter uses the HEG request to invoke the HEG client and have the granules converted from HDF-EOS format to GeoTIFF format.

In addition to running in conjunction with the Synchronizer or the OWS Registration Utility (described in a subsequent section of this lesson), the GeoTIFF Converter utility may be invoked manually from the command line by a DAAC operator, especially when operating on data from a time range prior to the most recent period. In such cases the operator specifies the ECS mode and the temporal extent of the conversion session.

When specifying the temporal extent of the conversion session, the operator specifies one of two mutually exclusive temporal modes:

- **Previous** mode, which specifies an amount of time before the time when the converter is executed (i.e., how long before the present time).
- **Range** mode, which specifies a period of time from a particular date/time to some other date/time.

The procedure for manually invoking the GeoTIFF Converter starts with the assumption that the Data Pool and OGC database servers are currently running.

Manually Invoking the DOWS GeoTIFF Converter Utility

---

1 Log in at the machine where the GeoTIFF Converter is installed (e.g., e0ogp01).

2 To change to the directory for starting the GeoTIFF Converter, type **cd /usr/ecs/<*MODE*>/CUSTOM/utilities** and then press the **Return/Enter** key.

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

NOTE: There are two mutually exclusive temporal modes in which the DOWS GeoTIFF Converter Utility can be run; i.e., "previous" or "range."

3 To run the converter in "range" mode, at the UNIX prompt enter:

EcOwGeotiffConverter -mode <*MODE*> -start <*STARTDATE*> [-stop <*STOPDATE*>]

- **<*MODE*>** is the operating mode in which the utility is being executed (e.g., OPS, TS1, or TS2).
- **<*STARTDATE*>** is the start date/time range for the data to be converted. It has the format ***MM:DD:YY HH:MM***.
- **<*STOPDATE*>** is the ending date/time range, in format ***MM:DD:YY HH:MM***, for the data to be converted.
- The **-start** option takes a timestamp value in ***MM:DD:YY HH:MM*** format. It is mandatory for running the converter in the "range" mode.
- The **-stop** option takes a timestamp value in ***MM:DD:YY HH:MM*** format. It is optional (not mandatory) for running the converter in the "range" mode.
- For example:

  EcOwGeotiffConverter -mode OPS ·-start 03/06/01 22:30

  – The converter would operate on OPS-mode data in the time period between March 6, 2001 at 10:30 p.m. and the time when the command is executed.

EcOwGeotiffConverter -mode OPS ·-start 03/06/01 22:30 -stop 03/07/01 01:20

– The converter would operate on OPS-mode data in the time period between on March 6, 2001 at 10:30 p.m. and March 7, 2001 at 1:20 a.m.

- The GeoTIFF Converter runs and records new operational-, debug-, and performance-related information in the **GeotiffConverter.ops.log**, **GeotiffConverter.debug.log**, and **GeotiffConverter.perf.log**.

4 To run the converter in "previous" mode, at the UNIX prompt enter:

EcOwGeotiffConverter -mode <*MODE*> -duration <*period*> [-count <*integer*>]

- **<*period*>** specifies the type of units (i.e., "day", "hour", or "minute") that define the temporal constraint on the execution of the converter.
- **<*integer*>** is a positive integer value that describes how many of the "duration" units (i.e., days, hours, or minutes) are to be applied.
- The **-duration** option can take the values "day", "hour", or "minute". It is mandatory for running the converter in the "previous" mode.
- The **-count** option takes a positive integer value that describes how many of the duration units (i.e., days, hours, or minutes) are to be applied. It is optional (not mandatory) for running the converter in the 'previous' mode. If no value for **–count** is specified, the assumed value is 1 (one).
- For example:

  EcOwGeotiffConverter -mode OPS -duration day -count 3

  – The converter would operate on OPS-mode data in the time period between three days prior to the time when the command is executed and the time when the command is executed.

  EcOwGeotiffConverter -mode OPS -duration hour

  – The converter would operate on OPS-mode data in the time period between one hour prior to the time when the command is executed and the time when the command is executed.

- The GeoTIFF Converter runs and records new operational-, debug-, and performance-related information in the **GeotiffConverter.ops.log**, **GeotiffConverter.debug.log**, and **GeotiffConverter.perf.log**.

Using the DOWS OWS Binding Utility (OBU)

Synergy V needs to provide OGC interoperability information to ECHO so that ECHO can serve as the OGC catalog for data pool holdings. To support this, a new configuration file for BMGT

has been created to make BMGT export collection-level metadata about OWS-enabled granules to ECHO. Also, a new metadata export utility, the OWS Binding Utility (OBU) exports granule-level information for OWS-enabled data in the Data Pool. This information indicates which granules from the OWS-enabled collections have been successfully indexed for WCS or WMS (or both) web service(s) by the DOWS, and already exported to ECHO via the BMGT. The Bulk URL Utility propagates the deletion of granules from the Data Pool to ECHO, to include OWS-enabled granules.

The DOWS Synchronization Utility performs web service indexing. The OWS Binding Utility detects which granules have been successfully indexed by the DOWS Synchronization Utility for either web service (i.e., WCS or WMS) before it exports the service metadata to ECHO. This is because the service on the granule will not have been validated prior to its successful processing by the WCS/WMS indexing function. Exporting the service metadata prematurely would result in failed service requests at the client; this is avoided by exporting web service metadata only after it has been validated through indexing. Consequently, there are two flags (i.e., one for WMS indexing status and another for WCS indexing status) in the Data Pool database that the DOWS Synchronization Utility sets to indicate granule indexing status. After the OWS Binding Utility detects a granule that has been successfully indexed for web services, it generates the appropriate XML metadata file, which includes separate URLs for WCS and WMS (as applicable). Since the granules will have been exported already to ECHO via BMGT, the OWS Binding Utility exports the WCS/WMS URLs for coverage/mapping to ECHO as an update for an existing granule, versus exporting an entire record as a new granule. ECHO supports (via a new API) the insertion of WCS/WMS URLs for an existing granule in its catalog.

In addition to running in conjunction with the Synchronizer, the OWS Binding Utility may be invoked manually from the command line by a DAAC operator.

The procedure for using the OWS Binding Utility (OBU) starts with the following assumptions:

- The appropriate version of the Java Runtime Environment (JRE) has been installed on the OBU host.
- Data Pool database server is currently running.

Using the DOWS OWS Binding Utility (OBU)

---

1 Log in at the machine where the OWS Binding Utility (OBU) is installed (e.g., e0ogp01).

2 To change to the directory for starting the OBU, type **cd /usr/ecs/<*MODE*>/CUSTOM/utilities** and then press the **Return/Enter** key.

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

3 At the UNIX prompt type **EcBmOwsBindingStart <*MODE*>** and then press the **Return/Enter** key.

- **<*MODE*>** is the operating mode in which the utility is being executed (e.g., OPS, TS1, or TS2).

- The OBU starts running and records new operational-, debug-, and performance-related information in the **EcBmOwsBinding.ops.log**, **EcBmOwsBinding.debug.log**, and **EcBmOwsBinding.perf.log**.

Standalone OGC (WCS/WMS) Archive

Synergy V includes a prototype of a stand-alone OGC solution that can serve as a low-cost alternative to the COTS products that implement OGC standards. In addition, it offers a path for integration and enabling of selected data holdings for OpenGIS Web Services (OWS) in order to promote interoperability. The initial deployment site is the GES DAAC.

One of the principal components of the Standalone OGC Archive is the Archive-to-OWS Gateway (AOG), which allows easy registration of existing data-provider holdings (as WCS coverages, WMS maps, and their corresponding metadata) with the OGC engine. A data provider that wishes to expose some or all of its WMS/WCS-compatible data has to perform a sequence of well-defined "OWS registration steps." After successful completion of the registration steps, the provider is able to respond to WMS and WCS HTTP requests via the OWS engine. In the GES DAAC prototype deployment the DAAC is the data provider and the WMS/WCS-compatible data is in the Data Pool.

A DAAC OWS utility, the OWS Registration Utility, provides the interface between the DAAC Data Pool and the AOG. The utility registers granules that have already been inserted into the GSFC Data Pool and unregisters granules that are no longer needed. The utility does not interfere with the operational Data Pool nor will it have any effect on the operational insert rate.

Using the Open Geospatial Consortium (OGC) Web Services (OWS) Registration Utility

The Open Geospatial Consortium (OGC) Web Services (OWS) Registration Utility provides the DAAC Operations Staff with a command-line interface for registering or unregistering Web Coverage Service (WCS) coverages and Web Mapping Service (WMS) layers for granules that reside in the Data Pool. It serves as an interface between the DAAC Data Pool and the Archive-to-OWS Gateway (AOG).

For registration (indicated by the command line parameter **-register**), the OWS Registration Utility requires a command-line parameter (i.e., **-granulefile**) that specifies an input file containing a list of granule IDs and another parameter (**-regtype**) specifying the type of registration; i.e., either WCS or WMS. The input file can be generated with the DOWS Synchronization Utility.

For each granule ID in the input file, the registration sequence is as follows:

1 The OWS Registration Utility converts each granule data file from HDF-EOS to GeoTIFF(s) for WCS coverage registrations. If the requested GeoTIFF files are already stored in the DPL area (i.e., they were preconverted on DPL insert), their existing location is registered and the conversion process is skipped. WMS layers are assumed to

be in jpg or png format. For multiple-file granules, a separate registration is performed for each file.

2 The utility extracts the metadata from the granule DPL metadata file and generates the AOG registration XML file containing the following information:

- OWS-style metadata.
- Type of registration being requested.
- Data file locations.

3 The utility places the AOG registration XML file, OWS metadata files, and OWS data files in a configured area. It notifies the AOG (via its socket interface) of the registration, passing the location of the registration file to the AOG. The AOG handles the remainder of the OWS registration process.

4 A group name can be registered (**-group** parameter). Coverages and layers can then be registered as part of that group.

During unregistration (indicated by the command line parameter **-unregister**) OWS-enabled data are unregistered from the OWS inventory. There is an option (**-deletedata** parameter) to remove the associated data files (WCS coverages and WMS layers). The OWS Registration Utility requires a command-line parameter that specifies an input file containing a list of granule IDs whose coverages or layers are to be removed. Alternately, a command-line parameter (**-resourcefile**) that specifies an input file containing a list of resource IDs may be used. (Coverages and layers are called "resources" in the OGC database.) If the registered data files do not reside under the dedicated OWS storage area, they are not removed via the AOG unregistration services. The unregistration is provided via an AOG XML file that contains unregistration instructions.

The procedure for using the OWS Registration Utility starts with the following assumptions:

- Data Pool and OGC database servers are currently running.
- AOG server is currently running.

Using the Open Geospatial Consortium (OGC) Web Services (OWS) Registration Utility

1 Log in at the machine where the OWS Registration Utility is installed.

- The OWS Registration Utility runs on a LINUX host (as determined locally) that runs the HEG tool.

2 To change to the directory for starting the OWS Registration Utility, type **cd /*path*** and then press the **Return/Enter** key.

- ***path*** is the path to the working directory (determined at installation) where the OWS Registration Utility (i.e., EcDlOWSUtility) is installed.

NOTE: There are two types of runs that can be performed with the OWS Registration Utility; i.e., registration or unregistration.

NOTE: There is no required sequence of parameters on the command line but not all combinations are valid. The parameters are validated by the utility. A command line input error results in a "usage" syntax display that (in most cases) explains why the command input was incorrect.

3 To perform a "registration" run of the OWS Registration Utility, at the UNIX prompt enter:

EcDlOWSUtility -register -regtype <*type*> -granulefile <*file name*> [-groupbytheme <*theme*>]

OR

EcDlOWSUtility -register -regtype <*type*> -geofile <*file name*>

OR

EcDlOWSUtility -register -regtype <*type*> -granulefile <*file name*> [-groupbycollgroup]

OR

EcDlOWSUtility -register -regtype <*type*> -granulefile <*file name*> [-groupbycoll]

OR

EcDlOWSUtility -register -regtype <*type*> -granulefile <*file name*> [-group <*groupname*>] [-parent <*groupname*>]

- **<*type*>** is the value specified for the type of registration or unregistration. The value is either **WCS** or **WMS**.
- **<*file name*>** is the name of a file to be associated with either the **-granulefile** option or the **-geofile** option. When used with the **-granulefile** option, it specifies the name of the file containing granule IDs to be used as input to the utility. When used with the **-geofile** option, it specifies the name of the file containing geoTIFF and xml (DPL style) pathnames to be used as input to the utility.
- **<*theme*>** is the name of a theme to be associated with the **-groupbytheme** option, which specifies that each granule in the input file is to be grouped under three levels; i.e., the Data Pool theme, followed by its collection group, followed by its collection (in accordance with the group tree structure maintained in the OGC database). The **-groupbytheme** parameter is valid only when both **-register** and **-granulefile** are specified.
- **<*groupname*>** is the name of a group to be associated with either the **-group** option or the **-parent** option. When used with the **-group** option, it specifies that a group is to be registered. If a **-parent** is not specified, the specified group has no parent (i.e.,

it is the highest level group). When used with the **-parent** option, the group name specifies the parent group.

- The **-groupbycollgroup** option specifies that each granule in the input file is to be grouped under two levels; i.e., its collection group followed by its collection (in accordance with the group tree structure maintained in the OGC database). The **-groupbycollgroup** parameter is valid only when both –register and –granulefile are specified.
- The **-groupbycoll** option specifies that each granule in the input file is to be grouped under its collection (in accordance with the group tree structure maintained in the OGC database). The **-groupbycoll** parameter is valid only when both **-register** and **-granulefile** are specified.
- The OWS Registration Utility runs and the log file **EcDlOWSUtility.log.*n*** records update events and errors.

4 To perform an "unregistration" run of the OWS Registration Utility, at the UNIX prompt enter:

EcDlOWSUtility -unregister -regtype <type> -granulefile <file name> [-deletedata]

OR

EcDlOWSUtility -unregister -regtype <type> -resourcefile <file name> [-deletedata]

- ***<file name>*** is the name of a file to be associated with either the **-granulefile** option or the **-resourcefile** option. When used with the **-granulefile** option, it specifies the name of the file containing granule IDs to be used as input to the utility. When used with the **-resourcefile** option, it specifies the name of the file containing resource IDs to be used as input to the utility.
- The **-deletedata** option specifies that the WCS or WMS files will be deleted during unregistration. The **-deletedata** parameter is valid with **-unregister** only.
- The OWS Registration Utility runs and the log file **EcDlOWSUtility.log.*n*** records update events and errors.

Using the Spatial Subscription Server (NBSRV) GUI

Spatial Subscription Server (NBSRV) GUI

Although there are some Data Pool maintenance tasks that are done with scripts and utilities, as described in a subsequent section of this lesson, the **Spatial Subscription Server (NBSRV) GUI** provides a convenient means of modifying the period of retention in a Data Pool insert subscription and designating a subscription for secure distribution. This can be done by User Services and/or science personnel, but the archive support personnel should also be familiar with the GUI and its use. Other tasks done with the **Spatial Subscription Server GUI** may also be of use to archive personnel (e.g., viewing the acquire and notification actions currently being processed by the Spatial Subscription Server, viewing statistics on the processing of events and actions by the Spatial Subscription Server).

The **Spatial Subscription Server GUI** permits an operator to view subscribable events and review existing subscriptions in the NBSRV database. It allows an operator to add a subscription specifying Data Pool qualification and retention criteria, thus adding a subscription for Data Pool insert to the database. The operator may also create a standard subscription for notification and/or distribution of ECS data products and designate a subscription for secure distribution.

For an existing subscription, there may occur circumstances that make it desirable to retain certain data in the Data Pool for a longer period of time than originally specified (e.g., an unusual ground event, such as a volcanic eruption or flood, causes granules for that geographic area to be of more than usual interest). Data Pool maintenance personnel can run a script to update the expiration date for selected science granules. The **Spatial Subscription Server GUI** can be used to extend the period of retention in a Data Pool insert subscription for new granules from that area.

The **Spatial Subscription Server GUI** is also used to view, update, or delete (cancel) subscriptions in the NBSRV database. Finally, it permits an operator to view the acquire and notification actions currently being processed by the Spatial Subscription Server, and to view statistics on the processing of events and actions by the Spatial Subscription Server.

New operator GUI security standards require the following two levels of permissions for the **Spatial Subscription Server GUI**:

- Full Capability.
- Limited Capability.

An operator's level of permission is determined when the operator logs in to the GUI using the security login prompt.

Full-capability operators have the ability to configure parameters and perform all other actions that can be accomplished with the **Spatial Subscription Server 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.

This lesson provides instruction in the full-capability version of the **Spatial Subscription Server GUI**. In general, both full-capability operators and limited-capability operators can **view** the following items:

- Subscribable events.
- Subscriptions.
- Bundling orders.
- Action queue.
- Statistics relating to Spatial Subscription Server performance.

Full-capability operators only may perform the actions:

- Add, update, or delete (cancel) a subscription.
- Configure defaults for a bundling order.
- Add, update, or cancel a bundling order.

The **Spatial Subscription Server GUI** is a web application certified for use with Netscape 7 (and higher) browsers.

Figure 61 illustrates the **Spatial Subscription Server GUI Home Page**, from which the operator can navigate to other pages to begin specific tasks. As the figure shows, besides the **Home Page** link, it provides four links for access to pages supporting various tasks:

- **List Events**: access to pages for listing subscribable events.
- **Manage Subscriptions**: access to pages for managing subscriptions.
- **Manage Bundling Orders**: access to pages for managing bundling orders.
- **Monitor Queues**: access to pages for monitoring the action queue and listing statistics.

There is also a **Help** link providing descriptions of the NBSRV functions to provide the operator with assistance in navigating through the GUI.

Figure 62 illustrates the **List Events** page. It allows the operator to view all events for which a subscription can be created. The column headers are links permitting the operator to sort the list by collection, version, or event type. There are also three option buttons and a **Filter** button, permitting the operator to filter the list by collection, version, and event type.

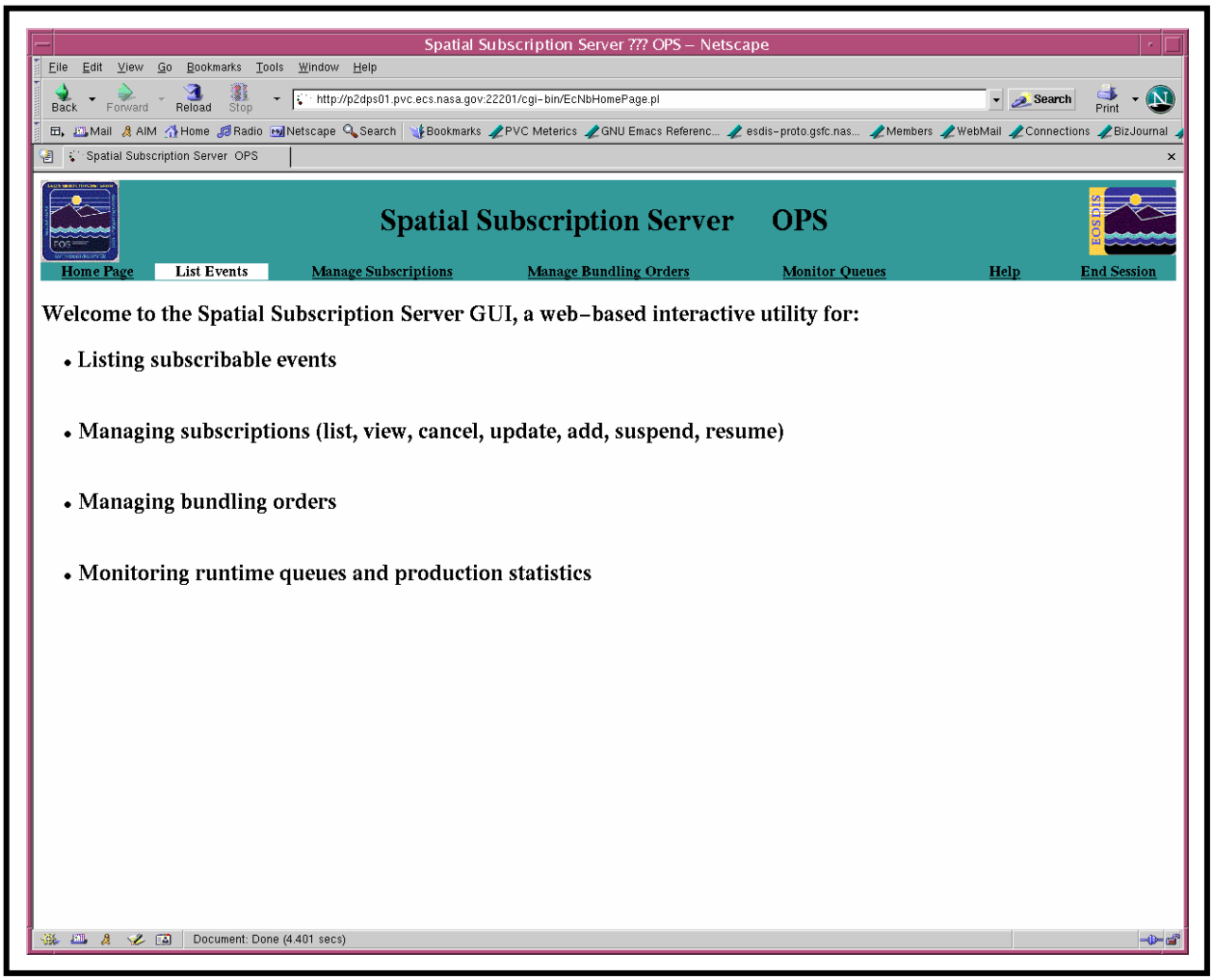


Figure 61. Spatial Subscription Server GUI Home Page

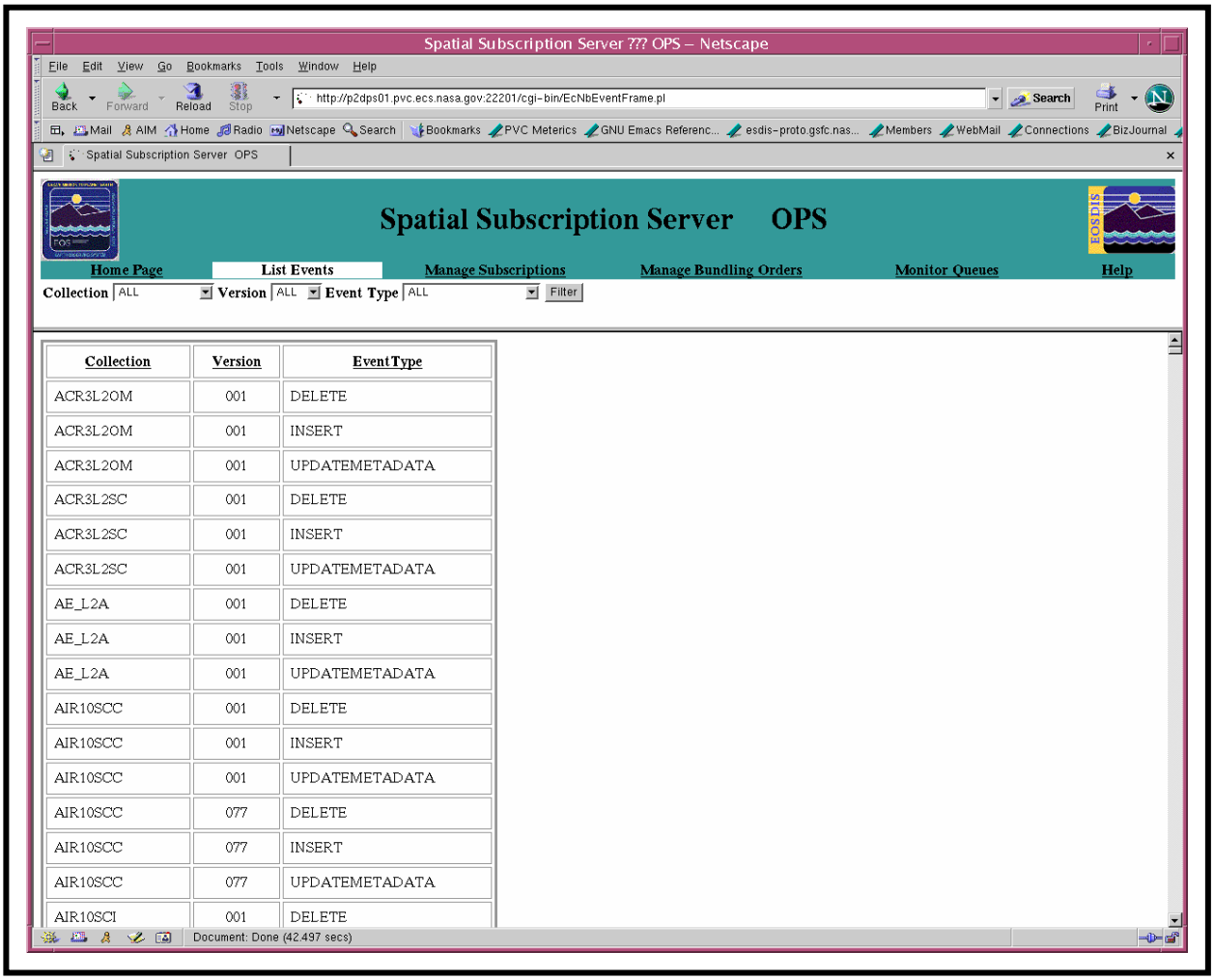
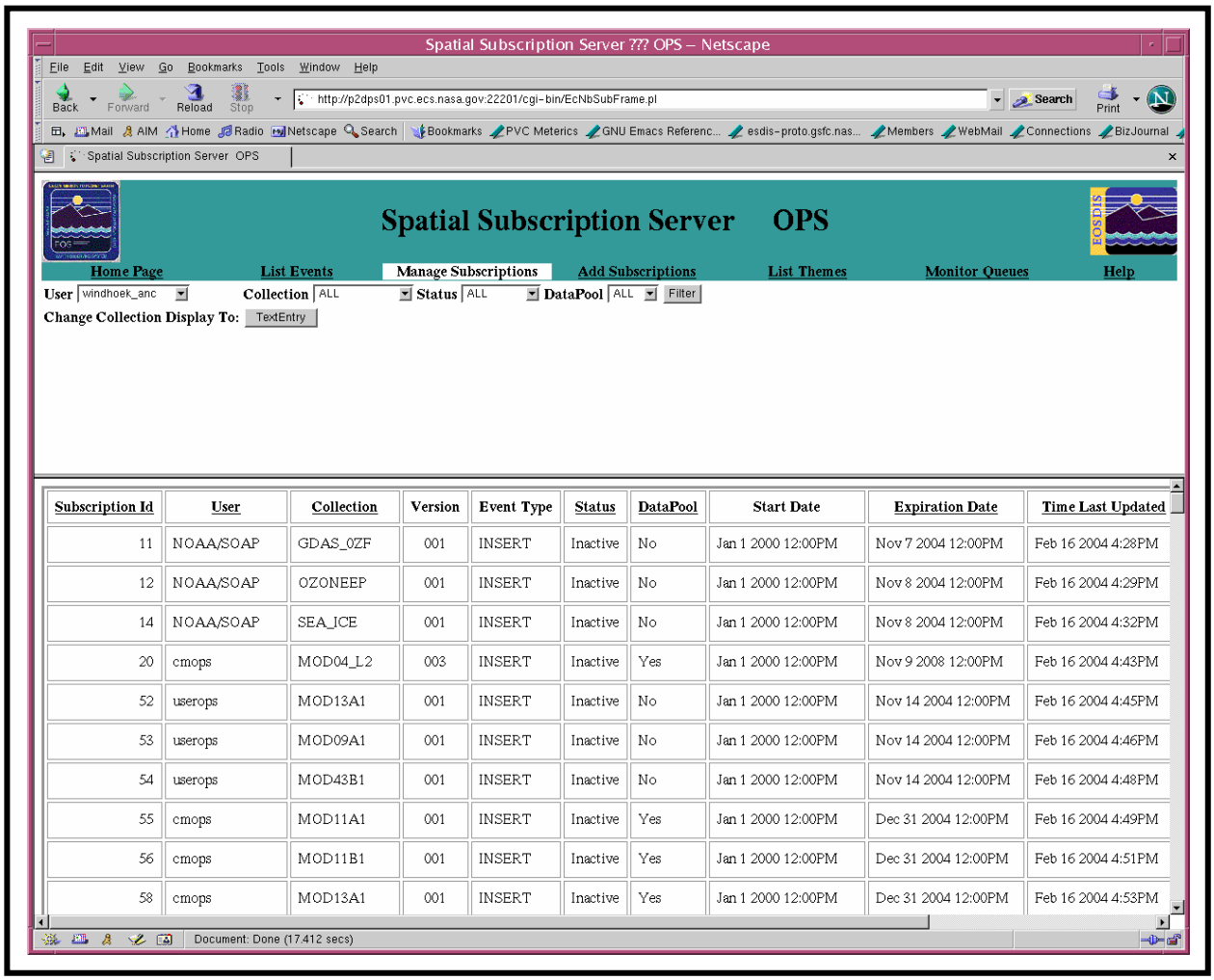


Figure 62. Spatial Subscription Server GUI List Events Page

Figure 63 shows the **Manage Subscriptions** page. This page lists subscriptions existing in the NBSRV database, with columns for **Subscription Id**, **User** for whom the subscription was created, **Collection**, **Version**, **Event Type**, **Status**, **Data Pool, Start Date, Expiration Date, Time Last Updated and Choose Subscription Action**. These column headers, except for **Version** and **Event Type**, are links on which the list can be sorted. There are also three option buttons and a **Filter** button, permitting the operator to filter the list by user, collection, and status. Finally, at the right side of the page, there is a **Choose Subscription Action** column permitting operators an option, for each listed subscription, to **View**, **Update**, or **Delete** the subscription, with an **Apply** button to implement a selected option. Note that when the **Manage Subscriptions** page appears, the GUI displays a new link, **List Themes**, as shown in the figure.



Subscription Id	User	Collection	Version	Event Type	Status	DataPool	Start Date	Expiration Date	Time Last Updated
11	NOAA/SOAP	GDAS_0ZF	001	INSERT	Inactive	No	Jan 1 2000 12:00PM	Nov 7 2004 12:00PM	Feb 16 2004 4:28PM
12	NOAA/SOAP	OZONEEP	001	INSERT	Inactive	No	Jan 1 2000 12:00PM	Nov 8 2004 12:00PM	Feb 16 2004 4:29PM
14	NOAA/SOAP	SEA_ICE	001	INSERT	Inactive	No	Jan 1 2000 12:00PM	Nov 8 2004 12:00PM	Feb 16 2004 4:32PM
20	cmops	MOD04_L2	003	INSERT	Inactive	Yes	Jan 1 2000 12:00PM	Nov 9 2008 12:00PM	Feb 16 2004 4:43PM
52	userops	MOD13A1	001	INSERT	Inactive	No	Jan 1 2000 12:00PM	Nov 14 2004 12:00PM	Feb 16 2004 4:45PM
53	userops	MOD09A1	001	INSERT	Inactive	No	Jan 1 2000 12:00PM	Nov 14 2004 12:00PM	Feb 16 2004 4:46PM
54	userops	MOD43B1	001	INSERT	Inactive	No	Jan 1 2000 12:00PM	Nov 14 2004 12:00PM	Feb 16 2004 4:48PM
55	cmops	MOD11A1	001	INSERT	Inactive	Yes	Jan 1 2000 12:00PM	Dec 31 2004 12:00PM	Feb 16 2004 4:49PM
56	cmops	MOD11B1	001	INSERT	Inactive	Yes	Jan 1 2000 12:00PM	Dec 31 2004 12:00PM	Feb 16 2004 4:51PM
58	cmops	MOD13A1	001	INSERT	Inactive	Yes	Jan 1 2000 12:00PM	Dec 31 2004 12:00PM	Feb 16 2004 4:53PM

Figure 63. Spatial Subscription Server GUI Manage Subscriptions Page

When the **Manage Subscriptions** page is displayed, the GUI also offers a link for access to a page permitting the operator to **Add Subscriptions**. Figures 64 and 65 show the **Add Subscriptions** page as it appears after an operator has entered a **User Id** and **Expiration Date**, and selected an **ESDT Short Name/Version/Event Type**. If the subscription is to specify an **Acquire** or **Notify** action, it is necessary to enter information in the **Acquire Information** or **E-Mail Notification Information** portion of the **Actions** block of the form. It is also possible to specify a bundling order associated with acquire actions, to accumulate granules until the achievement of some criterion for completion of a bundle before the data are actually distributed. The entered information is necessary to support notification and distribution for ECS standard products. Additionally, a subscription can be designated for secure distribution by selecting the scp **Media Type** and entering the required FTP information.

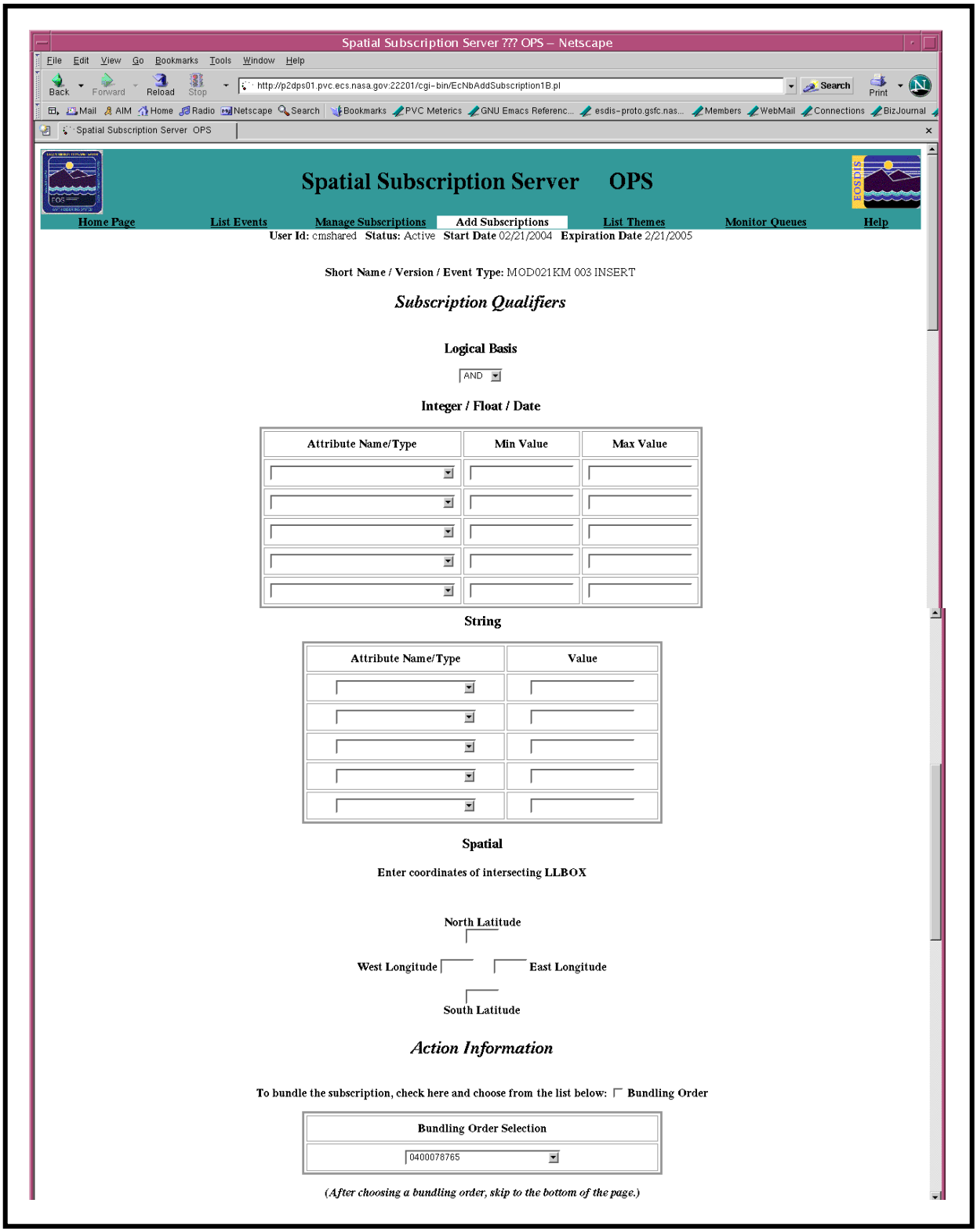


Figure 64. Spatial Subscription Server GUI Add Subscriptions Page (Top)

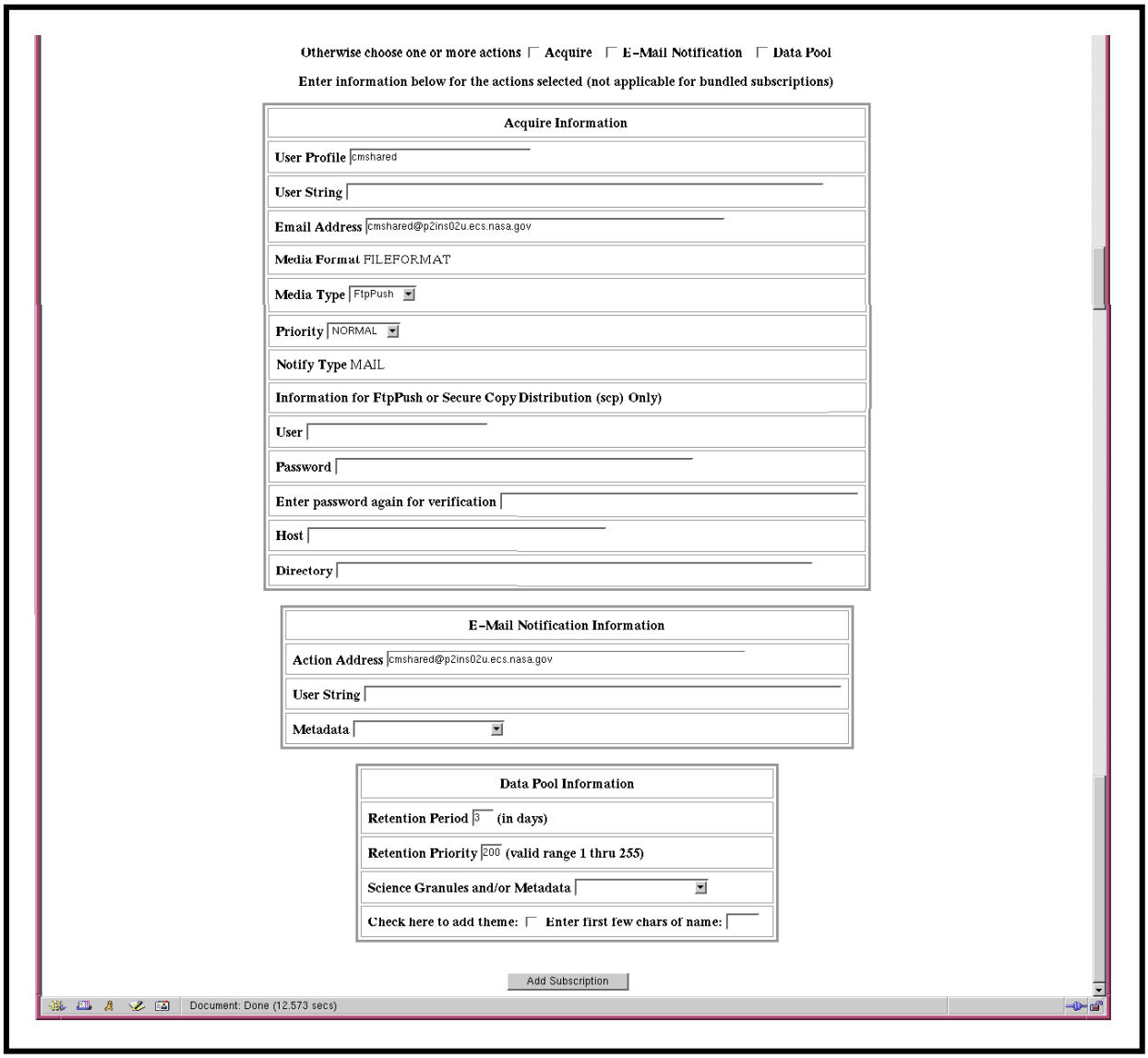


Figure 65. Spatial Subscription Server GUI Add Subscriptions Page (Bottom)

If an operator selects and implements the **View** option for a subscription (listed on the **Manage Subscriptions** page), the GUI provides a display similar to that illustrated in Figure 66. The page lists detailed information about the selected subscription, including information concerning any associated theme and any associated bundling order.

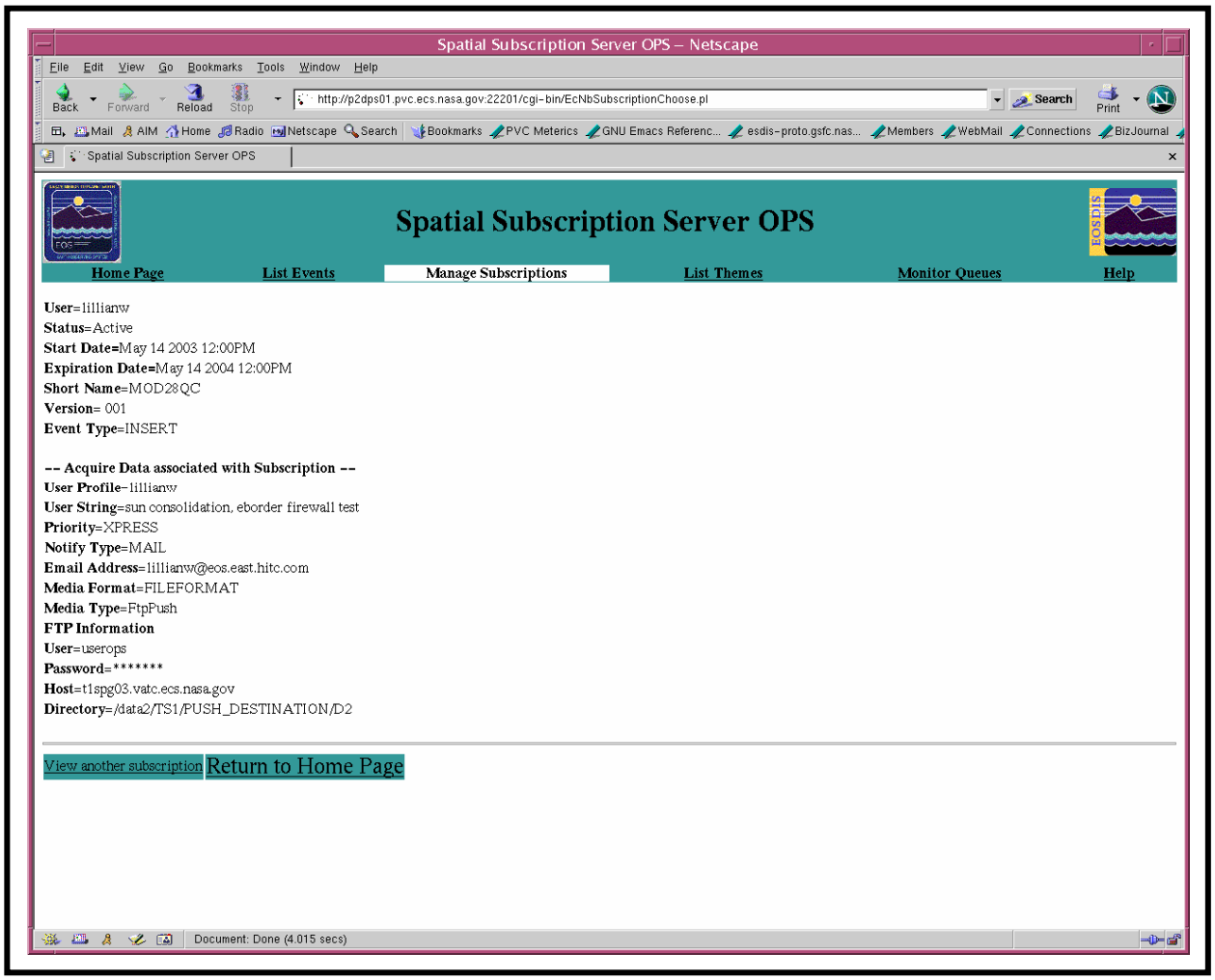


Figure 66. Spatial Subscription Server GUI View Subscriptions Page

If an operator selects and implements the **Update** option for a subscription (listed on the **Manage Subscriptions** page), the GUI provides a display similar to that illustrated in Figures 67 and 68. This page is structured very similarly to the **Add Subscription** page, permitting the operator to add or change qualifiers and change the action selection for the subscription. It also permits the operator to specify a bundling order for the subscription as well as providing the fields required to designate secure distribution subscriptions (**Media Type**, **FTP User**, **FTP Password**, **FTP Password Verification**, **FTP Host**, and **FTP Directory**). Changes are submitted by clicking on the **Update Subscription** button at the bottom of the page.

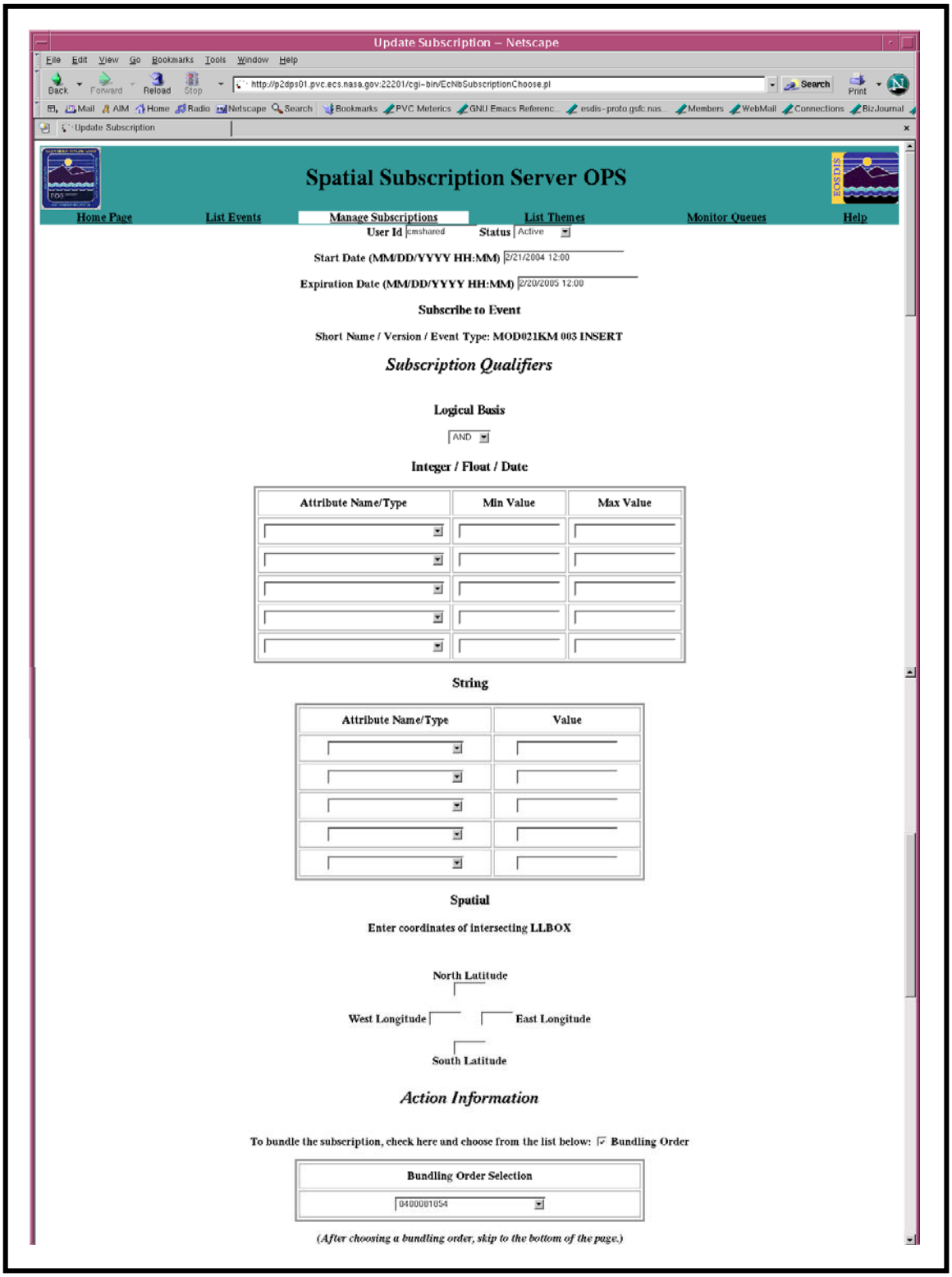


Figure 67. Spatial Subscription Server GUI Update Subscriptions Page (Top)

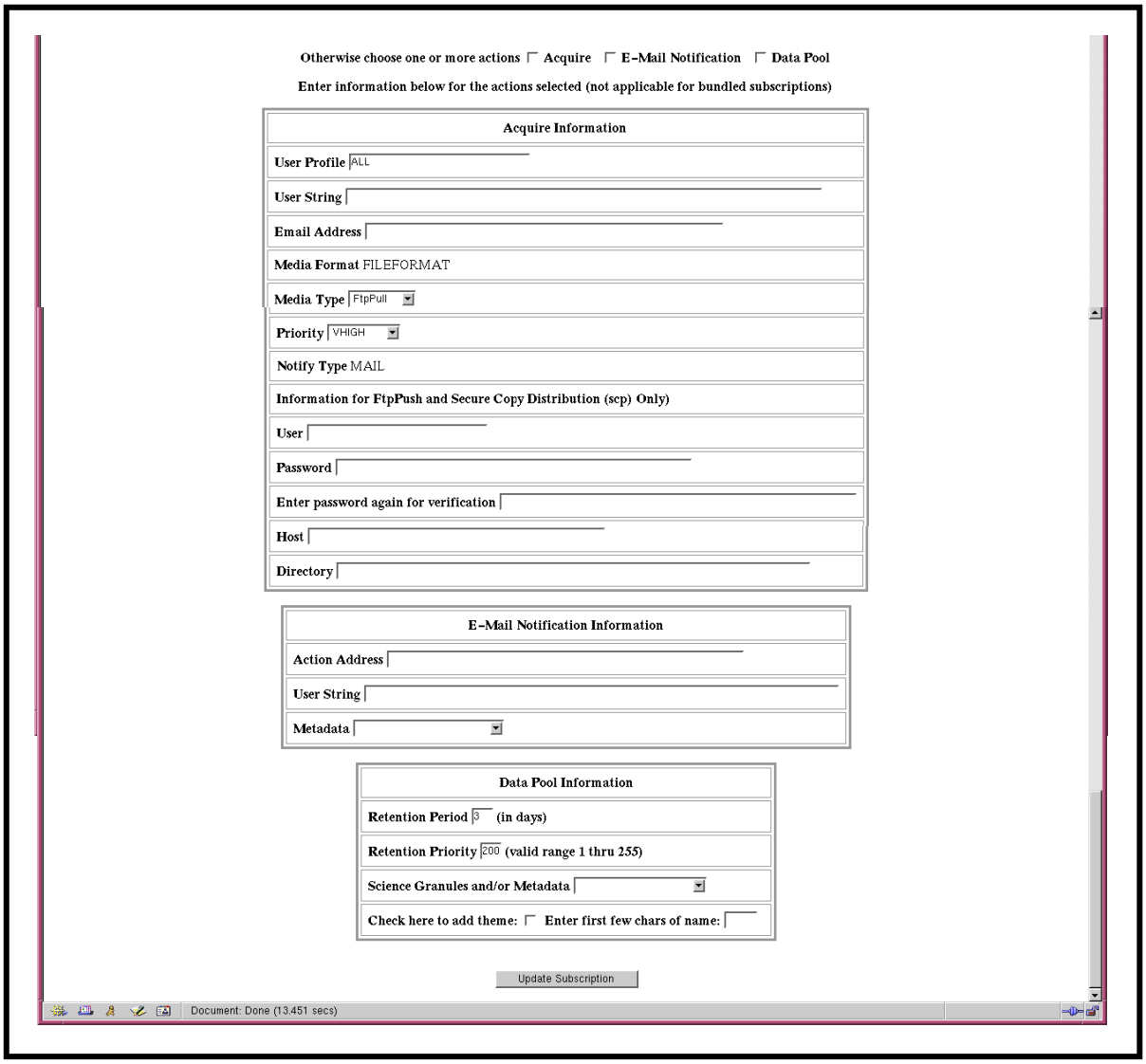


Figure 68. Spatial Subscription Server GUI Update Subscriptions Page (Bottom)

If the operator clicks on the **List Themes** link while managing subscriptions, the GUI displays the **List Themes Request** page illustrated in Figure 69. By clicking on the **Apply** button the operator obtains a **Theme List** page, as shown in Figure 70. The list may be filtered (to select a particular theme) by entering the first few characters of the theme name in the text box on the **List Themes Request** page (Figure 69) before clicking on the **Apply** button.

By clicking on a check box on the **Theme List** page (to specify a theme for which subscriptions should be listed) and then clicking on the **Apply** button the operator obtains a **List Subscriptions for Theme** page, as shown in Figure 71.

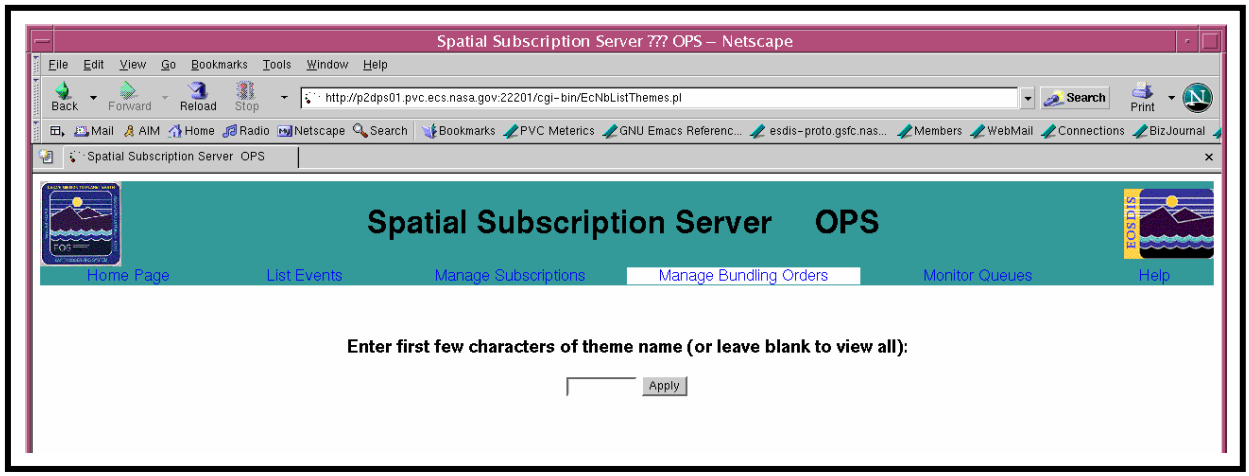


Figure 69. Spatial Subscription Server GUI List Themes Request Page

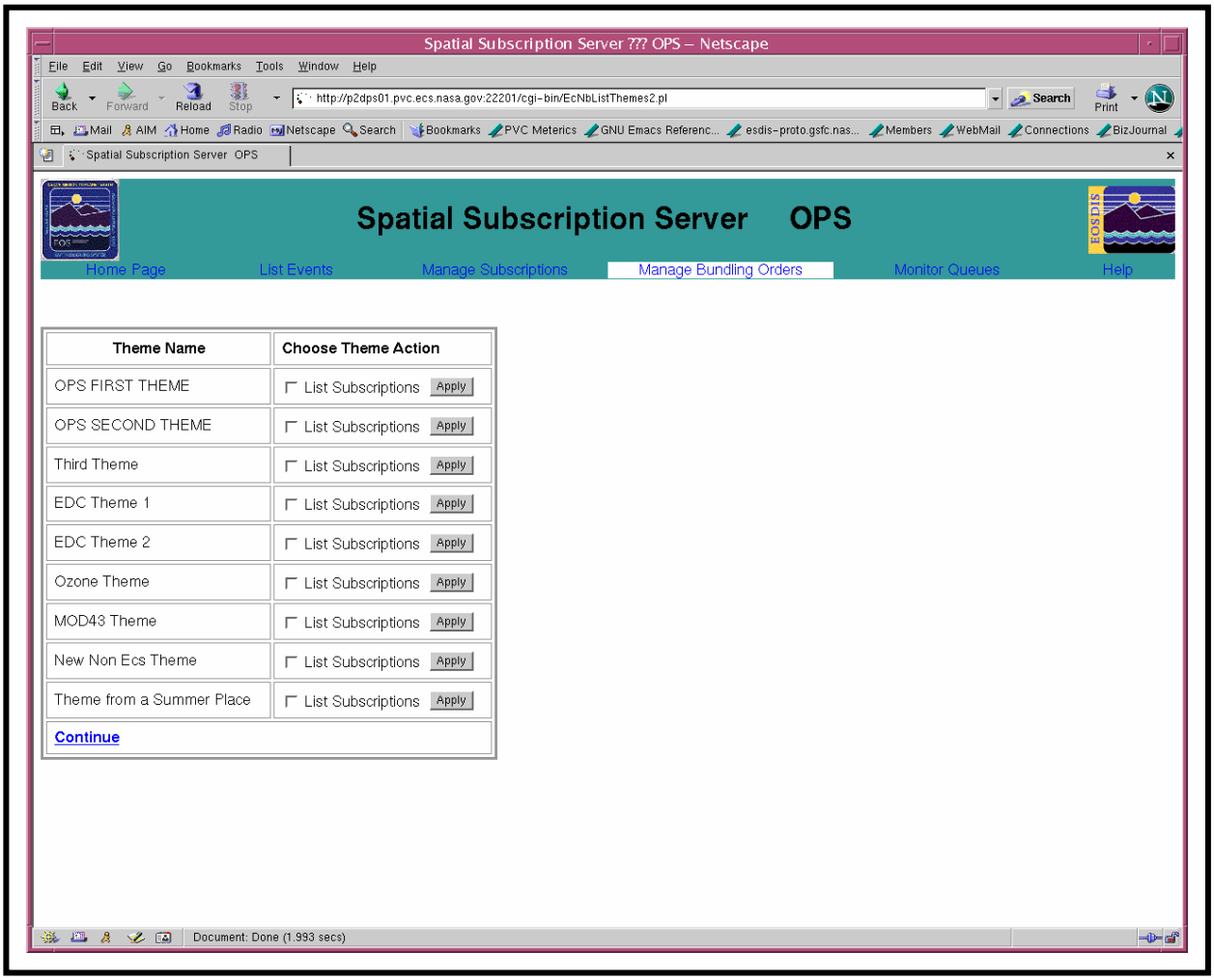


Figure 70. Spatial Subscription Server GUI Theme List Page

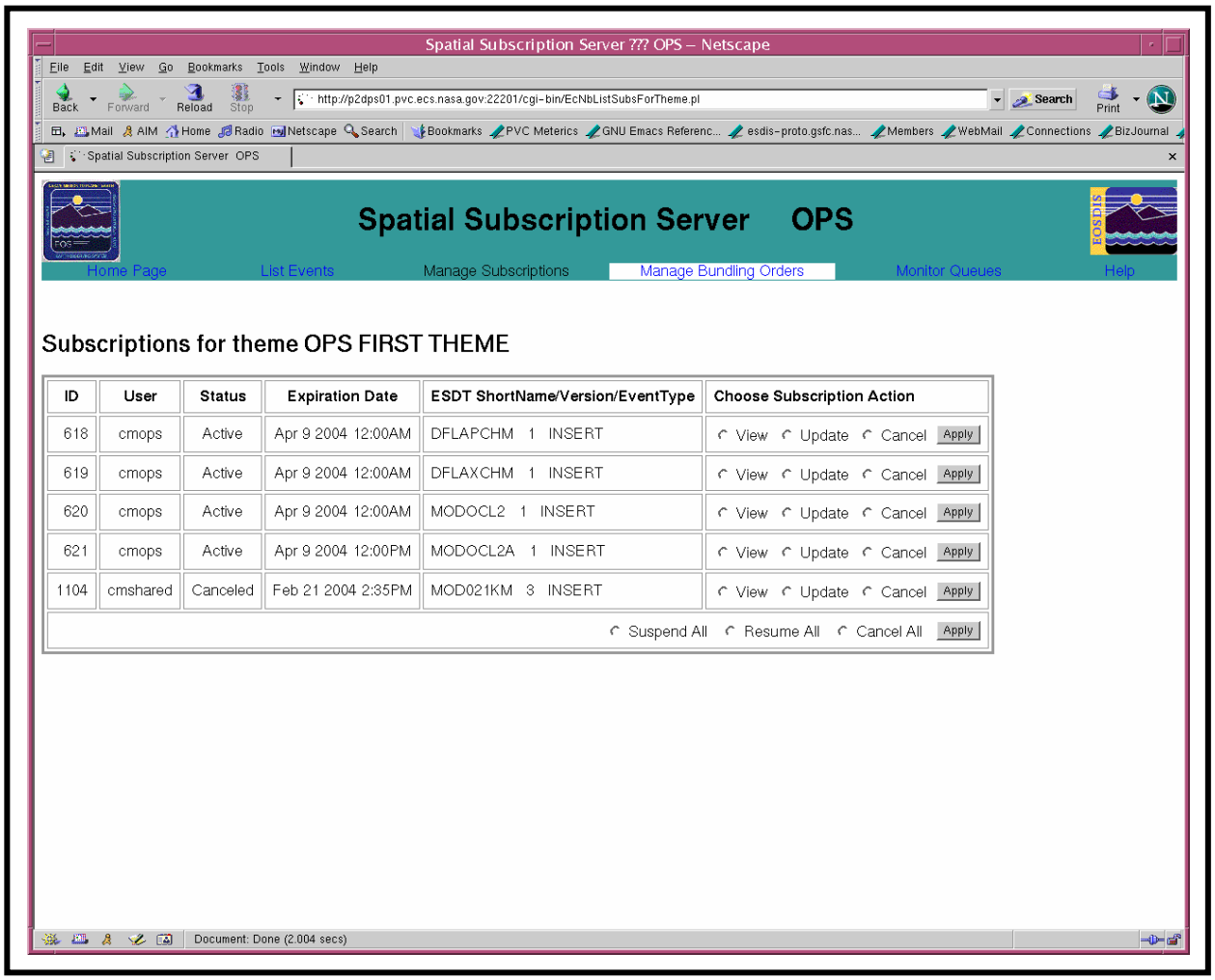


Figure 71. Spatial Subscription Server GUI List Subscriptions for Theme Page

If the operator clicks on the **Manage Bundling Orders** link on the **Home Page** or other page where the link is available, the GUI displays the page illustrated in Figure 72, listing bundling orders. The list includes a **Bundling Order** identifying number for each bundling order, the **User** for whom the bundling order was created, the **Creation Date** and **Expiration Date**, the **Media Type**, and **Status**. There is an **Add Bundling Order** link to a page where the operator can add a bundling order and a **Configure Defaults** link to a page where the operator can set default parameters for configuring bundling orders. There are also three option buttons and a **Filter** button, permitting the operator to filter the list by user, media type, and status. Finally, at the right side of the page, there is a **Choose Bundling Order Action** column permitting operators an option, for each listed bundling order, to **View**, **Update**, **Cancel**, or **List Subs** (list the subscriptions) associated with the bundling order, with an **Apply** button to implement a selected option. The figure shows the window that is obtained by selecting and implementing the **View** option.

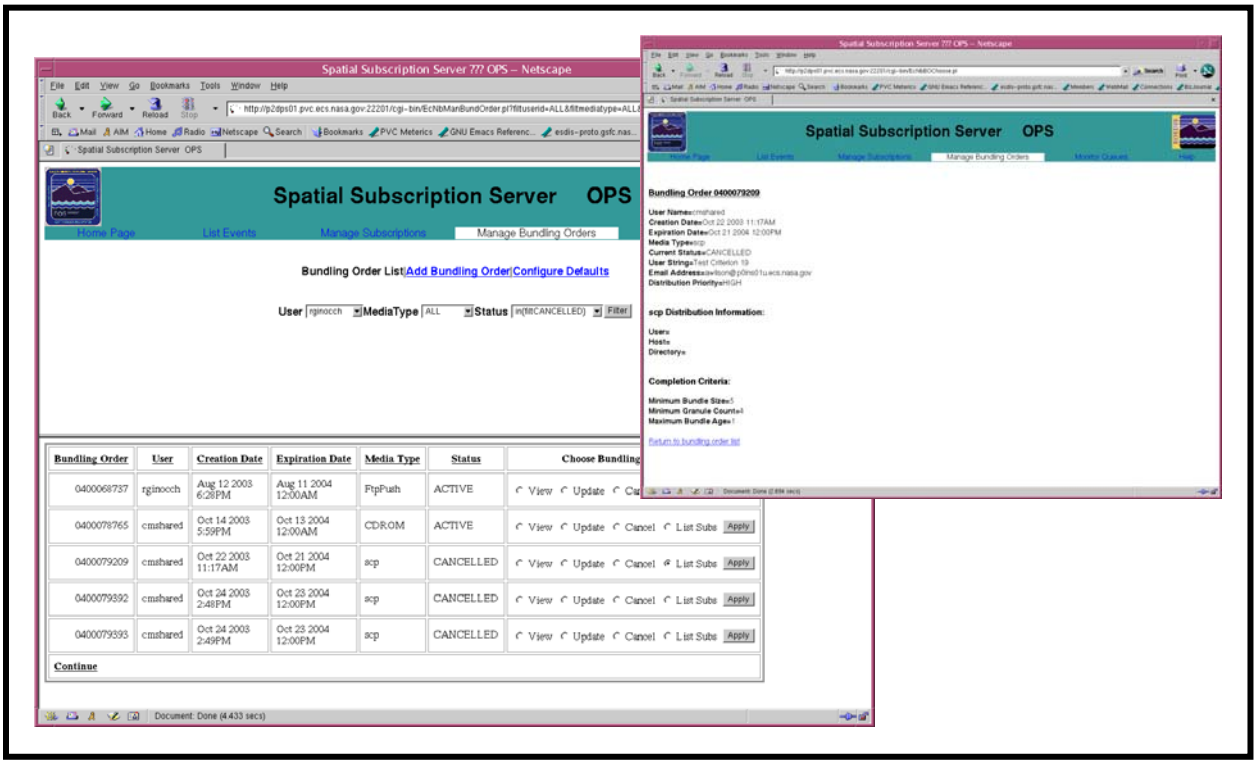


Figure 72. Spatial Subscription Server GUI Manage Bundling Orders Page

If the operator clicks on the **Add Bundling Order** link, the GUI displays the page illustrated in Figure 73. This page contains a field for entry of the **UserID**, a field for specification of an **Expiration Date** for the bundling order (with a default entry), and an option button for selection of a media type (CD ROM, 8 mm tape, DVD, DLT, FtpPush, FtpPull, or scp). There is a **continue** button at the bottom to continue to a page for further specification of data related to the bundling order; the data to be entered on the resulting page varies with the media type selection. There is also a **clear** button that can be used to clear the **UserID** field and restore the default **Expiration Date** and **Media Type**.

Figures 74 and 75 show the data pages that result from clicking on the **continue** button at the **Add Bundling Order** page when the **Media Type** chosen is a hard media type, an FTP Push, an FTP Pull or a scp (Secure Copy). These pages permit entry of information necessary for full specification of the selected **Media Type**. Each one also requires specification of three criteria for completion of a bundle, whereby fulfillment of any one of the three criteria will trigger distribution of the bundle. The criteria include:

- Minimum bundle size (GB).
- Minimum Granule Count.
- Maximum Bundle Age (days).

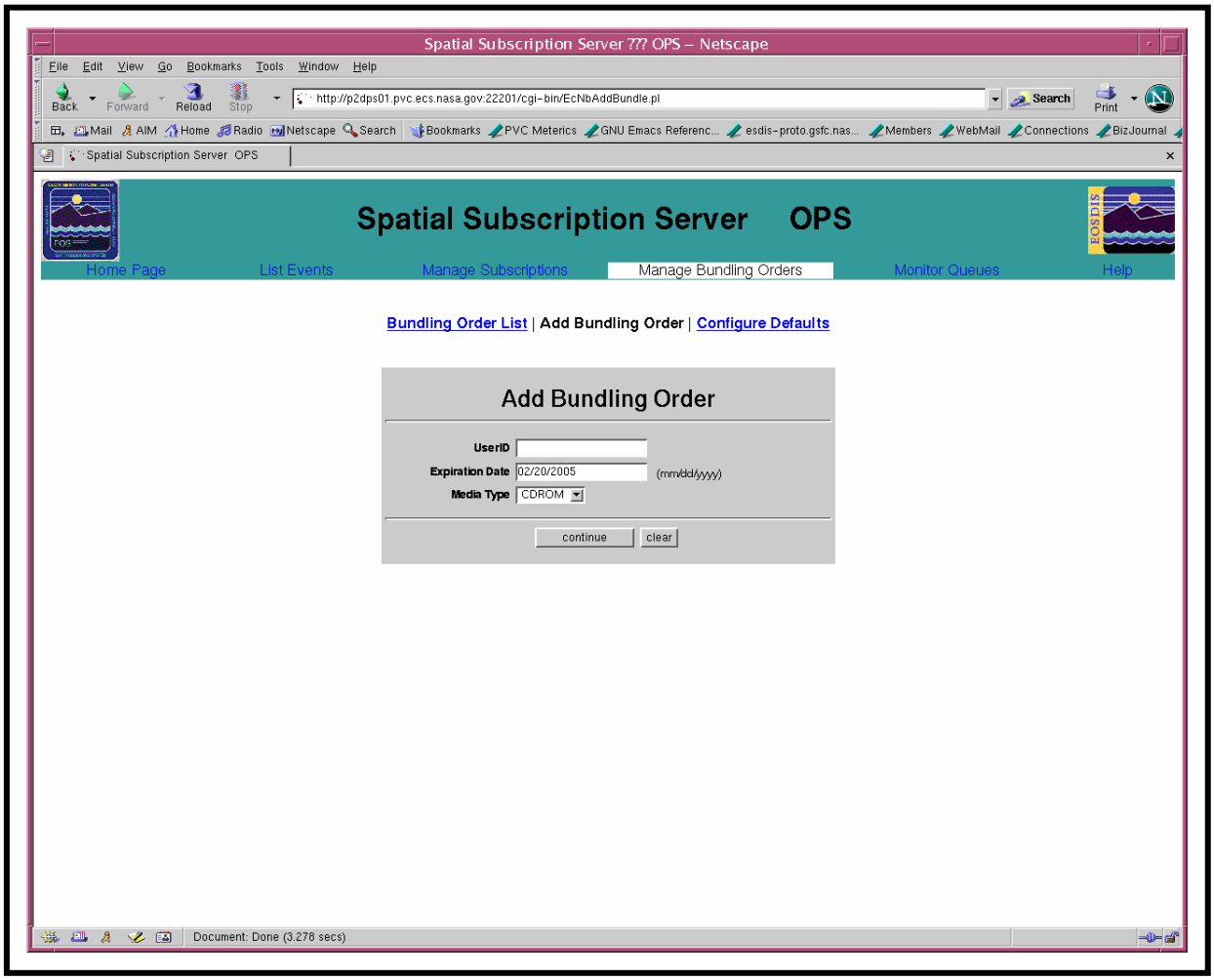


Figure 73. Spatial Subscription Server GUI Add Bundling Order Page

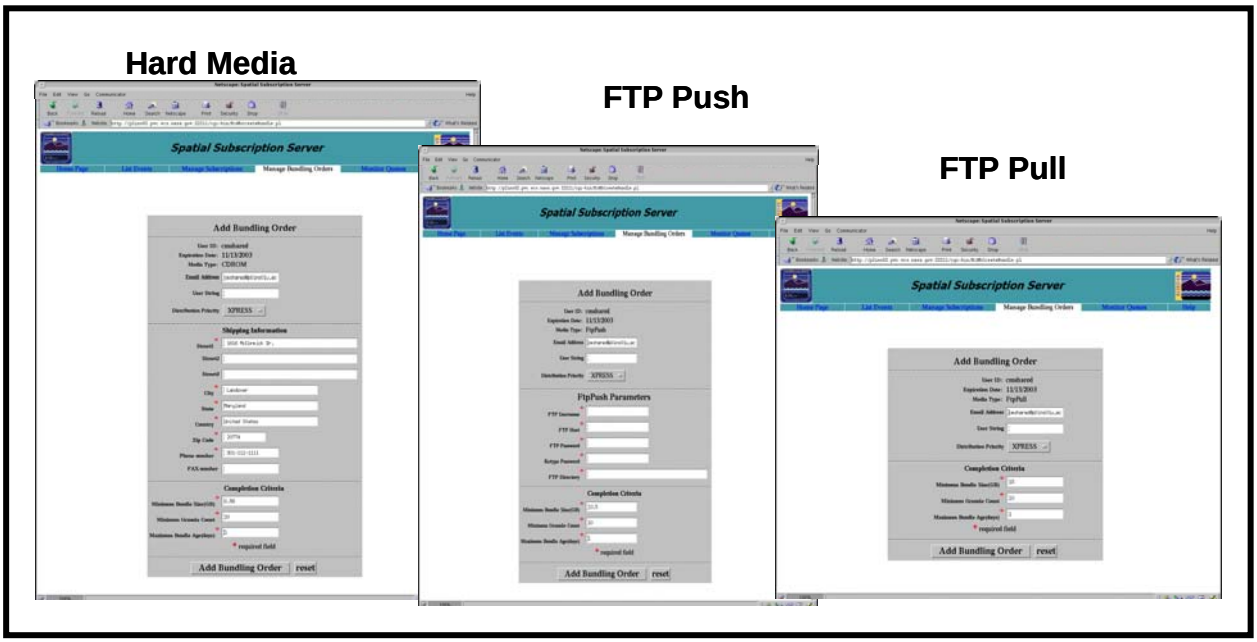


Figure 74. Spatial Subscription Server GUI Data Pages for Adding Bundling Orders (Hard Media, FTP Push, or FTP Pull)

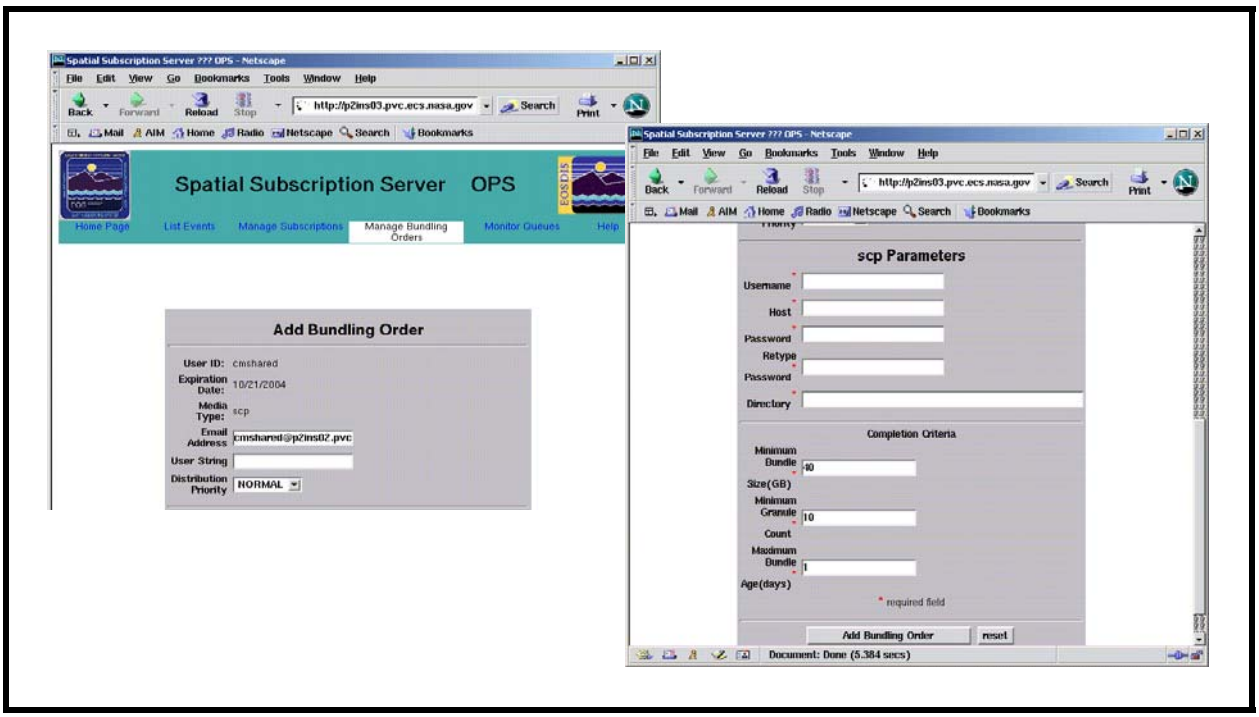


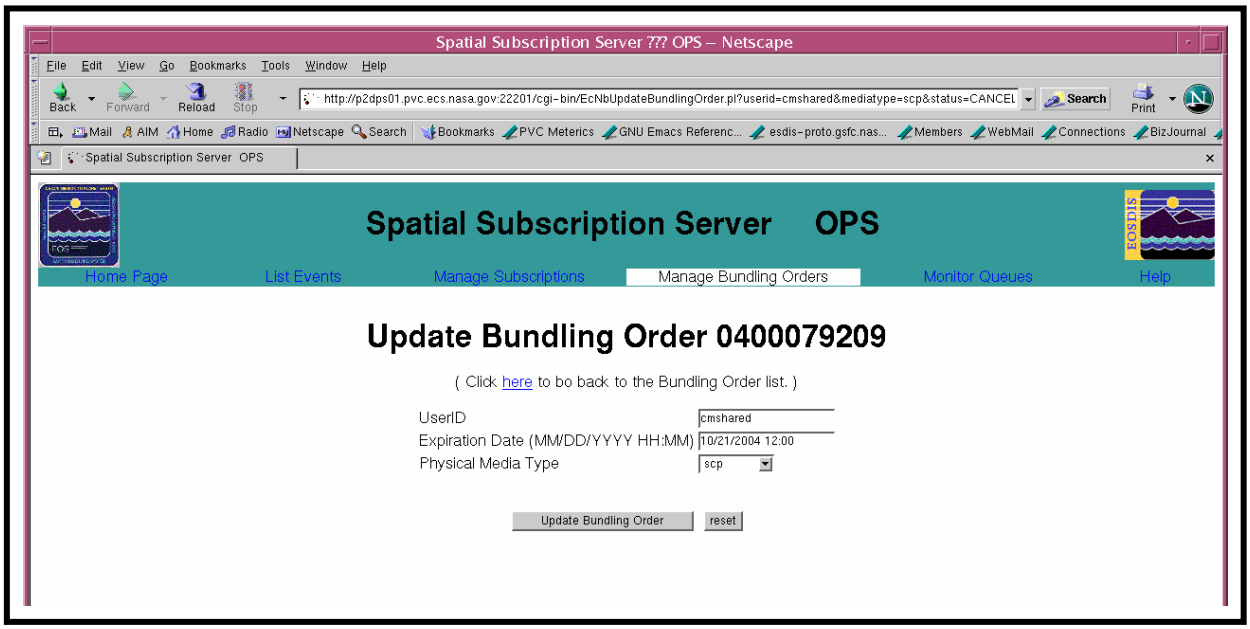
Figure 75. Spatial Subscription Server GUI Data Pages for Adding Bundling Orders (Secure Distribution)

When adding a secure distribution bundling order, scp must be selected as the Media Type and **FTP User, FTP Host** and **FTP Password** must be completed.

If the operator clicks on the **Update Bundling Order** link on the **Manage Bundling Orders** page, the GUI displays the **Update Bundling Order** page (Figure 76), which allows the updating of an existing bundling order. The page contains a field for entry of the **UserID**, a field for specification of an **Expiration Date** for the bundling order (with a default entry), and an option button for selection of a type of distribution medium (e.g., CD ROM, 8 mm tape, DVD, DLT, FtpPush, FtpPull, or scp). There is an **Update Bundling Order** button at the bottom to continue to a data page for specification of additional data related to the bundling order. There is a **reset** button on the **Update Bundling Order** page that can be used to clear the **UserID** field and restore the default **Expiration Date** and **Media Type**.

The fields displayed on the data pages (Figure 77) vary with the medium selected. The data pages permit entry of information necessary for full specification of the selected **Media Type**. Each one also requires specification of three criteria for completion of a bundle, whereby fulfillment of any one of the three criteria will trigger distribution of the bundle. The criteria include:

- Minimum bundle size (GB).



- Minimum Granule Count.

Figure 76. Spatial Subscription Server GUI Update Bundling Order Page

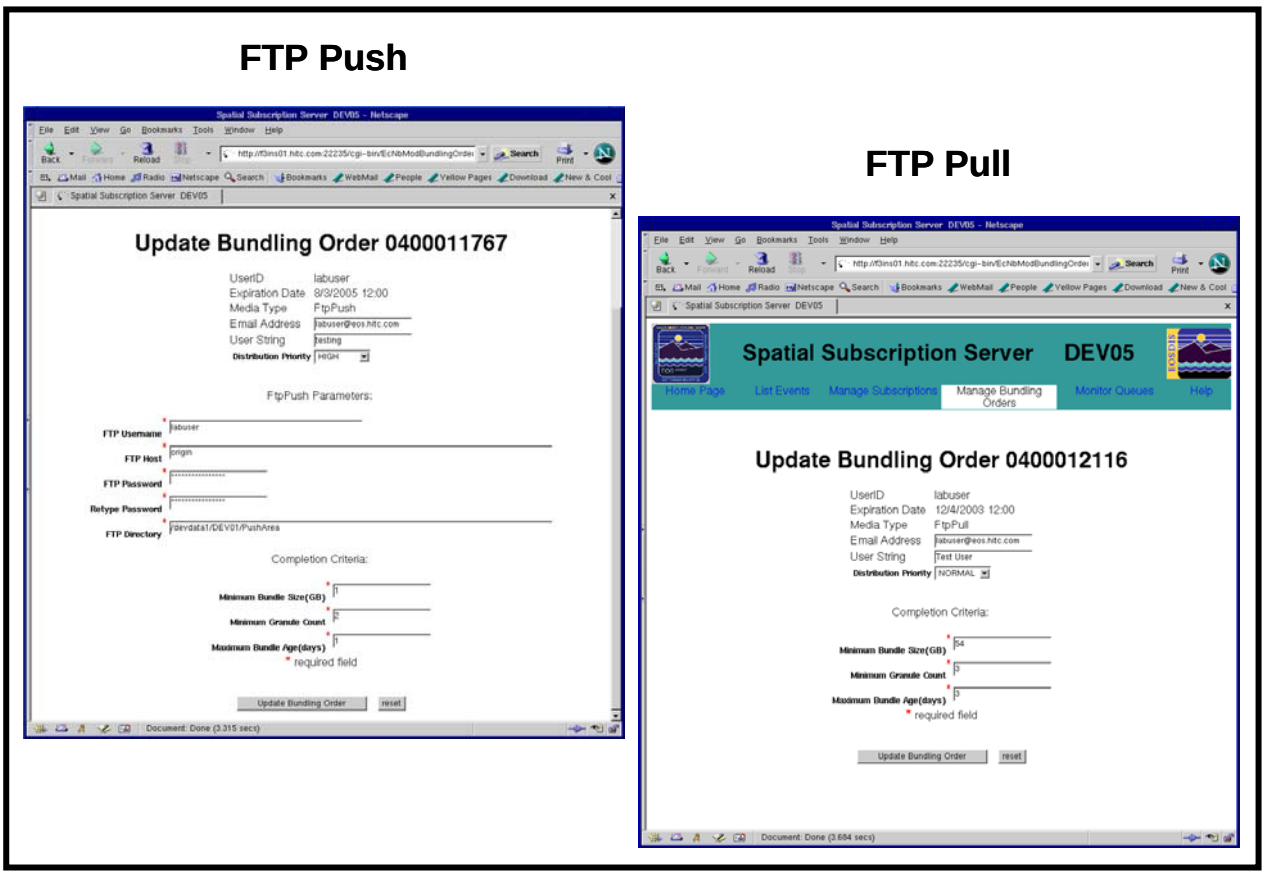


Figure 77. Spatial Subscription Server GUI Data Pages for Updating Bundling Order

- Maximum Bundle Age (days).

In addition to the links for listing and adding bundling orders, there is a link on the **Manage Bundling Orders** link to Configure Defaults. This link provides access to the page illustrated in Figure 78, permitting the operator to set default values for bundle completion criteria. These defaults set the values that appear at the bottom of the data page for each media type. As shown in Figure 78, the GUI displays the current values for the minimum granule count and maximum bundle age (days). It also shows the current value for minimum bundle size (GB) for each media and FTP type. The other parameter shown, Bundle Expiration Period (days), refers to the expiration period for the bundling order. There is a **Change to . . .** data entry field for each parameter. The figure also shows an example of the confirmation dialog box that appears when an operator makes an entry in one or more of the data entry fields and then clicks on the **Change Bundling Criteria** button at the bottom. There is also a **Reset** button to clear the data entry fields.

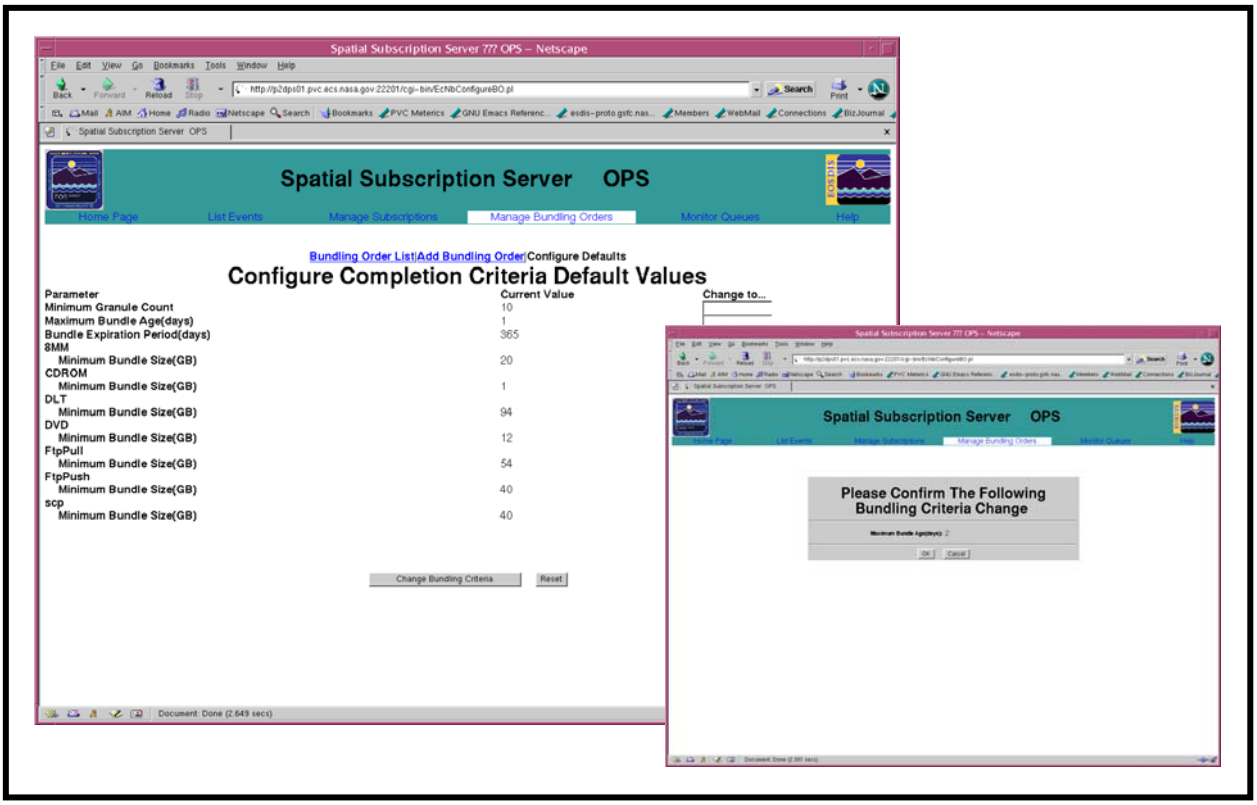
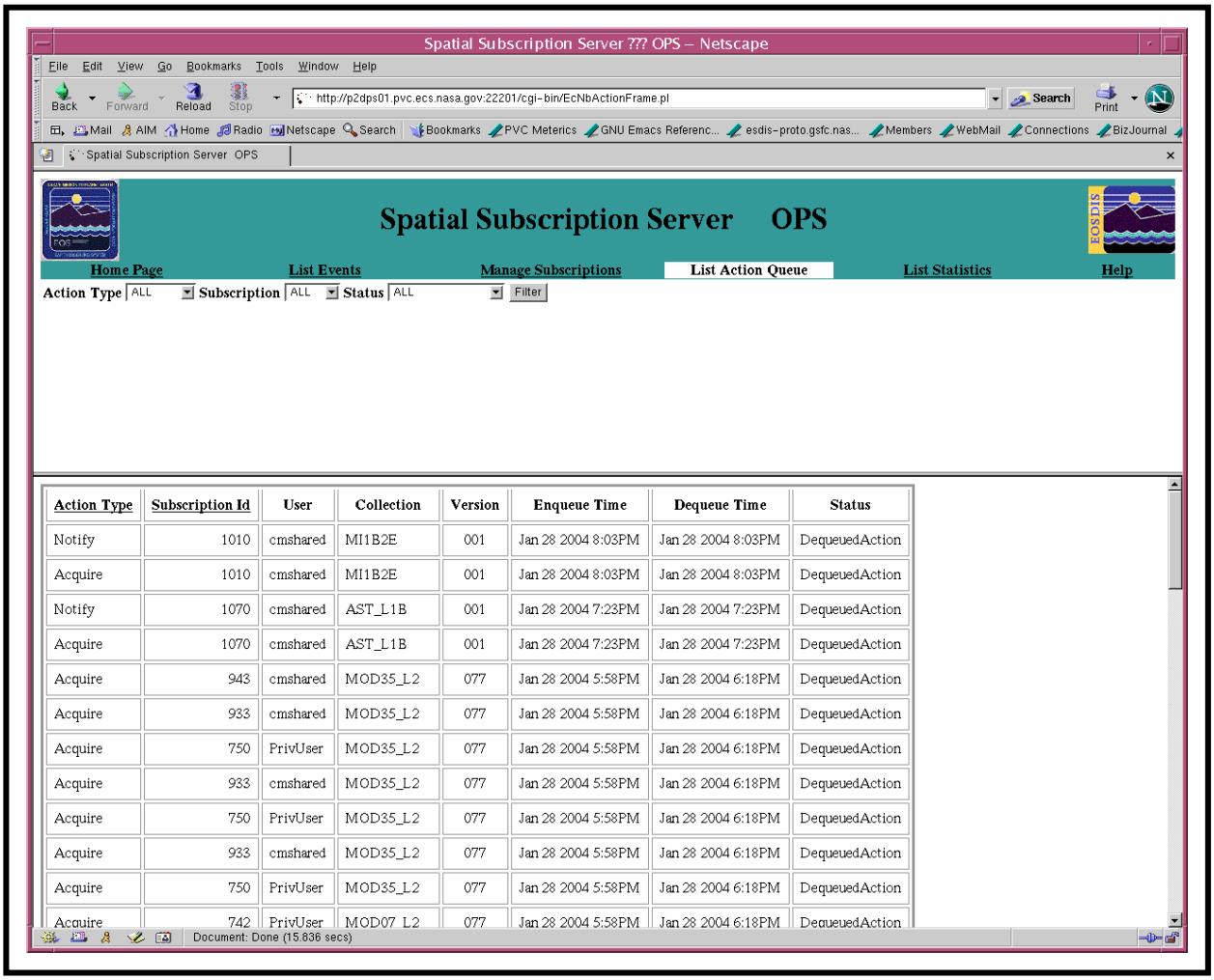


Figure 78. Spatial Subscription Server GUI Page for Configuring Bundling Order Completion Criteria Default Values

There are two other major pages of the **Spatial Subscription Server GUI** accessible from the **Monitor Queues** link on the **Home Page**. The **List Action Queue** page, illustrated in Figure 79, provides a table listing acquire and notification actions that are being processed. On this page, the **Action Type** and **Subscription Id** column headers are links for sorting the list, and there are also **Action Type**, **Subscription**, and **Status** option buttons and a filter button for filtering the list.

The **List Statistics** page, shown in Figure 80, provides summary information concerning the processing of events and actions related to subscriptions. Using this page, the operator can monitor subscription processing activity, such as numbers of notifications and actions, total and average times for notifications and acquires, and other information.

The **List Failed Action** page (Figure 81) provides information concerning failed actions. The page has buttons that the operator can use to remove failed actions.



Action Type	Subscription Id	User	Collection	Version	Enqueue Time	Dequeue Time	Status
Notify	1010	cmshared	MI1B2E	001	Jan 28 2004 8:03PM	Jan 28 2004 8:03PM	DequeuedAction
Acquire	1010	cmshared	MI1B2E	001	Jan 28 2004 8:03PM	Jan 28 2004 8:03PM	DequeuedAction
Notify	1070	cmshared	AST_L1B	001	Jan 28 2004 7:23PM	Jan 28 2004 7:23PM	DequeuedAction
Acquire	1070	cmshared	AST_L1B	001	Jan 28 2004 7:23PM	Jan 28 2004 7:23PM	DequeuedAction
Acquire	943	cmshared	MOD35_L2	077	Jan 28 2004 5:58PM	Jan 28 2004 6:18PM	DequeuedAction
Acquire	933	cmshared	MOD35_L2	077	Jan 28 2004 5:58PM	Jan 28 2004 6:18PM	DequeuedAction
Acquire	750	PrivUser	MOD35_L2	077	Jan 28 2004 5:58PM	Jan 28 2004 6:18PM	DequeuedAction
Acquire	933	cmshared	MOD35_L2	077	Jan 28 2004 5:58PM	Jan 28 2004 6:18PM	DequeuedAction
Acquire	750	PrivUser	MOD35_L2	077	Jan 28 2004 5:58PM	Jan 28 2004 6:18PM	DequeuedAction
Acquire	933	cmshared	MOD35_L2	077	Jan 28 2004 5:58PM	Jan 28 2004 6:18PM	DequeuedAction
Acquire	750	PrivUser	MOD35_L2	077	Jan 28 2004 5:58PM	Jan 28 2004 6:18PM	DequeuedAction
Acquire	742	PrivUser	MOD07_L2	077	Jan 28 2004 5:58PM	Jan 28 2004 6:18PM	DequeuedAction

Figure 79. Spatial Subscription Server GUI List Action Queue Page

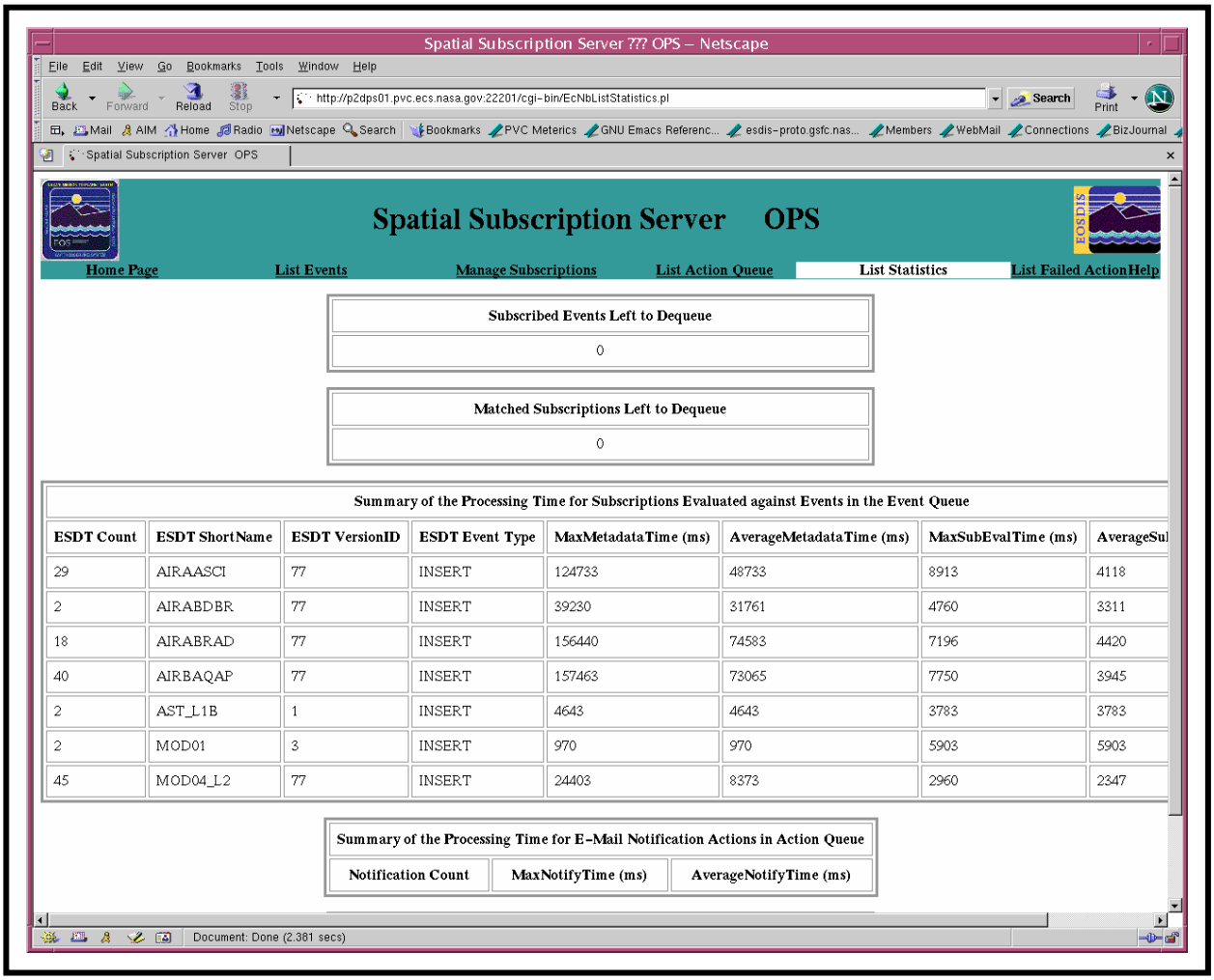


Figure 80. Spatial Subscription Server GUI List Statistics Page

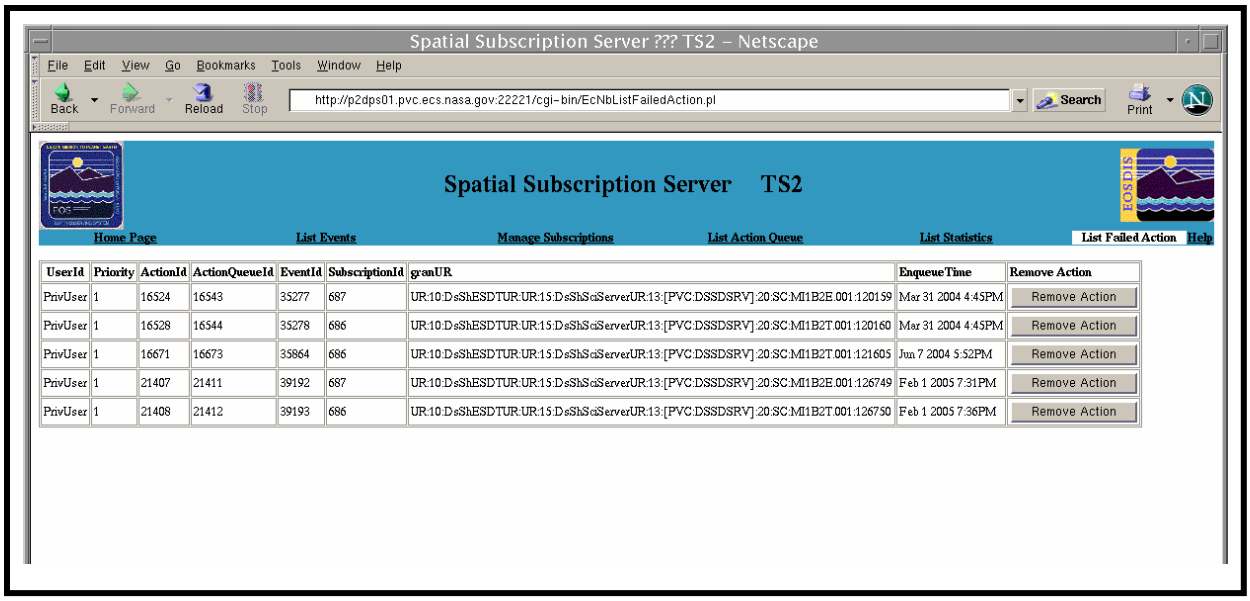


Figure 81. Spatial Subscription Server GUI List Failed Action Page

Procedures for Using the Spatial Subscription Server GUI

Launching and Shutting Down the Spatial Subscription Server GUI

Let's examine how the **Spatial Subscription Server GUI** is used for managing subscriptions. 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.

Launch the Spatial Subscription Server GUI

1 At the UNIX command shell prompt, type **setenv DISPLAY *clientname*:0.0** and then press the **Return/Enter** key.

- For ***clientname***, use either the local terminal/workstation IP address or its machine name.

2 Start the log-in to a Netscape host by typing **/tools/bin/ssh *hostname*** (e.g., e0ins01, g0ins01, l0ins01, or n0ins02) at the UNIX command shell prompt, and press the **Return/Enter** key.

- 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)?** type **yes** ("**y**" alone does 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, type your ***Passphrase*** and then press the **Return/Enter** key. Go to Step 5.

4 At the **<*user@remotehost*>'s password:** prompt, type your ***Password*** and then press the **Return/Enter** key.

- You are logged in and a UNIX command shell prompt is displayed.

5 Type **netscape &** then press **Return/Enter**.

- It may be necessary to type the path as well as the netscape command (e.g., **/tools/bin/netscape &**).
- 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 Netscape web browser (Figure 29) is displayed.

6 If a bookmark has been created for the **Spatial Subscription Server GUI**, select the appropriate bookmark from those listed on the browser's **Bookmarks** button (or the **Communicator → Bookmarks** pull-down menu).

- The security login **Prompt** (Figure 30) is displayed.

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

- For example:

  http://x0dps01.daac.ecs.nasa.gov:54321/cgi-bin/EcNbHomePage.pl

- The security login **Prompt** (Figure 30) is displayed.

8 Type the appropriate user name in the **User Name** box of the security login **Prompt**.

9 Type the appropriate password in the **Password** box of the security login **Prompt**.

NOTE: If the security login prompt reappears after the first time the user name and password have been entered (and the **OK** button has been clicked), it may not be due to a data entry problem. Try again to log in using the same user name and password. Sometimes it is necessary to enter the user name and password for the GUI more than once.

10 Click on the appropriate button from the following selections:

- **OK -** to complete the log-in and dismiss the dialogue box.
  - The dialogue box is dismissed.
  - The **Spatial Subscription Server GUI Home Page** (Figure 61) is displayed.
- **Cancel -** to dismiss the dialogue box without logging in.
  - The dialogue box is dismissed.

– The Netscape web browser (Figure 29) is displayed.

At some point it becomes necessary to shut down the **Spatial Subscription Server GUI** (end a **Spatial Subscription Server GUI** session). The procedure that follows is recommended and is applicable to both full-capability and limited-capability operators.

Shut Down the Spatial Subscription Server GUI (End a Spatial Subscription Server GUI Session)

---

1 Click on the **Home Page** link at the top of the **Spatial Subscription Server GUI**.

- The **Home Page** is displayed.

2 Click on the **End Session** link at the top of the **Spatial Subscription Server GUI**.

- A log-out page containing the following message is displayed.

  Click on Button Below to End Session: NOTE: THIS WOULD ALSO SHUT DOWN THE BROWSER :

NOTE: To abort the log-out and return to the **Home Page**, click on the browser **Back** button.

3 Click on the **ShutDown** button.

- The Netscape browser is dismissed.

Listing Subscribable Events and Subscriptions

Suppose you want to obtain and examine a list of events for which subscriptions may be established, and to find insert events for MOD01 data. Use the following procedure, which is applicable to both full-capability and limited-capability operators.

List Subscribable Events Using the Spatial Subscription Server GUI

---

1 Launch the **Spatial Subscription Server GUI**.

- For detailed instructions refer to the **Launch the Spatial Subscription Server GUI** procedure (previous section of this lesson).
- The **Home Page** is the default display, offering links for access to Spatial Subscription Server function pages (i.e., **List Events**, **Manage Subscriptions**, **Manage Bundling Orders**, **Monitor Queues**, **Help**, and **End Session**).

2 Click on the **List Events** link.

- The **List Events** page is displayed with a table of all events for which a subscription can be created.

3 Observe information displayed on the **Manage Subscriptions** page.

- The table on the **List Events** page has columns containing the following types of subscription information:
  - **Collection.**
  - **Version.**
  - **Event Type.**
- The column headers in the table are links for sorting the list. There are also buttons for filtering the list.
- There are option lists for filtering the table data by **Collection**, **Version**, and/or **Event Type**.
  - Event Type options include **ALL, DELETE, INSERT, UNDELETE, DELETEPHYSICAL,** and **UPDATEMETADATA**.

4 To **sort** the list click on the appropriate column header link (i.e., **Collection**, **Version**, or **Event Type**).

- If the list were sorted by **Event Type**, it would display events grouped by type: **DELETE** events would be listed first, followed by **INSERT** events, etc.

5 To **filter** the list to display certain types of information only, click on the appropriate option button (i.e., **Collection**, **Version**, or **Event Type**) and then click to select the desired option from the option list.

- The selected choice is displayed in the option field.

6 Repeat Step 5 to select an additional filter, if applicable.

7 Click on the **Filter** button to implement the selected filter.

- A list is displayed showing MOD01 events only.

8 Return to Step 3.

Suppose that you now want to view subscriptions related to MOD01 data, and then to review detailed information about a particular subscription. To obtain the list you will need to use the **Manage Subscriptions** link, filter on **Collection**, and then select and apply an option to **View** the particular subscription. Use the following procedure, which is applicable to both full-capability and limited-capability operators.

List and View Subscriptions in the NBSRV Database Using the Spatial Subscription Server GUI

---

1 Launch the **Spatial Subscription Server GUI**

- For detailed instructions refer to the **Launch the Spatial Subscription Server GUI** procedure (previous section of this lesson).
- The **Home Page** is the default display, offering links for access to Spatial Subscription Server function pages (i.e., **List Events**, **Manage Subscriptions**, **Manage Bundling Orders**, **Monitor Queues**, **Help**, and **End Session**).

2 Click on the **Manage Subscriptions** link.

- The **Manage Subscriptions** page is displayed with a table listing all subscriptions in the NBSRV database.

3 Observe information displayed on the **Manage Subscriptions** page.

- The table on the **Manage Subscriptions** page has columns containing the following types of subscription information:
  - **Subscription Id.**
  - **User.**
  - **Collection.**
  - **Version.**

- **Event Type.**
- **Status.**
- **Data Pool.**
- **Start Date.**
- **Expiration Date.**
- **Time Last Updated.**
- **Choose Subscription Action.**

- The column headers in the table, except for **Version**, **Event Type**, **Start Date**, and **Choose Subscription Action** are links for sorting the list.
- The **Choose Subscription Action** column has radio buttons for taking the following actions with respect to the corresponding subscription:
  - **View.**
  - **Update.**
  - **Cancel.**
- There are option lists for filtering the table data by **User**, **Collection**, **Status**, and/or **Data Pool**.
- There is a "**Change Collection Display to**" option button for switching the display between **TextEntry** and **SelectableList**.

4 To filter the list to display certain types of subscriptions only, click on the appropriate option button (i.e., **User**, **Collection**, **Status**, or **Data Pool**) and then click to select the desired option from the option list.

- The selected choice is displayed in the option field.

5 Repeat Step 4 to select an additional filter, if applicable.

6 Click on the **Filter** button to implement the selected filter.

- A list is displayed showing subscriptions that meet the filter criteria only.

7 To view a particular subscription first click on the **View** radio button in the **Choose Subscription Action** column for the subscription.

- The button is filled to indicate selection of the option.

8 To implement the selected action with respect to the particular subscription click on the **Apply** button in the **Choose Subscription Action** column for the subscription.

- A **View Subscriptions** page is displayed.

9 Observe information displayed on the **View Subscriptions** page.

- The table on the **Manage Subscriptions** page has columns containing the following types of subscription information:
  - **User.**
  - **Status.**
  - **Start Date.**
  - **Expiration Date.**
  - **Short Name.**
  - **Version.**
  - **Event Type.**
  - **Retention Period** (if applicable).
  - **Retention Priority** (if applicable).
  - **Science Granules and/or Metadata** (if applicable).
  - **Data Pool Action is associated with theme *x*** (if applicable).
  - etc.
- There are **View another subscription** and **Return to Home Page** links on the page.

10 Click on the **View another subscription** link or the **Return to Home Page** link (as appropriate) when finished viewing the subscription data on the **View Subscriptions** page.

11 Repeat Steps 3 through 10 as necessary to view additional subscriptions.

Adding a Subscription

A user who wants to have data made available for downloading through the Data Pool must contact the DAAC to request that a subscription be placed so that when data of the specified type are inserted in ECS, a copy is inserted in the Data Pool. User Services or Science Data Specialists then create the subscription, adding Data Pool qualification and retention criteria to meet the user's requirement within any constraints imposed for the Data Pool at the DAAC.

Suppose that a registered user requests Data Pool insert for MOD04_L2 003 data, specifying that the data granules inserted during the next two years be kept in the Data Pool for a period of 30 days. The user requests granules from daytime collection, in a geographic area bounded by 0 degrees North Latitude, -50 degrees East Longitude, -20 degrees South Latitude, and -70 degrees West Longitude. The user also requests notification as well as the Data Pool insert. Full-capability operators (only) can use the following procedure to create the necessary subscription.

Add a Subscription to the NBSRV Database Using the Spatial Subscription Server GUI

---

1 Launch the **Spatial Subscription Server GUI**.

- For detailed instructions refer to the **Launch the Spatial Subscription Server GUI** procedure (previous section of this lesson).
- The **Home Page** is the default display, offering links for access to Spatial Subscription Server function pages (i.e., **List Events**, **Manage Subscriptions**, **Manage Bundling Orders**, **Monitor Queues**, **Help**, and **End Session**).

NOTE: At various points in this procedure, you may encounter a security information warning. Unless you know of a specific potential danger that you must avoid, click on the **Continue submission** button when the warning is displayed.

2 Click on the **Manage Subscriptions** link.

- The **Manage Subscriptions** page is displayed, providing a table of subscription information showing 11 columns: **Subscription Id**, **User**, **Collection**, **Version**, **Event Type**, **Status**, **Data Pool**, **Start Date**, **Expiration Date**, **Time Last Updated**, and **Choose Subscription Action** (containing radio buttons for selecting an action to take and **Apply** buttons for implementing the selected actions).

3 Click on the **Add Subscriptions** link.

- The **Add Subscriptions** page is displayed.

4 Type the User Id for the requesting user in the **User Id** text entry field.

- The entered User Id must be a valid registered user (i.e., must be listed in the User Profile database).
- The typed entry is displayed in the **User Id** field.

5 To select a **Status** option click on the appropriate choice from the option list.

- **Active**.
- **Inactive**.

6 If the start date for the subscription is different from the default (current date and time) type the appropriate date and time in the **Start Date (mm/dd/yyyy hh:mm)** text entry box.

- The date and time should be entered in ***mm/dd/yyyy hh:mm*** format

7 If the expiration date for the subscription is different from the default (one year from the current date and time) type the appropriate date and time in the **Expiration Date (mm/dd/yyyy hh:mm)** text entry box.

- The date and time should be entered in ***mm/dd/yyyy hh:mm*** format

8 Type the first few characters of the name of the collection for which the subscription is to be created (e.g., **MOD04**) in the **Collection Short Name** text entry box.

9 Click on the **Apply** button.

- A **SELECT Short Name/Version/Event Type** option button and a **Numbers of String Qualifiers** button are displayed, along with an **Apply** button.

10 To select a **Short Name/Version/Event Type** option click on the appropriate choice (e.g., **MOD04_L2 003 INSERT**) from the option list.

11 To select a **Numbers of String Qualifiers** option click on the appropriate choice (e.g., **1**) from the option list.

12 Click on the **Apply** button.

- Option buttons and fields are displayed to permit the entry of data in two general categories; i.e., **Subscription Qualifiers** and **Action Information**.
- In the **Subscription Qualifiers** section it is possible to select **Attribute Name/Type** and enter **Min Value** and **Max Value** for any valid **Integer/Float/Date** qualifiers, string **Value** qualifiers, and **Latitude** and **Longitude** coordinates to define a bounding rectangle spatial qualifier for the subscription to be created.
- In the **Action Information** section there is a button to specify that the subscription is to be associated with a bundling order; if that button is not selected, one has the choice of one or more of three actions (i.e., **Acquire**, **E-Mail Notification**, or **Data Pool**). There are blocks with option lists and text entry fields for defining the selected action(s) to be taken upon occurrence of the event for which the subscription is to be created.
- At the bottom of the page is an **Add Subscription** button for submitting the subscription.

13 To select a different **Logical Basis** option (if applicable) click on the appropriate choice (e.g., **AND**) from the option list.

14 To select an integer, float, or date **Attribute Name/Type** option (if applicable) click on the appropriate choice from the **Integer/Float/Date Attribute Name/Type** option list.

15 To specify a minimum value for an integer, float, or date attribute (if applicable) type the appropriate value in the **Integer/Float/Date Min Value** text entry box.

16 To specify a maximum value for an integer, float, or date attribute (if applicable) type the appropriate value in the **Integer/Float/Date Max Value** text entry box.

17 Repeat Steps 14 through 16 as necessary to specify additional integer, float, or date attributes.

18 To select a string **Attribute Name/Type** option (if applicable) click on the appropriate choice from the **String Attribute Name/Type** option list.

19 To specify a value for a string attribute (if applicable) type the appropriate value in the **String Value** text entry box.

20 Repeat Steps 18 and 19 as necessary to specify additional string attributes.

21 To specify spatial coordinates of intersecting LLBox (if applicable) type the appropriate values in the following text entry boxes.

- **North Latitude.**
- **West Longitude.**
- **East Longitude.**
- **South Latitude.**
- For example:

  North Latitude 0

  West Longitude -70

  East Longitude -50

  South Latitude -20

NOTE: Every subscription must have at least one action specified and may have more than one.

22 To associate the subscription with a bundling order (if applicable) first click in the **Bundling Order** box.

23 To continue the process of associating a subscription with a bundling order (if applicable) click on the appropriate choice from the **Bundling Order Selection** option list.

- If a subscription is associated with a bundling order, skip Steps 24 through 27 and go to Step 28.

24 To select subscription action(s) click in the following boxes as applicable (a check mark in a box indicates that the action has been selected):

- **Acquire** – to request shipment of the data specified in the subscription.
- **E-Mail Notification** - to request e-mail notification of the event specified in the subscription.
- **Data Pool** - to request insertion of data specified in the subscription into the Data Pool.

NOTE: Multiple actions can be selected.

25 If **Acquire** was selected in Step 24, either type the appropriate data or click on the appropriate choice from the option list in the relevant fields of the **Acquire Information** block:

- **User Profile.**
- **User String** [if applicable (to distinguish the subscription from others)].
- **Email Address**.
- **Media Type**.
- **Priority**.
- **User** [applicable to ftp push or secure copy distributions only].
- **Password** [applicable to ftp push or secure copy distributions only].
- **Enter password again for verification** [applicable to ftp push or secure copy distributions only].
- **Host** [applicable to ftp push or secure copy distributions only].
- **Directory** [applicable to ftp push or secure copy distributions only].

26 If **E-Mail Notification** was selected in Step 24, either type the appropriate data or click on the appropriate choice from the option list in the relevant fields of the **E-Mail Notification Information** block:

- **Action Address.**
- **User String** [if applicable (to distinguish the subscription from others)].
- **Metadata.**

27 If **Data Pool** was selected in Step 24, either type the appropriate data or click on the appropriate choice from the option list in the relevant fields of the **Data Pool Information** block:

- **Retention Period (in days)** [e.g., **30**].
- **Retention Priority (valid range 1 thru 255)** [e.g., **200**].

- **Science Granules and/or Metadata** [e.g., **science and metadata**].
- **Click here to add theme:** [click in the box if the granules to be added to the Data Pool as a result of the subscription are to be associated with a theme].
- **Enter first few characters of name** [optional - if associating the granules from the subscription with a theme].

28 Click on the **Add Subscription** button.

- If the **Click here to add theme:** box was checked, a **Select Theme for Data Pool Action** page is displayed.
- If the **Click here to add theme:** box was not checked, a message is displayed confirming that "Subscription *x* was added" to the database and buttons permit **Add another subscription** or **Return to Home Page**.

29 To select a theme for Data Pool action (if applicable) first click on the appropriate choice from the **Select Theme for Data Pool Action** option list.

30 To make the theme association retroactive (if applicable) first click in the **To make theme association retroactive check here:** box.

31 To implement the theme association (if applicable) click on the **Apply** button on the **Select Theme for Data Pool Action** page.

- A message is displayed confirming that "Subscription *x* was updated/added" to the database and buttons permit **Add another subscription** or **Return to Home Page**.

32 Click on the appropriate link from the following selections:

- **Add… another subscription**.
  - The **Manage Subscriptions** page is displayed.
- **Update another subscription**.
  - The **Manage Subscriptions** page is displayed.
- **Return to Home Page**.

– The **Home Page** is displayed.

Working with Existing Subscriptions

On occasion, it may be necessary to modify or delete existing subscriptions in the NBSRV database. The **Spatial Subscription Server GUI** web access supports these requirements. However, only full-capability operators can modify or delete existing subscriptions using the **Spatial Subscription Server GUI**. Limited-capability operators can view subscriptions but may not modify or delete them.

It may become desirable to extend the period of retention in a Data Pool insert subscription. For example, if conditions arise that increase user interest in the data that will be inserted as a result of the subscription, keeping the data in the Data Pool longer is a way to be responsive to that user interest. Such conditions might be unusual weather activity, flooding, volcanic activity, or other events in the area covered by the subscription.

The update procedure for extending the period of retention in a Data Pool insert overlaps extensively with procedures for updating other information in an existing subscription, either for Data Pool insert (for example, to change qualifiers or to associate the subscription with a theme) or for notification/distribution of standard ECS products (for example, to change the distribution media type or other information defining the distribution). The operator accesses the subscription in the same way, makes the necessary inputs to specify the changes, and uses the **Update Subscription** button to submit the changes.

Update a Subscription in the NBSRV Database Using the Spatial Subscription Server GUI

1 Launch the **Spatial Subscription Server GUI**.

- For detailed instructions refer to the **Launch the Spatial Subscription Server GUI** procedure (previous section of this lesson).
- The **Home Page** is the default display, offering links for access to Spatial Subscription Server function pages (i.e., **List Events**, **Manage Subscriptions**, **Manage Bundling Orders**, **Monitor Queues**, **Help**, and **End Session**).

NOTE: At various points in this procedure, you may encounter a security information warning. Unless you know of a specific potential danger that you must avoid, click on the **Continue submission** button when the warning is displayed.

2 Click on the **Manage Subscriptions** link.

- The **Manage Subscriptions** page is displayed, providing a table of subscription information showing 11 columns: **Subscription Id**, **User**, **Collection**, **Version**, **Event Type**, **Status**, **Data Pool**, **Start Date**, **Expiration Date**, **Time Last Updated**, and **Choose Subscription Action** (containing radio buttons for selecting an action to take and **Apply** buttons for implementing the selected actions).

3 Find the subscription to be updated in the list of subscriptions on the **Manage Subscriptions** page.

- Scroll, sort, and/or filter the list as necessary.
  - For detailed instructions refer to the **List and View Subscriptions in the NBSRV Database Using the Spatial Subscription Server GUI** procedure (previous section of this lesson).

4 Click on the **Update** radio button in the **Choose Subscription Action** column for the subscription to be updated.

- The button is filled to indicate selection of the option.

5 To implement the "update" action with respect to the subscription click on the **Apply** button in the **Choose Subscription Action** column for the subscription.

- An **Update Subscriptions** page is displayed.

6 To change the User Id (if applicable) type the User Id for the requesting user in the **User Id** text entry field.

- The entered User Id must be a valid registered user (i.e., must be listed in the User Profile database).
- The typed entry is displayed in the **User Id** field.

7 To change the status (if applicable) click on the appropriate choice from the **Status** option list.

- **Active**.
- **Inactive**.

8 To change the start date for the subscription (if applicable) type the appropriate date and time in the **Start Date (mm/dd/yyyy hh:mm)** text entry box.

- The date and time should be entered in ***mm/dd/yyyy hh:mm*** format

9 To change the expiration date for the subscription (if applicable) type the appropriate date and time in the **Expiration Date (mm/dd/yyyy hh:mm)** text entry box.

- The date and time should be entered in ***mm/dd/yyyy hh:mm*** format

NOTE: It is not possible to update the subscription event [**Short Name/Version/Event Type** (e.g., **MOD04_L2 003 INSERT**)].

NOTE: Option buttons and fields are displayed to permit the entry of data in two general categories; i.e., **Subscription Qualifiers** and **Action Information**. In the **Subscription Qualifiers** section it is possible to select **Attribute Name/Type** and enter **Min Value** and **Max Value** for any valid **Integer/Float/Date** qualifiers, string **Value** qualifiers, and **Latitude** and **Longitude** coordinates to define a bounding rectangle spatial qualifier for the subscription to be created. In the **Action Information** section there is a button to specify that the subscription is to be associated with a bundling order; if that button is not selected, one has the choice of one or more of three actions (i.e., **Acquire**, **E-Mail Notification**, or **Data Pool**). There are blocks with option lists and text entry fields for defining the selected action(s) to be taken upon occurrence of the event for which the subscription was created.

10 To select a different **Logical Basis** option (if applicable) click on the appropriate choice (e.g., **AND**) from the option list.

11 To change, add, or delete an integer, float, or date **Attribute Name/Type** option (if applicable) click on the appropriate choice from the **Integer/Float/Date Attribute Name/Type** option list.

12 To change, add, or delete a minimum value for an integer, float, or date attribute (if applicable) type (or delete, if necessary) the appropriate value in the **Integer/Float/Date Min Value** text entry box.

13 To change, add, or delete a maximum value for an integer, float, or date attribute (if applicable) type (or delete, if necessary) the appropriate value in the **Integer/Float/Date Max Value** text entry box.

14 Repeat Steps 11 through 13 as necessary to change, add, or delete integer, float, or date attributes.

15 To change, add, or delete a string **Attribute Name/Type** option (if applicable) click on the appropriate choice from the **String Attribute Name/Type** option list.

16 To change, add, or delete a value for a string attribute (if applicable) type (or delete, if necessary) the appropriate value in the **String Value** text entry box.

17 Repeat Steps 15 and 16 as necessary to change or specify additional string attributes.

18 To change, add, or delete spatial coordinates of intersecting LLBox (if applicable) type (or delete, if necessary) the appropriate values in the following text entry boxes (as necessary).

- **North Latitude.**
- **West Longitude.**
- **East Longitude.**
- **South Latitude.**

NOTE: Every subscription must have at least one action specified and may have more than one.

19 To associate the subscription with a bundling order (if applicable) first click in the **Bundling Order** box.

20 To continue the process of associating a subscription with a bundling order (if applicable) click on the appropriate choice from the **Bundling Order Selection** option list.

- If a subscription is associated with a bundling order, skip Steps 21 through 24 and go to Step 25.

21 To select or deselect subscription action(s) click in the following boxes as applicable (a check mark in a box indicates that the action has been selected):

- **Acquire** – to request shipment of the data specified in the subscription.
- **E-Mail Notification** - to request e-mail notification of the event specified in the subscription.
- **Data Pool** - to request insertion of data specified in the subscription into the Data Pool.

NOTE: Multiple actions can be selected.

22 If **Acquire** was selected in Step 21, change or add shipment data (if applicable) by either typing the appropriate data or clicking on the appropriate choice from the option list in the relevant fields of the **Acquire Information** block:

- **User Profile.**
- **User String** [if applicable (to distinguish the subscription from others)].
- **Email Address**.
- **Media Type**.
- **Priority**.
- **User** [applicable to ftp push or secure copy distributions only].
- **Password** [applicable to ftp push or secure copy distributions only].
- **Enter password again for verification** [applicable to ftp push or secure copy distributions only].
- **Host** [applicable to ftp push or secure copy distributions only].
- **Directory** [applicable to ftp push or secure copy distributions only].

23 If **E-Mail Notification** was selected in Step 21, change or add e-mail notification data by either typing the appropriate data or clicking on the appropriate choice from the option list in the relevant fields of the **E-Mail Notification Information** block:

- **Action Address.**
- **User String** [if applicable (to distinguish the subscription from others)].
- **Metadata.**

24 If **Data Pool** was selected in Step 21, change or add Data Pool insertion data either by typing the appropriate data or clicking on the appropriate choice from the option list in the relevant fields of the **Data Pool Information** block:

- **Retention Period (in days)**.
- **Retention Priority (valid range 1 thru 255)**.

- **Science Granules and/or Metadata**.
- **Click here to add theme:** [click in the box if the granules to be added to the Data Pool as a result of the subscription are to be associated with a theme].
- **Enter first few characters of name** [optional - if associating the granules from the subscription with a theme].

25 Click on the **Update Subscription** button.

- If the **Click here to add theme:** box was checked, a **Select Theme for Data Pool Action** page is displayed.
- If the **Click here to add theme:** box was not checked, a message is displayed confirming that "Subscription *x* was updated" to the database and buttons permit **Update another subscription** or **Return to Home Page**.

26 To select a theme for Data Pool action (if applicable) first click on the appropriate choice from the **Select Theme for Data Pool Action** option list.

27 To make the theme association retroactive (if applicable) first click in the **To make theme association retroactive check here:** box.

28 To implement the theme association (if applicable) click on the **Apply** button on the **Select Theme for Data Pool Action** page.

- A message is displayed confirming that "Subscription *x* was updated/added" to the database and buttons permit **Add another subscription** or **Return to Home Page**.

29 Click on the appropriate link from the following selections:

- **Add… another subscription**.
  - The **Manage Subscriptions** page is displayed.
- **Update another subscription**.
  - The **Manage Subscriptions** page is displayed.
- **Return to Home Page**.

– The **Home Page** is displayed.

If it is necessary to delete (cancel) a subscription, the **Spatial Subscription Server GUI** provides the means. Suppose Dr. L. Abuser contacts User Services with a request to cancel the subscription for ASTER expedited data. A full-capability operator (only) can use the following procedure to respond to such a request.

Delete (Cancel) a Subscription in the NBSRV Database Using the Spatial Subscription Server GUI

---

1 Launch the **Spatial Subscription Server GUI**.

- For detailed instructions refer to the **Launch the Spatial Subscription Server GUI** procedure (previous section of this lesson).
- The **Home Page** is the default display, offering links for access to Spatial Subscription Server function pages (i.e., **List Events**, **Manage Subscriptions**, **Manage Bundling Orders**, **Monitor Queues**, **Help**, and **End Session**).

2 Click on the **Manage Subscriptions** link.

- The **Manage Subscriptions** page is displayed, providing a table of subscription information showing 11 columns: **Subscription Id**, **User**, **Collection**, **Version**, **Event Type**, **Status**, **Data Pool**, **Start Date**, **Expiration Date**, **Time Last Updated**, and **Choose Subscription Action** (containing radio buttons for selecting an action to take and **Apply** buttons for implementing the selected actions).

3 Observe information displayed on the **Manage Subscriptions** page.

- The table on the **Manage Subscriptions** page has columns containing the following types of subscription information:
  - **Subscription Id.**
  - **User.**
  - **Collection.**
  - **Version.**
  - **Event Type.**
  - **Status.**
  - **Data Pool.**
  - **Start Date.**
  - **Expiration Date.**
  - **Time Last Updated.**
  - **Choose Subscription Action.**

- The column headers in the table, except for **Version**, **Event Type**, **Start Date**, and **Choose Subscription Action** are links for sorting the list.
- The **Choose Subscription Action** column has radio buttons for taking the following actions with respect to the corresponding subscription:
  - **View.**
  - **Update.**
  - **Cancel.**
- There are option lists for filtering the table data by **User**, **Collection**, **Status**, and/or **Data Pool**.
- There is a "**Change Collection Display to**" option button for switching the display between **TextEntry** and **SelectableList**.

4 To filter the list to display certain types of subscriptions only, click on the appropriate option button and then click to select the desired option from the option list.

- The selected choice is displayed in the option field.

5 Repeat Step 4 to select an additional filter, if applicable.

6 Click on the **Filter** button to implement the selected filter.

- A list is displayed showing subscriptions that meet the filter criteria only.

7 Click on the **Cancel** radio button in the **Choose Subscription Action** column for the subscription to be canceled (deleted).

- The button is filled to indicate selection of the option.

8 Click on the **Apply** button in the **Choose Subscription Action** column for the subscription to be canceled (deleted).

- A confirmation page is displayed with the message "Are you sure that you want to cancel subscription *x*?"

9 Click on the appropriate link from the following selections:

- **Yes**- to confirm the deletion.
  - A "Subscription x was canceled" message is displayed.
- **No**- to abort the deletion.

10 Click on the appropriate link from the following selections:

- **Cancel another subscription**.
  - The **Manage Subscriptions** page is displayed.

- **Return to Home Page**.
  - The **Home Page** is displayed.

Both full-capability and limited-capability operators can use the following procedure to list subscriptions associated with a theme.

List Subscriptions Associated with a Theme

---

1 Launch the **Spatial Subscription Server GUI**.

- For detailed instructions refer to the **Launch the Spatial Subscription Server GUI** procedure (previous section of this lesson).
- The **Home Page** is the default display, offering links for access to Spatial Subscription Server function pages (i.e., **List Events**, **Manage Subscriptions**, **Manage Bundling Orders**, **Monitor Queues**, **Help**, and **End Session**).

2 Click on the **Manage Subscriptions** link.

- The **Manage Subscriptions** page is displayed with a table listing all subscriptions in the NBSRV database.

3 Click on the **List Themes** link.

- The **List Themes Request** page is displayed.
- An "Enter first few characters of theme name (or leave blank to view all):" message is displayed on the **List Themes Request** page.

4 To display all themes, click on the **Apply** button (without entering anything in the text box).

- The **Theme List** page is displayed with columns containing the following types of theme information:
  - **Theme Name.**
  - **Choose Theme Action.**
- All themes are displayed on the **Theme List** page.
- Go to Step 7.

5 To begin the process of displaying a particular theme, type the first few letters of the theme name in the text box on the **List Themes Request** page.

6 To continue the process of displaying a particular theme, click on the **Apply** button on the **List Themes Request** page.

- The **Theme List** page is displayed with columns containing the following types of theme information:
  - **Theme Name.**
  - **Choose Theme Action.**
- The relevant theme(s) is (are) displayed on the **Theme List** page.

7 To begin the process of viewing the list of subscriptions associated with a particular theme, click in the corresponding check box (in the **Choose Theme Action** column) on the **Theme List** page.

8 To view a list of the subscriptions associated with the specified theme, click on the corresponding **Apply** button (in the **Choose Theme Action** column) on the **Theme List** page.

- The **List Subscriptions for Theme** page is displayed.
  - The subscription(s) associated with the specified theme is (are) displayed.
  - Buttons are available for viewing, updating, or canceling each subscription being displayed or suspending, resuming, or canceling all subscriptions. (Refer to the applicable procedure for instructions on performing any of those actions.)

9 Observe information displayed on the **List Subscriptions for Theme** page.

- The table on the **List Subscriptions for Theme** page has columns containing the following types of subscription information:
  - **ID.**
  - **User.**
  - **Status.**
  - **Expiration Date.**
  - **ESDT ShortName/Version/EventType.**
  - **Choose Subscription Action.**
- The **Choose Subscription Action** column has radio buttons for taking the following actions with respect to the corresponding subscription:
  - **View.**
  - **Update.**
  - **Cancel.**

- At the bottom of the table there are radio buttons for selecting the following actions with respect to all listed subscriptions:
  - **Suspend All.**
  - **Resume All.**
  - **Cancel All.**

10 To take action with respect to a particular subscription first click on the appropriate radio button (i.e., **View**, **Update**, or **Cancel**) in the **Choose Subscription Action** column for the subscription.

- The button is filled to indicate selection of the option.

11 To implement a selected action with respect to a particular subscription click on the **Apply** button in the **Choose Subscription Action** column for the subscription.

- If **View** was the selected action, a **View Subscriptions** page is displayed.
  - Click on the **View another subscription** link or the **Return to Home Page** link (as appropriate) when finished viewing the subscription data on the **View Subscriptions** page.
- If **Update** was the selected action, an **Update Subscriptions** page is displayed.
  - Go to the **Update a Subscription in the NBSRV Database Using the Spatial Subscription Server GUI** procedure (previous section of this lesson).
- If **Cancel** was the selected action, a confirmation page is displayed with the message "Are you sure that you want to cancel subscription *x*?"
  - Click on the **Yes** link to confirm the deletion.
  - Click on the **No** link to abort the deletion.

12 To take action with respect to all subscriptions associated with the theme first click on the appropriate radio button (i.e., **Suspend All**, **Resume All**, or **Cancel All**) at the bottom of the table on the **List Subscriptions for Theme** page.

- The button is filled to indicate selection of the option.

13 To implement a selected action with respect to all subscriptions associated with the theme click on the **Apply** button at the bottom of the table on the **List Subscriptions for Theme** page.

- If **Suspend All** was the selected action, an "All associated subscriptions have been suspended" message is displayed.
- If **Resume All** was the selected action, an "All associated subscriptions have been resumed" message is displayed.

- If **Cancel All** was the selected action, a confirmation page is displayed with the message "Are you sure that you want to cancel subscription *x*?"
  - Click on the **Yes** link to confirm the deletion.
  - Click on the **No** link to abort the deletion.

14 Click on the **View refreshed subscription list** link (if applicable).

- The **List Subscriptions for Theme** page is displayed.

Managing Bundling Orders Using the Spatial Subscription Server GUI

The **Spatial Subscription Server GUI** is used to list, view, create, update, cancel, and list associated subscriptions for bundling orders. These functions are accessible through the **Manage Bundling Orders** link. The limited-capability operator may view bundling orders and the subscriptions associated with them. The full-capability operator can perform those same functions and create, update, or cancel bundling orders as well.

Obtain a List of Bundling Orders and View a Bundling Order Using the Spatial Subscription Server GUI

---

1 Launch the **Spatial Subscription Server GUI**.

- For detailed instructions refer to the **Launch the Spatial Subscription Server GUI** procedure (previous section of this lesson).
- The **Home Page** is the default display, offering links for access to Spatial Subscription Server function pages (i.e., **List Events**, **Manage Subscriptions**, **Manage Bundling Orders**, **Monitor Queues**, **Help**, and **End Session**).

NOTE: At various points in this procedure, you may encounter a security information warning. Unless you know of a specific potential danger that you must avoid, click on the **Continue submission** button when the warning is displayed.

2 Click on the **Manage Bundling Orders** link.

- The **Manage Bundling Orders** page is displayed.

3 Observe information displayed on the **Manage Bundling Orders** page.

- The table on the **Manage Bundling Orders** page has columns containing the following types of bundling order information:
  - **Bundling Order.**
  - **User.**
  - **Creation Date.**

- **Expiration Date.**
- **Media Type.**
- **Status.**
- **Choose Bundling Order Action.**

- The column headers in the table, except for **Choose Bundling Order Action** are links for sorting the list.
- The **Choose Bundling Order Action** column has radio buttons for taking the following actions with respect to the corresponding bundling order:
  - **View.**
  - **Update.**
  - **Cancel.**
  - **List Subs** [list associated subscriptions].
- There are option lists for filtering the table data by **User**, **Media Type**, and/or **Status**.
- There are links on the page to **Add Bundling Order** and **Configure Defaults**.

4 To filter the list to display certain types of bundling orders only, click on the appropriate option button (i.e., **User**, **Media Type**, or **Status**) and then click to select the desired option from the option list.

- The selected choice is displayed in the option field.

5 Repeat Step 4 to select an additional filter, if applicable.

6 Click on the **Filter** button to implement the selected filter.

- A list is displayed showing subscriptions that meet the filter criteria only.

7 To view a particular bundling order first click on the **View** radio button in the **Choose Bundling Order Action** column for the bundling order.

- The button is filled to indicate selection of the option.

8 To implement the selected action click on the **Apply** button in the **Choose Bundling Order Action** column for the bundling order.

- A **View Bundling Order** page is displayed.

9 Observe information displayed on the **View Bundling Order** page.

- The **View Bundling Order** page displays the following types of bundling order information:
  - **Bundling Order ID.**
  - **User Name.**

- **Creation Date.**
- **Expiration Date.**
- **Media Type.**
- **Current Status.**
- **User String.**
- **Email Address.**
- **Distribution Priority.**
- **scp distribution information: User** (if applicable).
- **scp distribution information: Host** (if applicable).
- **scp distribution information: Directory** (if applicable).
- **Completion criteria: Minimum Bundle Size.**
- **Completion criteria: Minimum Granule Count.**
- **Completion criteria: Maximum Bundle Age.**

- There is a **Return to bundling order list** link.

10 Click on the **Return to bundling order list** link when finished viewing the bundling order data on the **View Bundling Order** page.

11 To view subscriptions associated with a particular bundling order first click on the **List Subs** radio button in the **Choose Bundling Order Action** column for the bundling order.

- The button is filled to indicate selection of the option.

12 To implement the selected action click on the **Apply** button in the **Choose Bundling Order Action** column for the subscription.

- A **Bundling Order Subscriptions (Subscriptions for bundling order *x*)** page is displayed.

13 Observe information displayed on the **Bundling Order Subscriptions (Subscriptions for bundling order *x*)** page.

- The **Bundling Order Subscriptions (Subscriptions for bundling order *x*)** page displays the following types of subscription information:
  - **ID.**
  - **User.**
  - **Status.**
  - **Expiration Date.**

- **ESDT ShortName/Version/EventType.**
- **Choose Subscription Action.**

- The **Choose Subscription Action** column has radio buttons for taking the following actions with respect to the corresponding subscription:
  - **View.**
  - **Update.**
  - **Cancel.**

14 To take action with respect to a particular subscription first click on the appropriate radio button (i.e., **View**, **Update**, or **Cancel**) in the **Choose Subscription Action** column for the subscription.

- The button is filled to indicate selection of the option.

15 To implement a selected action with respect to a particular subscription click on the **Apply** button in the **Choose Subscription Action** column for the subscription.

- If **View** was the selected action, a **View Subscriptions** page is displayed.
  - Click on the **View another subscription** link or the **Return to Home Page** link (as appropriate) when finished viewing the subscription data on the **View Subscriptions** page.
- If **Update** was the selected action, an **Update Subscriptions** page is displayed.
  - Go to the **Update a Subscription in the NBSRV Database Using the Spatial Subscription Server GUI** procedure (previous section of this lesson).
- If **Cancel** was the selected action, a confirmation page is displayed with the message "Are you sure that you want to cancel subscription *x*?"
  - Click on the **Yes** link to confirm the deletion.
  - Click on the **No** link to abort the deletion.

16 To return to the **View Bundling Order** page from the **Bundling Order Subscriptions (Subscriptions for bundling order *x*)** page click on the browser **Back** button.

17 Repeat Steps 3 through 16 as necessary to view additional bundling orders.

To cancel a bundling order and its associated subscriptions a full-capability operator (only) can use the following procedure:

Cancel a Bundling Order and Its Associated Subscriptions Using the Spatial Subscription Server GUI

---

1 Launch the **Spatial Subscription Server GUI**.

- For detailed instructions refer to the **Launch the Spatial Subscription Server GUI** procedure (previous section of this lesson).
- The **Home Page** is the default display, offering links for access to Spatial Subscription Server function pages (i.e., **List Events**, **Manage Subscriptions**, **Manage Bundling Orders**, **Monitor Queues**, **Help**, and **End Session**).

NOTE: At various points in this procedure, you may encounter a security information warning. Unless you know of a specific potential danger that you must avoid, click on the **Continue submission** button when the warning is displayed.

2 Click on the **Manage Bundling Orders** link.

- The **Manage Bundling Orders** page is displayed.

3 Observe information displayed on the **Manage Bundling Orders** page.

- The **Manage Bundling Orders** page is displayed, providing a table of bundling order information showing seven columns: **Bundling Order**, **User**, **Creation Date**, **Expiration Date**, **Media Type**, **Status**, and **Choose Bundling Order Action** (containing radio buttons for selecting an action to take and **Apply** buttons for implementing the selected actions).

4 To filter the list to display certain types of bundling orders only, click on the appropriate option button (i.e., **User**, **Media Type**, or **Status**) and then click to select the desired option from the option list.

- The selected choice is displayed in the option field.

5 Repeat Step 4 to select an additional filter, if applicable.

6 Click on the **Filter** button to implement the selected filter.

- A list is displayed showing subscriptions that meet the filter criteria only.

7 To cancel a particular bundling order first click on the **Cancel** radio button in the **Choose Bundling Order Action** column at the end of the row for the bundling order.

- The button is filled to indicate selection of the option.

8 To implement the selected action click on the **Apply** button in the **Choose Bundling Order Action** column for the bundling order.

- The following confirmation message is displayed:

  Note: Any associated subscriptions will also be cancelled. **Are you sure you wish to cancel bundling order <*nnnnnn*>?**

9 Click on the appropriate button from the following selections:

- **Yes -** to confirm cancellation of the bundling order and associated subscriptions.
  - A confirmation message displays: **Bundling order *x* has been cancelled.**
- **No -** to abort cancellation of the bundling order.

– The **Manage Bundling Orders** page is displayed.

10 Click on the **Return to bundling order list** link (if applicable).

- The **Manage Bundling Orders** page is displayed.

11 Repeat Steps 3 through 10 as necessary to cancel additional bundling orders.

To add a bundling order, a full-capability operator (only) can use the following procedure:

Add a Bundling Order Using the Spatial Subscription Server GUI

---

1 Launch the **Spatial Subscription Server GUI**.

- For detailed instructions refer to the **Launch the Spatial Subscription Server GUI** procedure (previous section of this lesson).
- The **Home Page** is the default display, offering links for access to Spatial Subscription Server function pages (i.e., **List Events**, **Manage Subscriptions**, **Manage Bundling Orders**, **Monitor Queues**, **Help**, and **End Session**).

NOTE: At various points in this procedure, you may encounter a security information warning. Unless you know of a specific potential danger that you must avoid, click on the **Continue submission** button when the warning is displayed.

2 Click on the **Manage Bundling Orders** link.

- The **Manage Bundling Orders** page is displayed, providing a table of bundling order information showing seven columns: **Bundling Order**, **User**, **Creation Date**, **Expiration Date**, **Media Type**, **Status**, and **Choose Bundling Order Action** (containing radio buttons for selecting an action to take and **Apply** buttons for implementing the selected actions).
- There are links on the page to **Add Bundling Order** and **Configure Defaults**.

3 Click on the **Add Bundling Order** link.

- The **Add Bundling Order** page is displayed.

4 Type the appropriate user ID in the **User ID** text entry box.

- The typed entry is displayed in the **User ID** field.

5 If the expiration date for the bundling order is different from the default (one year from the current date) type the appropriate date in the **Expiration Date** text entry box.

- The date should be entered in ***mm/dd/yyyy*** format

6 To select a **Media Type** option click on the appropriate choice (e.g., **scp**) from the option list.

7 Click on the **continue** button.

- A confirmation dialog box asks **Your present values have been entered. Continue?**

8 Click on the appropriate button from the following selections:

- **OK** - to confirm that the bundling order is to be added.
  - The GUI displays an **Add Bundling Order Detail** page that is appropriate for the selected **Media Type**.
- **Cancel**- to abort the process of adding a bundling order.

– The **Add Bundling Order** page is displayed.

– Return to Step 4.

9 Either type the appropriate data or click on the appropriate choice from the option list in the relevant fields (depending on the distribution medium selected) of the **Add Bundling Order Detail** page:

- **Email Address.**
- **User String.**
- **Distribution Priority.**
- **Shipping Information: Street1.**
- **Shipping Information: Street2.**
- **Shipping Information: Street3.**
- **Shipping Information: City.**
- **Shipping Information: State.**
- **Shipping Information: Country.**

- **Shipping Information: Zip Code.**
- **Shipping Information: Phone Number.**
- **Shipping Information: FAX Number.**
- **FTP Push Parameters: FTP Node.**
- **FTP Push Parameters: FTP Address.**
- **FTP Push Parameters: Password.**
- **FTP Push Parameters: Confirm Password.**
- **FTP Push Parameters: User String.**
- **FTP Push Parameters: Destination Directory.**
- **scp Distribution Information: User.**
- **scp Distribution Information: Host.**
- **scp Distribution Information: Password.**
- **scp Distribution Information: Retype Password.**
- **scp Distribution Information: Directory.**
- **Completion Criteria: Minimum Bundle Size.**
- **Completion Criteria: Minimum Granule Count.**
- **Completion Criteria: Maximum Bundle Age.**

NOTE: Required fields are identified by an asterisk on the **Add Bundling Order Detail** page.

10 Click on the **Add Bundling Order** button.

- A “Remember Values” Confirmation dialogue box is displayed.

11 If a **“Remember Values” Confirmation** dialogue box is displayed, click on the appropriate button from the following selections:

- **Yes**.
- **Never for this site**.
- **No**.

12 On the add bundling order confirmation page click on the appropriate button from the following selections:

- **OK -** to confirm that the bundling order is to be added.
  - A message is displayed confirming that "Bundling Order *x* was created" and there are links to **Create another Bundling Order** or **Return to Home Page**.
- **Cancel-** to abort the process of adding a bundling order.
- The **Add Bundling Order** page is displayed.
- Return to Step 4.

13 Click on the **Create another Bundling Order** link (if applicable).

- The **Add Bundling Order** page is displayed.

14 Repeat Steps 4 through 13 as necessary to add another bundling order.

A full-capability operator (only) can use the following procedure to update a bundling order:

Update a Bundling Order Using the Spatial Subscription Server GUI

---

1 Launch the **Spatial Subscription Server GUI**.

- For detailed instructions refer to the **Launch the Spatial Subscription Server GUI** procedure (previous section of this lesson).
- The **Home Page** is the default display, offering links for access to Spatial Subscription Server function pages (i.e., **List Events**, **Manage Subscriptions**, **Manage Bundling Orders**, **Monitor Queues**, **Help**, and **End Session**).

NOTE: At various points in this procedure, you may encounter a security information warning. Unless you know of a specific potential danger that you must avoid, click on the **Continue submission** button when the warning is displayed.

2 Click on the **Manage Bundling Orders** link.

- The **Manage Bundling Orders** page is displayed.

3 Observe information displayed on the **Manage Bundling Orders** page.

- The **Manage Bundling Orders** page is displayed, providing a table of bundling order information showing seven columns: **Bundling Order**, **User**, **Creation Date**, **Expiration Date**, **Media Type**, **Status**, and **Choose Bundling Order Action** (containing radio buttons for selecting an action to take and **Apply** buttons for implementing the selected actions).

4 To filter the list to display certain types of bundling orders only, click on the appropriate option button (i.e., **User**, **Media Type**, or **Status**) and then click to select the desired option from the option list.

- The selected choice is displayed in the option field.

5 Repeat Step 4 to select an additional filter, if applicable.

6 Click on the **Filter** button to implement the selected filter.

- A list is displayed showing subscriptions that meet the filter criteria only.

7 To update a particular bundling order first click on the **Update** radio button in the **Choose Bundling Order Action** column at the end of the row for the bundling order.

- The button is filled to indicate selection of the option.

8 To implement the selected action click on the **Apply** button in the **Choose Bundling Order Action** column for the bundling order.

- An **Update Bundling Order *x*** page is displayed.

9 To change the User ID (if applicable) type the User ID for the requesting user in the **User ID** text entry field.

- The typed entry is displayed in the **User Id** field.

10 To change the expiration date for the bundling order (if applicable) type the appropriate date in the **Expiration Date** text entry box.

- The date should be entered in ***mm/dd/yyyy*** format

11 To select a different **Media Type** option click on the appropriate choice (e.g., **CDROM**) from the option list.

12 Click on the **Update Bundling Order** button.

- A confirmation dialog box asks **Your present values have been entered. Continue?**

13 Click on the appropriate button from the following selections:

- **OK** - to confirm that the bundling order is to be updated.
  - The GUI displays an **Update Bundling Order *x* Detail** page that is appropriate for the selected **Media Type**.
- **Cancel**- to abort the process of updating the bundling order.

– The **Update Bundling Order *x*** page is displayed.

– Return to Step 9.

14 To change, add or delete bundling order data either type the appropriate data or click on the appropriate choice from the option list in the relevant fields (depending on the distribution medium selected) of the **Update Bundling Order *x* Detail** page:

- **Email Address.**
- **User String.**
- **Distribution Priority.**
- **Shipping Information: Street1.**
- **Shipping Information: Street2.**
- **Shipping Information: Street3.**
- **Shipping Information: City.**
- **Shipping Information: State.**
- **Shipping Information: Country.**
- **Shipping Information: Zip Code.**
- **Shipping Information: Phone Number.**
- **Shipping Information: FAX Number.**
- **FTP Push Parameters: FTP Node.**
- **FTP Push Parameters: FTP Address.**
- **FTP Push Parameters: Password.**
- **FTP Push Parameters: Confirm Password.**
- **FTP Push Parameters: User String.**
- **FTP Push Parameters: Destination Directory.**
- **scp Distribution Information: User.**
- **scp Distribution Information: Host.**
- **scp Distribution Information: Password.**
- **scp Distribution Information: Retype Password.**
- **scp Distribution Information: Directory.**
- **Completion Criteria: Minimum Bundle Size.**
- **Completion Criteria: Minimum Granule Count.**
- **Completion Criteria: Maximum Bundle Age.**

NOTE: Required fields are identified by an asterisk on the **Update Bundling Order *x* Detail** page.

15 Click on the **Update Bundling Order** button.

- A confirmation page is displayed.

16 On the update bundling order confirmation page click on the appropriate button from the following selections:

- **OK** - to confirm that the bundling order is to be updated.
  - A message is displayed confirming that "Bundling Order *x* was updated" and there are links to **Update another Bundling Order** or **Return to Home Page**.
- **Cancel**- to abort the process of adding a bundling order.
- The **Manage Bundling Orders** page is displayed.
- Return to Step 3.

17 Click on the **Update another Bundling Order** link (if applicable).

- The **Manage Bundling Orders** page is displayed.

18 Repeat Steps 3 through 17 as necessary to update another bundling order.

The **Configure Completion Criteria Default Values** page on the **Spatial Subscription Server GUI** allows a full-capability operator to set or change values assigned to bundling order completion criteria. Limited-capability operators have read-only access to the page.

The following parameters are examples of the types of bundling order completion criteria parameters that the full-capability operator can modify:

- **Minimum Granule Count**.
- **Maximum Bundle Age (days)**.
- **Bundle Expiration Period (days)**.
- **Minimum Bundle Size (GB)** [for each type of distribution medium].

Configure Bundling Order Completion Criteria Default Values

---

1 Launch the **Spatial Subscription Server GUI**.

- For detailed instructions refer to the **Launch the Spatial Subscription Server GUI** procedure (previous section of this lesson).

- The **Home Page** is the default display, offering links for access to Spatial Subscription Server function pages (i.e., **List Events**, **Manage Subscriptions**, **Manage Bundling Orders**, **Monitor Queues**, **Help**, and **End Session**).

NOTE: At various points in this procedure, you may encounter a security information warning. Unless you know of a specific potential danger that you must avoid, click on the **Continue submission** button when the warning is displayed.

2 Click on the **Manage Bundling Orders** link.

- The **Manage Bundling Orders** page is displayed, providing a table of bundling order information showing seven columns: **Bundling Order**, **User**, **Creation Date**, **Expiration Date**, **Media Type**, **Status**, and **Choose Bundling Order Action** (containing radio buttons for selecting an action to take and **Apply** buttons for implementing the selected actions).
- There are links on the page to **Add Bundling Order** and **Configure Defaults**.

3 Click on the **Configure Defaults** link.

- The **Configure Completion Criteria Default Values** page is displayed.

4 Observe information displayed on the **Configure Completion Criteria Default Values** page.

- The **Configure Completion Criteria Default Values** page has columns containing the following types of information:
  - **Parameter.**
  - **Current Value.**
  - **Change to…** (containing text entry boxes for entering new values).
- The rows on the page indicate the current values of the following types of parameters:
  - **Minimum Granule Count.**
  - **Maximum Bundle Age (days).**
  - **Bundle Expiration Period (days).**
  - **8MM Minimum Bundle Size (GB).**
  - **CDROM Minimum Bundle Size (GB).**
  - **DLT Minimum Bundle Size (GB).**
  - **DVD Minimum Bundle Size (GB).**
  - **FtpPull Minimum Bundle Size (GB).**
  - **FtpPush Minimum Bundle Size (GB).**
  - **scp Minimum Bundle Size (GB).**

5 To change the value assigned to a parameter first type the new value in the **Change to…** text entry box at the end of the line for the appropriate parameter.

- The typed entry is displayed in the field.

6 Repeat Step 5 as necessary to change any additional parameters.

7 Click on the **Change Bundling Criteria** button.

- A confirmation page is displayed with the message "Please Confirm The Following Bundling Criteria Change" and the parameter(s) for which changes were entered with the entered value(s).

8 On the change bundling criteria confirmation page click on the appropriate button from the following selections:

- **OK -** to confirm that the bundling criteria are to be changed.
  - The confirmation page is closed and the **Current Value** column on the **Configure Completion Criteria Default Values** page reflects the change(s).
- **Cancel-** to abort the process of changing bundling criteria.
  - The confirmation page is closed and the entries in the **Current Value** column on the **Configure Completion Criteria Default Values** page are restored to their original values.

Monitoring the Spatial Subscription Server Using the Spatial Subscription Server GUI

The **Spatial Subscription Server GUI** provides pages that can be used to keep track of current actions as the NBSRV processes Acquires and Notifications, and to track NBSRV performance. These capabilities are accessible through links on the **Spatial Subscription Server GUI Home Page**.

The **List Action Queue** link is used to view Acquire and Notification actions being processed by the NBSRV. The procedure that follows is applicable to both full-capability and limited-capability operators.

View the Acquire and Notification Actions Being Processed

---

1 Launch the **Spatial Subscription Server GUI**.

- For detailed instructions refer to the **Launch the Spatial Subscription Server GUI** procedure (previous section of this lesson).

- The **Home Page** is the default display, offering links for access to Spatial Subscription Server function pages (i.e., **List Events**, **Manage Subscriptions**, **Manage Bundling Orders**, **Monitor Queues**, **Help**, and **End Session**).

2 Click on the **Monitor Queues** link.

3 Click on the **List Action Queue** link.

- The **List Action Queue** page is displayed with a table listing acquire and notification actions that are being processed.

4 Observe information displayed on the **List Action Queue** page.

- The table on the **List Action Queue** page has columns containing the following types of information:
  - **Action Type.**
  - **Subscription Id.**
  - **User.**
  - **Collection.**
  - **Version.**
  - **Enqueue Time.**
  - **Dequeue Time.**
  - **Status.**
- The **Action Type** and **Subscription Id** column headers are links for sorting the list.
- There are option lists for filtering the table data by **Action Type**, **Subscription**, and/or **Status**.

5 To filter the list to display certain types of actions only, click on the appropriate option button and then click to select the desired option from the option list.

- The selected choice is displayed in the option field.

6 Repeat Step 5 to select an additional filter, if applicable.

7 Click on the **Filter** button to implement the selected filter.

- A list is displayed showing items that meet the filter criteria only.

8 Return to Step 4.

The **List Failed Action** link is used to view and remove failed actions processed by the NBSRV. The full-capability operator can both view and remove failed actions. The limited-capability operator has read-only access to the **List Failed Action** page.

View and Remove Failed Actions Using the Spatial Subscription Server GUI

1 Launch the **Spatial Subscription Server GUI**.

- For detailed instructions refer to the **Launch the Spatial Subscription Server GUI** procedure (previous section of this lesson).
- The **Home Page** is the default display, offering links for access to Spatial Subscription Server function pages (i.e., **List Events**, **Manage Subscriptions**, **Manage Bundling Orders**, **Monitor Queues**, **Help**, and **End Session**).

2 Click on the **Monitor Queues** link.

3 Click on the **List Failed Action** link.

- The **List Failed Action** page is displayed with a table listing acquire and notification actions that are being processed.

4 Observe information displayed on the **List Failed Action** page.

- The table on the **List Failed Action** page has columns containing the following types of information:
  - **Priority.**
  - **ActionId.**
  - **ActionQueueId.**
  - **SubscriptionId.**
  - **granUR.**
  - **Remove Action.**
- The **Remove Action** column contains **Remove Action** buttons for removing failed actions.

5 To remove a failed action click on the **Remove Action** button at the end of the row for the action.

6 Return to Step 4.

The **List Statistics** link is used to review summary information concerning the processing of events and subscriptions. The procedure that follows is applicable to both full-capability and limited-capability operators.

View Statistics on Processing of Events and Actions by the NBSRV

1 Launch the **Spatial Subscription Server GUI**.

- For detailed instructions refer to the **Launch the Spatial Subscription Server GUI** procedure (previous section of this lesson).
- The **Home Page** is the default display, offering links for access to Spatial Subscription Server function pages (i.e., **List Events**, **Manage Subscriptions**, **Manage Bundling Orders**, **Monitor Queues**, **Help**, and **End Session**).

2 Click on the **Monitor Queues** link.

3 Click on the **List Statistics** link.

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

4 Observe information displayed on the **List Statistics** page.

- The **List Statistics** page has a summary showing the number of **Subscribed Events Left to Dequeue**.
- The **List Statistics** page has a summary showing the number of **Matched Subscriptions Left to Dequeue**.
- The **List Statistics** page has a **Summary of the Processing Time for Subscriptions Evaluated against Events in the Event Queue** table that has columns containing the following types of information:
  - **ESDT Count.**
  - **ESDT Short Name.**
  - **ESDT VersionID.**
  - **ESDT Event Type.**
  - **MaxMetadataTime (ms).**
  - **AverageMetadataTime (ms).**
  - **MaxSubEvalTime (ms).**
  - **AverageSubEvalTime (ms).**
- The **List Statistics** page has a **Summary of the Processing Time for E-Mail Notification Actions in Action Queue** table that has columns containing the following types of information:
  - **Notification Count.**
  - **MaxNotifyTime (ms).**

  - **AverageNotifyTime (ms).**

- The **List Statistics** page has a **Summary of the Processing Time for Distribution Actions in Action Queue** table that has columns containing the following types of information:
  - **Acquire Count.**
  - **MaxAcquireTime (ms).**
  - **AverageAcquireTime (ms).**

This page intentionally left blank.

Practical Exercise

Introduction

This exercise is designed to practice key elements of the Archive procedures. Perform the tasks identified in the exercise.

Equipment and Materials

One workstation per student.

STK Powderhorn storage facility.

Statement of the requirements for the exercise.

Release 7.11 Operations Tools Manual for the EMD Project, 609-EMD-001, one copy per student.

Release 7.11 Mission Operation Procedures for the EMD Project, 611-EMD-001, one copy per student.

Perform Activities Related to Archive Processing

1. Locate the STK Powderhorn storage facility and the AMASS host. Point out the elements and sequence involved in starting AMASS.
2. At the STK Powderhorn, locate the control panels necessary for power up (or down) and identify all Power Switches.
3. Following all safety precautions, vary the STK Powderhorn offline and enter the unit; leave the unit and restore it to online status.
4. Launch the Data Distribution Graphical User Interface (GUI); examine the list of distribution requests. Then filter the list to examine only those requests that are staging.
5. Insert several granules that can be used to exercise the granule deletion capability, and then use that capability to delete the granules from the inventory and archive, specifying a lag time other than 0 in the Deletion Cleanup Utility script.
6. Use automatic loading procedures to load a 9940 tape into the STK Powderhorn. Then use automatic unloading procedures to remove the media you just loaded.
7. Experiment with the **vollist**, **dirfilelist**, and **volfilelist** commands for AMASS. Describe for yourself how the commands can be used to help you manage the archive.

8. Use the **vgexport -q** command to create a backup for the AMASS database.
9. Check for the existence of required running AMASS daemons; run **healthcheck** to determine the health of basis AMASS functions.
10. Look at the AMASS queue by using the **quedisplay** command.
11. Run the **amass_log** script to display AMASS messages in **/var/adm/SYSLOG** system log file (on SGI machines) or in **/var/adm/messages** (on Sun machines).
12. Launch the Data Pool Maintenance (DPM) GUI and review the Data Pool active insert processes. Use the **Refresh** button to obtain an immediate screen refresh. Change the automatic screen refresh rate and observe that the new rate takes effect.
13. Use the DPM GUI to suspend Data Pool Insert actions. Then use the GUI to resume Data Pool Insert actions.
14. Use the DPM GUI to check the Data Pool Insert Queue.
15. Use the DPM GUI **Manage Configuration Parameters** tab and increase the number of allowed active insert processes by 5. Then decrease the number by 5. Change the Default Retention Priority to a number of your choosing.
16. Use the DPM GUI to display a list of collection groups for data in the Data Pool. Select one of the groups and display the list of its collections. Then pick one of the collections in the group and display a description for it.
17. Use the DPM GUI to obtain a list and descriptions of themes in the Data Pool database. Then add a theme that is insert enabled but not web visible. Delete the theme you created.
18. Launch the Spatial Subscription Server (NBSRV) GUI and display a list of events for which a subscription can be created.
19. Using the Spatial Subscription Server GUI, display subscriptions in the NBSRV database and filter the list to show only subscriptions for data with a selected short name. Filter the list again to display only subscriptions for data with a different short name.
20. Use the Spatial Subscription Server GUI to obtain a list of bundling orders and view the details of one of them. Then add a bundling order, but do not associate it with a subscription. Delete the bundling order you created.
21. Use the Spatial Subscription Server GUI to view any acquire and notification actions in the Action Queue.
22. Use the Spatial Subscription Server GUI to display statistics on NBSRV processing of events and actions.
23. Run the **Update Granule Utility** to extend the period of retention for a single granule in the Data Pool. Examine the log file for the utility and review the entries for the update.
24. Invoke the **Data Pool Cleanup Utility** from the command line.
25. Execute the **Data Pool Access Statistics Utility** from the command line.

26. Use the **Data Pool Archive Access Statistics Data Utility** to archive statistics on Data Pool access for the last seven days.
27. Use the **Batch Insert Utility** to insert several granules from the archive into the Data Pool.

This page intentionally left blank.

Slide Presentation

Slide Presentation Description

The following slide presentation represents the slides used by the instructor during the conduct of this lesson.

This page intentionally left blank.