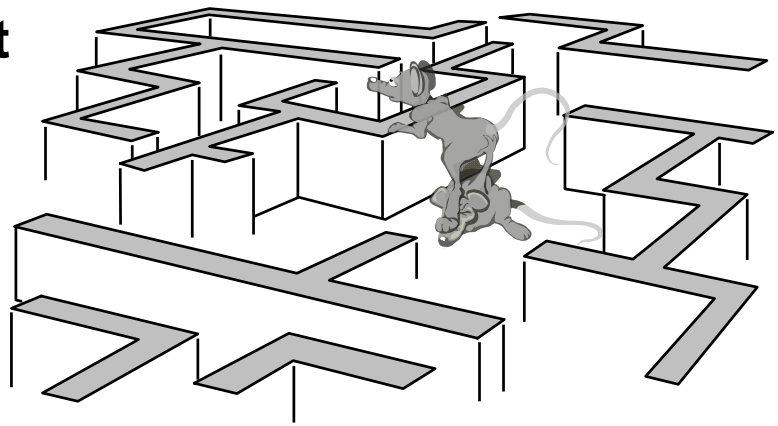
INGEST

EMD Training

625-EMD-008

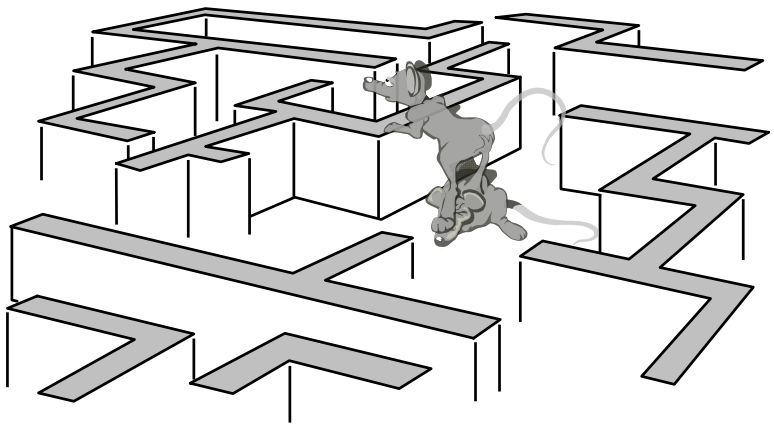
Overview of Lesson

- Introduction
- Ingest Topics
  - Ingest Concepts
  - Logging in to ECS Hosts
  - Launching the ECS Ingest and Storage Management Control GUIs
  - Handling Cross-DAAC or Cross-Mode Ingest
  - Monitoring Ingest Status



Overview of Lesson (Cont.)

- **Ingest Topics (Cont.)**
  - **Cleaning Directories**
  - **Performing Hard Media Ingest**
  - **Modifying Ingest Tunable Parameters and Performing File Transfers**
  - **Troubleshooting Ingest Problems**
- **Practical Exercise**



Objectives

- **OVERALL:**
  - **Develop proficiency in the procedures that apply to ingest operations**
- **SPECIFIC:**
  - **Describe the ingest function, including a general statement of the ingest responsibility in ECS and an overview of the ingest process**
  - **Perform the steps involved in...**
    - » **logging in to ECS hosts**
    - » **launching the ECS Ingest GUI**
    - » **launching the Storage Management Control GUI**
    - » **handling cross-DAAC or cross-mode ingest**
    - » **monitoring and controlling ingest requests**
    - » **viewing the Ingest History Log**
    - » **verifying the archiving of ingested data**

Objectives (Cont.)

- **SPECIFIC (Cont.):**
  - **Perform the steps involved in...**
    - » **cleaning the polling directories**
    - » **performing hard media ingest**
    - » **modifying external data provider information**
    - » **modifying Ingest Subsystem parameters using the Ingest GUI Operator Tools: Modify System Parameters screen**
    - » **transferring files using the Ingest GUI File Transfer screen**
    - » **modifying Ingest Subsystem parameters using interactive structured query language (isql)**
    - » **troubleshooting and recovering from ingest problems**

Objectives (Cont.)

- **STANDARDS:**
  - **Lesson content (procedures in the lesson)**
  - **Mission Operation Procedures for the EMD Project (611-EMD-001)**

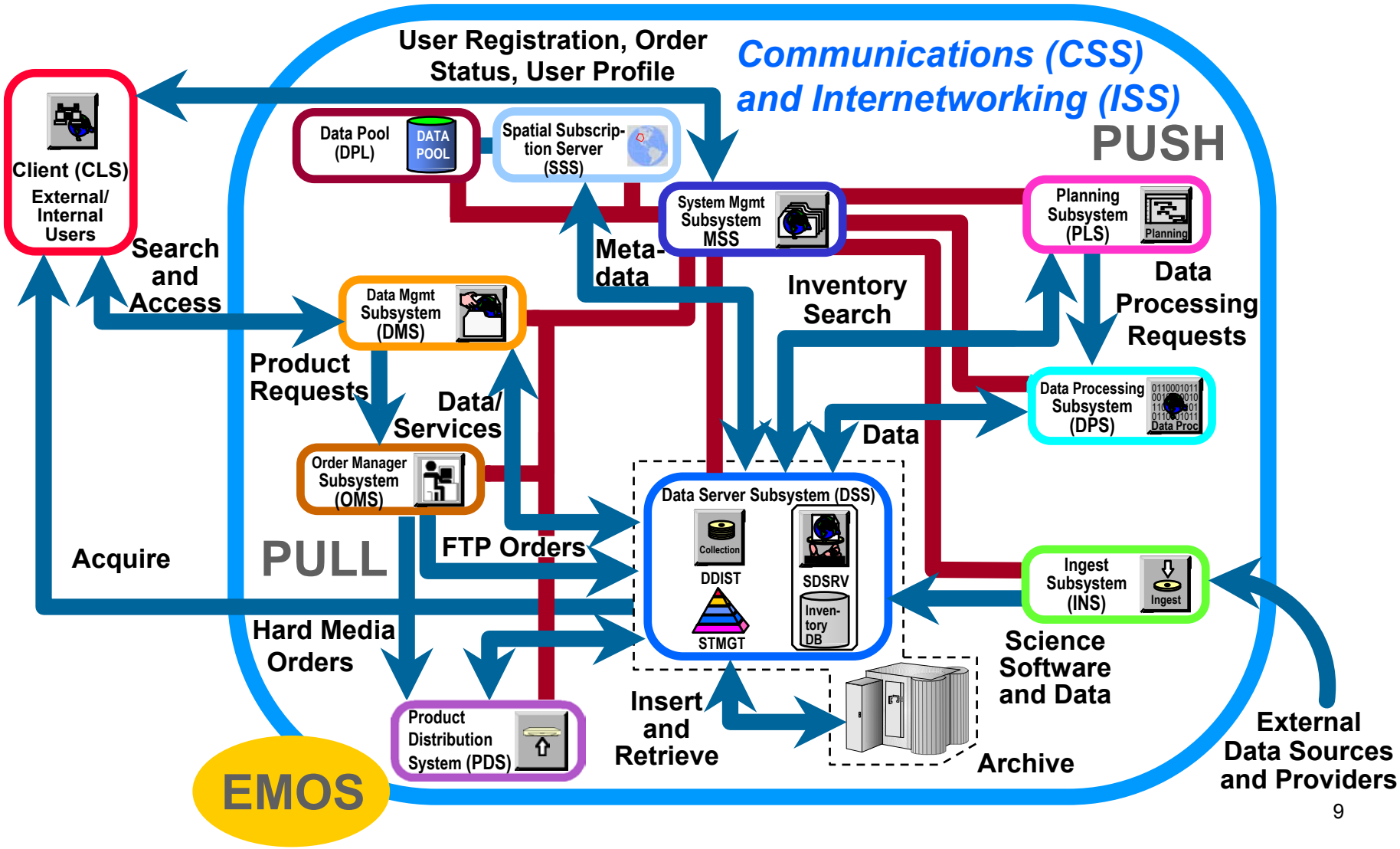
Ingest Concepts

- **ECS Context**
  - Ingest for ECS is accomplished at the Distributed Active Archive Centers (DAACs)
  - People involved in Ingest activities are Ingest Technicians
  - Ingest Subsystem (INS) is point of entry to ECS for data from external data providers
  - Data Server Subsystem (DSS) manages access to the data repositories, where ingested data are stored

Ingest Concepts (Cont.)

- ECS Context (Cont.)
  - Ingest transfers data into ECS, performs preprocessing, and forwards the data to DSS for archiving
  - STMGT CSCI in DSS stores, manages, and retrieves data files
    - » Provides interfaces and peripheral devices (e.g., tape drives)
    - » Provides for the copying of files into the archive for permanent storage
  - SDSRV CSCI in DSS manages and provides user access to collections of Earth Science data
    - Checks/verifies metadata
    - Issues requests to STMGT to perform storage services, such as insertion of data into the archive

ECS Context Diagram



Ingest Concepts (Cont.)

- **Ingest Subsystem: INGST CSCI**
  - Automated Polling Ingest Client Interface (EcInPolling)
  - Ingest Request Manager (EcInReqMgr)
  - Ingest Granule Server (EcInGran)
  - Ingest E-Mail Parser (EcInEmailGWServer)
  - ECS Ingest GUI (EcInGUI)
  - INGEST Media Tape Reader GUI (EcInTapeReaderGUI)
  - Sybase Adaptive Server Enterprise (ASE) Server

Ingest Concepts (Cont.)

- **INGST (Cont.)**
  - **Start-up script used by Ingest personnel (/usr/ecs/*MODE*/CUSTOM/utilities directory on the Operations Workstation)**
    - » **EcInGUIStart**
  - **Start-up script used by Ingest personnel (/usr/ecs/*MODE*/CUSTOM/utilities directory on the Sun Consolidation Internal Server host)**
    - » **EcInTapeReaderGUI**

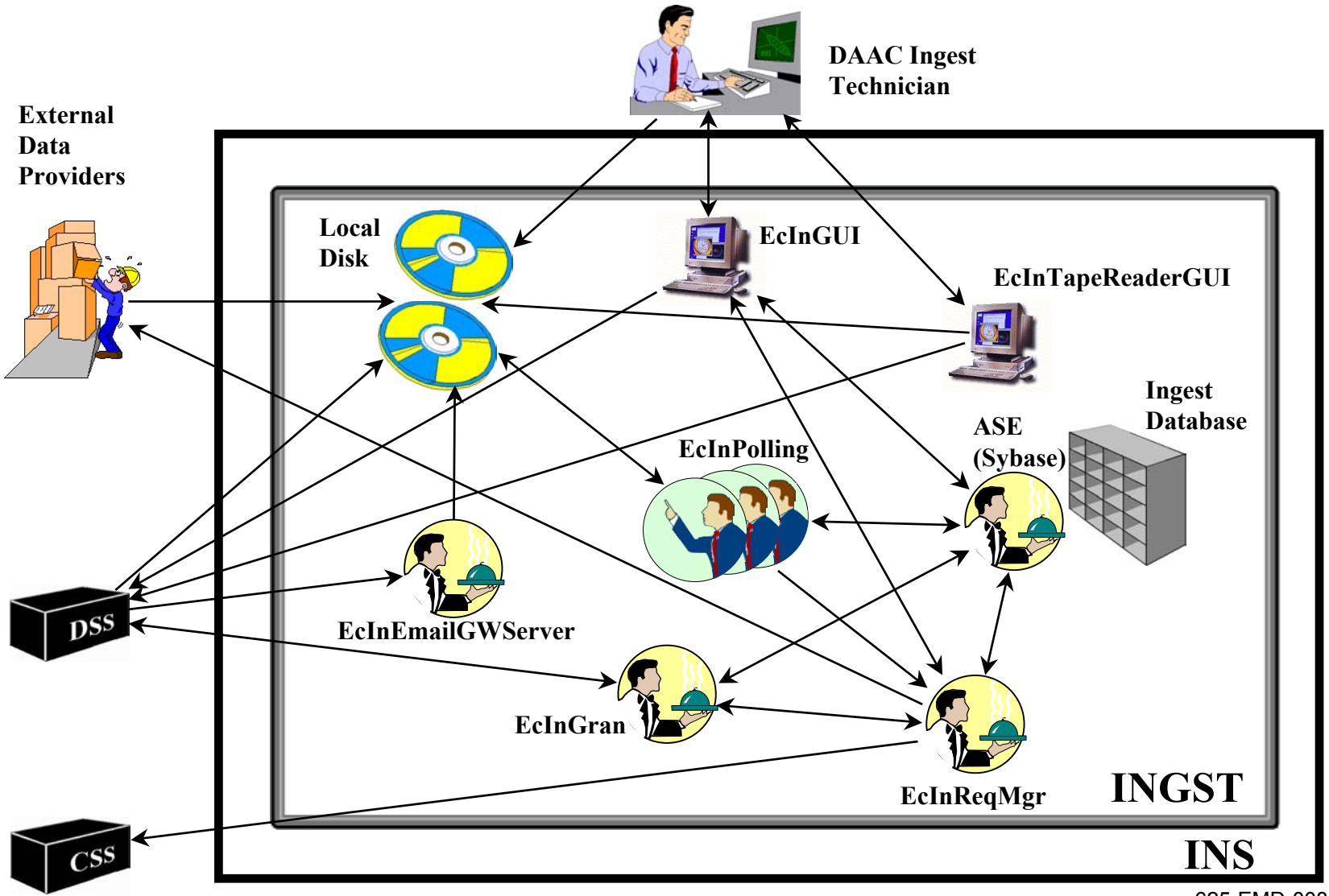
Ingest Concepts (Cont.)

- **INGST (Cont.)**
  - **Start-up scripts called by other applications (not normally invoked directly by Ingest personnel)**
    - » **EcInGranStart**
    - » **EcInIngestAppStart**
    - » **EcInInitPasswdStart**
    - » **EcInPollingStart**
    - » **EcInReqMgrStart**
    - » **EcInStart**
    - » **EcInEmailGWServerStart**

Ingest Concepts (Cont.)

- INGST (Cont.)
  - Other scripts
    - » EcInEDOSCleanupMain
    - » EcInPollClean
    - » EcInCopyAM1ANC
    - » EcInDbBuild
    - » EcInDbDrop
    - » EcInDbDump
    - » EcInDbDumpTrans
    - » EcInDbLoad
    - » EcInDbLoadTrans
    - » EcInDbPatch

Ingest Subsystem Architecture and Interfaces



Ingest Concepts (Cont.)

- **Data Server Subsystem: Storage Management (STMGT) CSCI**
  - **Archive Server (EcDsStArchiveServer)**
  - **Staging Servers**
    - » **Cache Manager Server (EcDsStCacheManagerServer)**
    - » **Pull Monitor (EcDsStPullMonitorServer) [symbolic link to the Cache Manager Server]**
    - » **Staging Disk Server (EcDsStStagingDiskServer)**
  - **Resource Managers**
    - » **8mm Server (EcDsSt8MMServer)**
    - » **DTF-2 Server (EcDsStDTFServer)**
    - » **FTP Server (EcDsStFtpServer)**
    - » **Copy Server (EcDsStCopyServer)**

Ingest Concepts (Cont.)

- **STMGT (Cont.)**
  - **Storage Management Request Manager (EcDsStRequestManagerServer)**
  - **Storage Management Control GUI (EcDsStmgtGui)**
  - **Sybase ASE Server**
  - **Archival Management and Storage System (AMASS)**

Ingest Concepts (Cont.)

- **STMGT (Cont.)**
  - **Start-up script used by Ingest personnel (/usr/ecs/*MODE*/CUSTOM/utilities directory on the Operations Workstation)**
    - » **EcDsStmgtGuiStart**

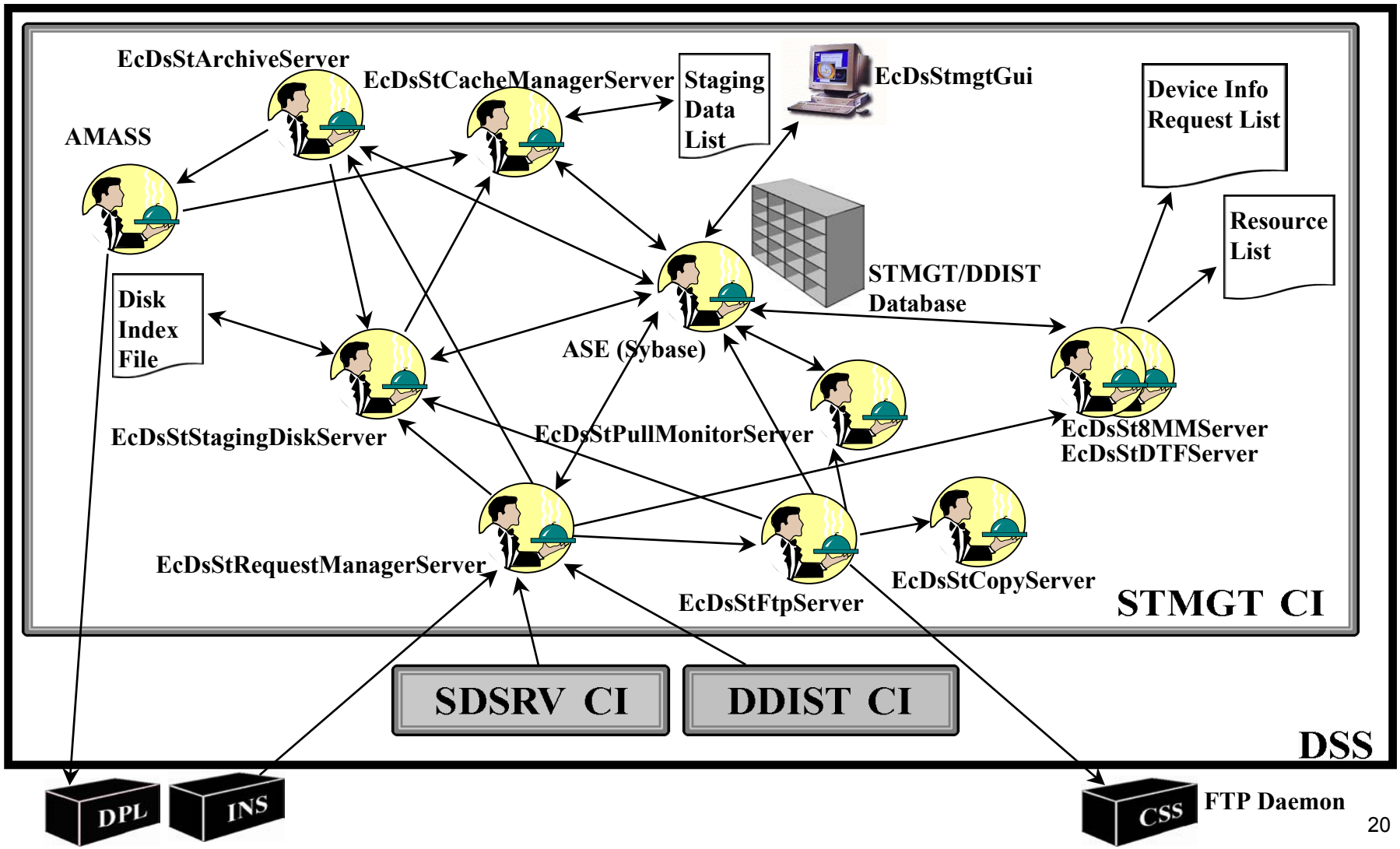
Ingest Concepts (Cont.)

- **STMGT (Cont.)**
  - **Start-up scripts called by other applications (not normally invoked directly by Ingest personnel)**
    - » **EcDsStFtpServerStart**
    - » **EcDsStStagingDiskServerStart**
    - » **EcDsStStart**
    - » **EcDsStStorageMgmtAppStart**
    - » **EcEcsAppStart**
    - » **EcDsStArchiveServerStart**
    - » **EcDsStCacheManagerServerStart**
    - » **EcDsStRequestManagerServerStart**
    - » **EcDsSt8MMServerStart**
    - » **EcDsStDTFServerStart**

Ingest Concepts (Cont.)

- STMGT (Cont.)
  - Other scripts
    - » EcDsCheckArchive
    - » EcDsStConfigVolGrps
    - » EcDsStDbBuild
    - » EcDsStDbDrop
    - » EcDsStDbDump
    - » EcDsStDbDumpTrans
    - » EcDsStDbLoad
    - » EcDsStDbLoadTrans
    - » EcDsStDbPatch
    - » EcDsStFilesPerTapeUtility
    - » EcDsStVolGrpCreateMain.pl

Data Server Subsystem: STMGT Architecture and Interfaces



Ingest Concepts (Cont.)

- **Data Server Subsystem: Science Data Server (SDSRV) CSCI**
  - Science Data Server (EcDsScienceDataServer)
  - Hierarchical Data Format (HDF) EOS Server (EcDsHdfEosServer)
  - Granule Deletion Administration Tool (EcDsGranuleDelete)
  - Science Data Server GUI (EcDsSdSrvGui)
  - Science Data Server (SDSRV) Command Line Interface (SCLI) (EcDsSCLI)
  - Autometric Spatial Query Server (SQS)
  - Sybase ASE Server

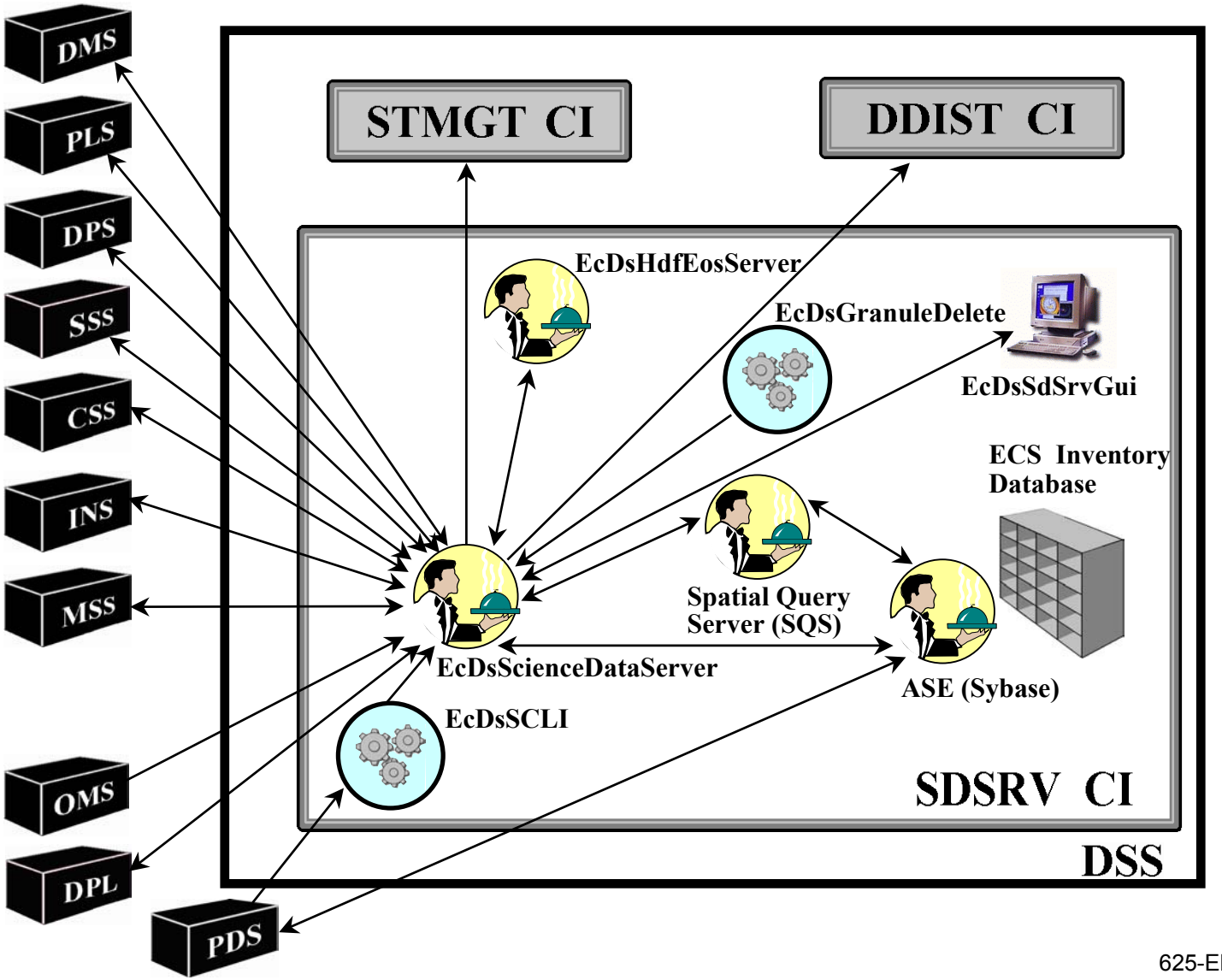
Ingest Concepts (Cont.)

- **SDSRV (Cont.)**
  - **Start-up script (/usr/ecs/*MODE*/CUSTOM/utilities directory on the Operations Workstation)**
    - » **EcDsSdSrvGuiStart**

Ingest Concepts (Cont.)

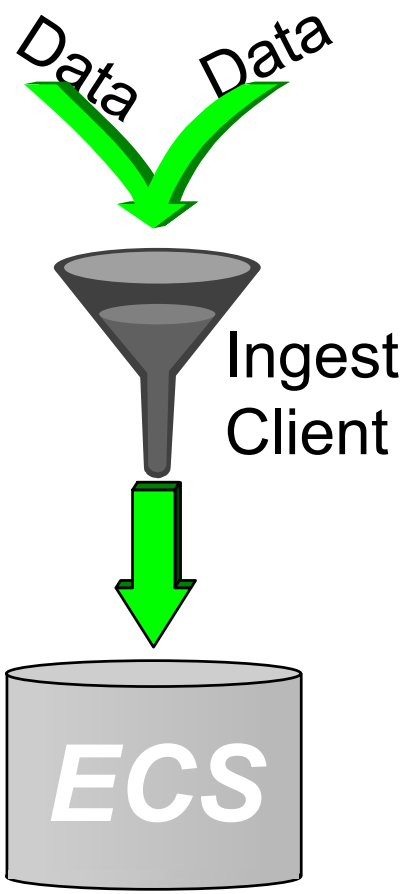
- **SDSRV (Cont.)**
  - **Other scripts**
    - » **EcTsDsClientDriverStart**
    - » **EcDsSrConvertEvt**
    - » **EcDsSrDbBuild**
    - » **EcDsSrDbDrop**
    - » **EcDsSrDbDump**
    - » **EcDsSrDbLoad**
    - » **EcDsSrDbMigrate**
    - » **EcDsSrDbPatch**
    - » **EcDsSrDbValids**

Data Server Subsystem: SDSRV Architecture and Interfaces



Ingest Concepts (Cont.)

- **Ingest Process**
  - **Hardware and software for Ingest**
    - » **Receipt and storage of data from multiple sources into ECS**
    - » **Sets stage for archiving and/or processing of the data**
  - **Provides tools**
    - » **Selected configuration: *Ingest client***
    - » **Single virtual interface point for receipt of all external data to be archived**
    - » **Performs ingest data preprocessing, metadata extraction, and metadata validation**



Ingest Concepts (Cont.)

- **Ingest Activities**
  - Ingest function brings data into ECS from external data providers
  - Representative data providers
    - » Landsat 7 Archive Management System (LAMS)
    - » Landsat 7 Image Assessment System (IAS)
    - » Landsat 7 International Ground Stations (IGSs)
    - » Ground Data System (GDS) for the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) instrument on the Terra (AM-1) satellite
    - » EOS Data and Operations System (EDOS)
    - » Science Computing Facilities (SCFs)

Ingest Concepts (Cont.)

- **Ingest Activities (Cont.):**
  - **Representative data providers (Cont.):**
    - » **Science Investigator-Led Processing Systems (SIPS); for example:**

      Advanced Cavity Radiometer Irradiance Monitor (ACRIM) III

      Advanced Microwave Scanning Radiometer (AMSR-E)

      ASTER Observation Schedule File (OSF)

      ASTER Digital Elevation Model (DEM)

      Measurements of Pollution in the Troposphere (MOPITT)

      MODIS Adaptive Processing System (MODAPS)

      Stratospheric Aerosol and Gas Experiment (SAGE) III

Ingest Concepts (Cont.)

- **Ingest Activities (Cont.)**
  - **Representative data providers (Cont.)**
    - » **National Oceanic and Atmospheric Administration (NOAA) National Environmental Satellite, Data, and Information Service (NESDIS)**
    - » **NOAA National Weather Service (NWS) National Centers for Environmental Prediction (NCEP)**
    - » **Flight Dynamics System (FDS)**

Ingest Concepts (Cont.)

- **Ingest Activities (Cont.)**
  - **Ingest activities include…**
    - » **Data transfer and transmission checking**
    - » **Data preprocessing [including data conversions (if required), metadata extraction (as required), metadata validation (as required)]**
    - » **Transferring ingested data to the Data Server Subsystem for long-term storage in the archive**

Ingest Concepts (Cont.)

- **Ingest Activities (Cont.)**
  - **Ingest provides a single point for monitoring and control of data ingested from external data providers**
  - **Nominal ingest process is fully automated with minimal operator intervention**

Ingest Concepts (Cont.)

- **Ingest Categories**
  - **Automated polling ingest with delivery record**
    - » **ECS periodically checks a network location for a delivery record file, which indicates the availability of data for ingest**
    - » **ECS "gets" data from the applicable directory on an ECS staging server, where the data provider will have put the data**
    - » **Data providers include EDOS, LAMS, IAS, SCFs, SIPS, and NOAA NCEP**

Ingest Concepts (Cont.)

- **Ingest Categories (Cont.)**
  - **Automated polling ingest without delivery record**
    - » **ECS periodically checks a network location for available data**
    - » **All data at the location are treated as one specific data type, one file per granule**
    - » **If the particular polling client is configured to perform file comparisons, each file is compared with the last version that was ingested**
    - » **If the new file is different from the previous one, it is ingested as a new file**
    - » **If it is identical to the previous one, it is not ingested**
    - » **ECS "gets" data from the network location**
    - » **Data providers include FDS and NOAA NESDIS CEMSCS**

Ingest Concepts (Cont.)

- **Ingest Categories (Cont.)**
  - **Hard media ingest by the Ingest Technician**
    - » **Ingest from hard media (e.g., tape cartridges); from authorized institutions or other providers, or as backup**
    - » **Requires file/record information equivalent to a PDR**
    - » **Data providers include ASTER GDS, Landsat-7 (L7) International Ground Stations (IGSs), EDOS, NOAA NCEP, and NOAA NESDIS NCDC**

Ingest Concepts (Cont.)

- **Ingest Categories (Cont.)**
  - **Cross-Mode Ingest Interface**
    - » **Ingest from other DAACs or other modes at the same DAAC**
    - » **Ingest receives a distribution notice (via e-mail) of data files transferred via the FTP service**
    - » **Distribution notification is used to create a Delivery Record File that is put in an agreed-upon network location**
    - » **Polling-with-delivery-record process checks the location for the delivery record files**

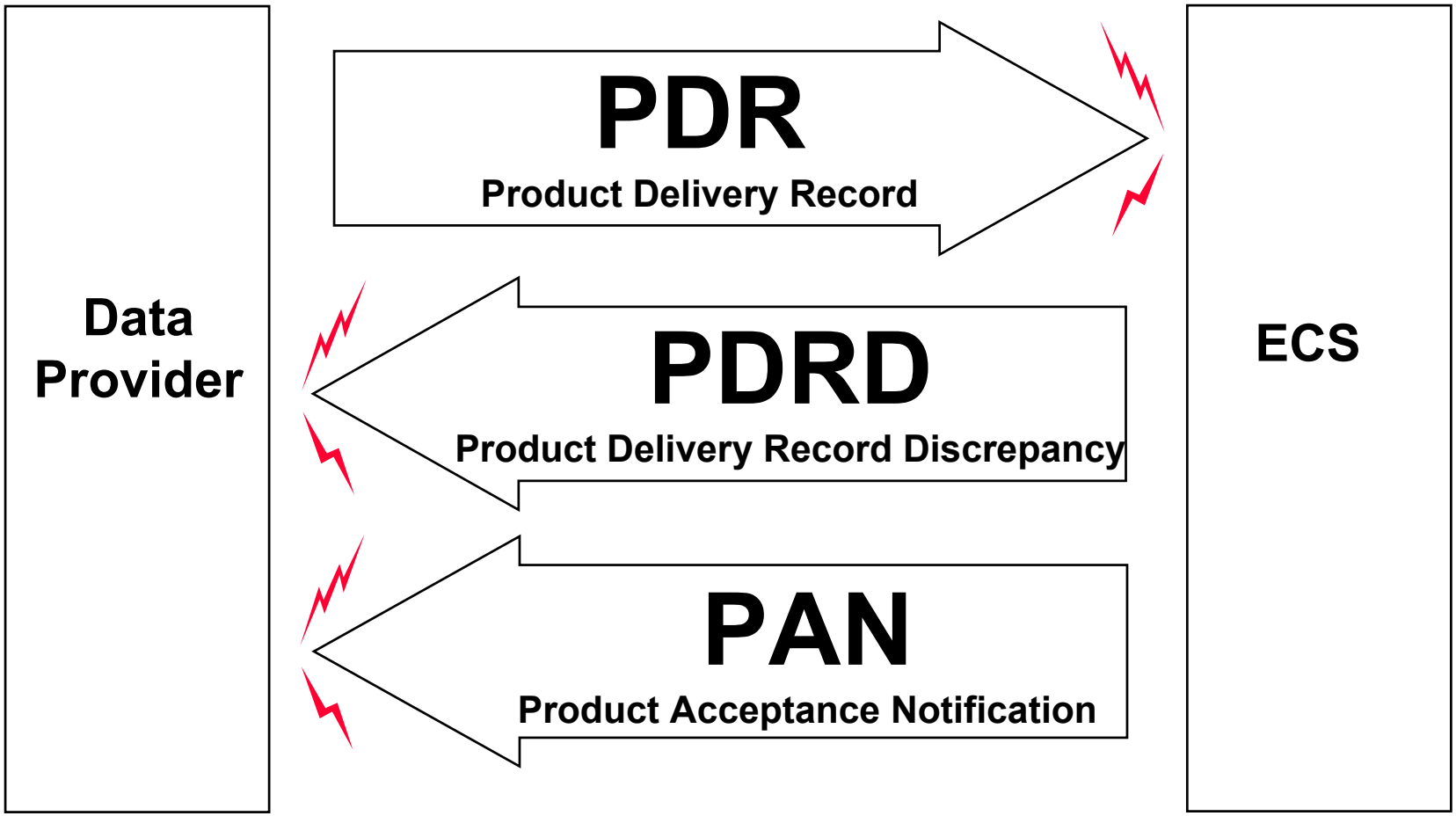
Ingest Concepts (Cont.)

- **Ingest Polling Messages (polling ingest with delivery record only)**
  - **Product Delivery Record (PDR)**
    - » **Notice that is sent to Ingest by an external data provider specifying data that are available for ingest**
    - » **Alternatives include the EDOS Production Data Set (PDS) Delivery Record (PDR) and Expedited Data Set (EDS) Delivery Record (EDR)**
    - » **EDOS provides a signal file to indicate that EDOS has completed transfer of all data files and the accompanying delivery record so the data can be ingested; the signal file is identified by an ".XFR" extension to the delivery record file name**

Ingest Concepts (Cont.)

- **Ingest Polling Messages (Cont.)**
  - **Product Delivery Record Discrepancy (PDRD)**
    - » **Notice sent from Ingest to the data provider via ftp or e-mail indicating that the PDR cannot be successfully validated**
    - » **There is no PDRD in EDOS polling**
  - **Product Acceptance Notification (PAN)**
    - » **Message sent via ftp or e-mail to the data provider from Ingest announcing the completion of data transfer and archiving and identifying any problems with any of the files specified in the PDR**
    - » **Alternatives include the EDOS PDS Acceptance Notification (PAN) and EDS Acceptance Notification (EAN)**

Ingest Polling Messages



Ingest Concepts (Cont.)

- **Ingest Polling Messages (Cont.)**
  - **No messages are passed in polling ingest without delivery record**

Ingest Concepts (Cont.)

- **Physical (Hard) Media Messages (exact nomenclature of messages depends on the particular data provider's agreement with ECS)**
  - **Physical Media Product Delivery Record (PMPDR)**
    - » **Notice provided on the physical medium (i.e., 8mm tape) and hard copy (if needed) to identify the data on the tape that are available for ingest**
  - **Physical Media PDR Discrepancy (PMPDRD)**
    - **Notice sent from Ingest to the data provider via e-mail only in the event of an error in the PMPDR**
  - **Physical Media Production Acceptance Notification (PMPAN)**
    - **Message sent via e-mail to the data provider from Ingest announcing the completion of data transfer and archiving and identifying any problems**

Ingest Concepts (Cont.)

- **Data Transfer and Staging**
  - **Data transfer from external data providers uses one of three methods:**
    - » **File transfer protocol (ftp) “get” by ECS**
    - » **Ftp “put” by external source**
    - » **Hard media transfer**

Ingest Concepts (Cont.)

- Data are staged to a working storage area
  - In general level 0 (L0) data received from external data providers and other selected data (EDOS ancillary data) is staged to the INS working storage area
    - » Metadata is extracted and the format is validated in the working storage area then the L0 data is transferred to an archive data repository in the DSS for long-term storage
  - Non-L0 data (such as non-EDOS ancillary data and L1A - L4 data from external facilities) is staged directly to the working storage area in the DSS
    - » Extraction of metadata is performed on the data by the INS software then the DSS (SDSRV CSCI) is called by the INS software residing in the INS processor hardware to perform metadata validation and the non-L0 data is transferred to a DSS archive data repository for long-term storage

Ingest Concepts (Cont.)

- Data are staged to a working storage area (Cont.)
  - Many types of ingest use “icl” (Ingest Client) staging areas
  - Hard media ingest (e.g., from 8mm tape) may involve staging in a “dip” (Distribution and Ingest Peripherals) area
  - Polling ingest for data from EDOS usually entails the use of the polling directory as the staging area
  - Some data are staged directly to working storage (“wks”) in the Data Server Subsystem

Ingest Concepts (Cont.)

- **Ingest Scenarios - Scenarios are presented for the following types of Ingest:**
  - **Polling ingest with delivery record**
  - **Polling ingest without delivery record**
  - **Hard media ingest**
  - **Cross-DAAC ingest**

ECS Ingest GUI Intro Screen

Logging in to ECS Hosts

- Logging in to ECS hosts is accomplished from a UNIX command line prompt
  - It is an initial set of steps that is performed when accomplishing many other Ingest tasks
- Procedure
  - Access the command shell
  - Set the DISPLAY environmental variable
  - Log in to the specified host using secure shell and the specified user ID

Launching the ECS Ingest and Storage Management Control GUIs

- Software applications associated with Ingest:
  - Automated Polling Ingest (EcInPolling)
  - Request Manager (EcInReqMgr)
  - Granule Server (EcInGran)
  - ECS Ingest GUI (EcInGUI)
  - Ingest E-Mail Parser (EcInEmailGWServer)
  - Sybase ASE Server
- Normally multiple instances of some Ingest servers
- Ingest depends on other servers, especially Storage Management and Science Data Server

Launching the ECS Ingest and Storage Management GUIs (Cont.)

- Use UNIX command line to gain access to graphical user interfaces (GUIs)
- Procedure (Launching the ECS Ingest GUI)
  - Access a terminal window logged in to the Operations Workstation host
  - Change directory to the directory containing the Ingest GUI startup script
  - Type command to start ECS Ingest GUI

Launching the ECS Ingest and Storage Management GUIs (Cont.)

- **Software applications associated with Storage Management as it relates to Ingest:**
  - Storage Management Control GUI (EcDsStmgtGui)
  - Archive Server (EcDsStArchiveServer)
  - Cache Manager Server (EcDsStCacheManagerServer)
  - Pull Monitor (EcDsStPullMonitorServer)
  - Staging Disk Server (EcDsStStagingDiskServer)
  - 8mm Server (EcDsSt8MMServer)
  - DTF-2 Server (EcDsStDTFServer)
  - FTP Server (EcDsStFtpServer)
  - Storage Management Request Manager (EcDsStRequestManagerServer)

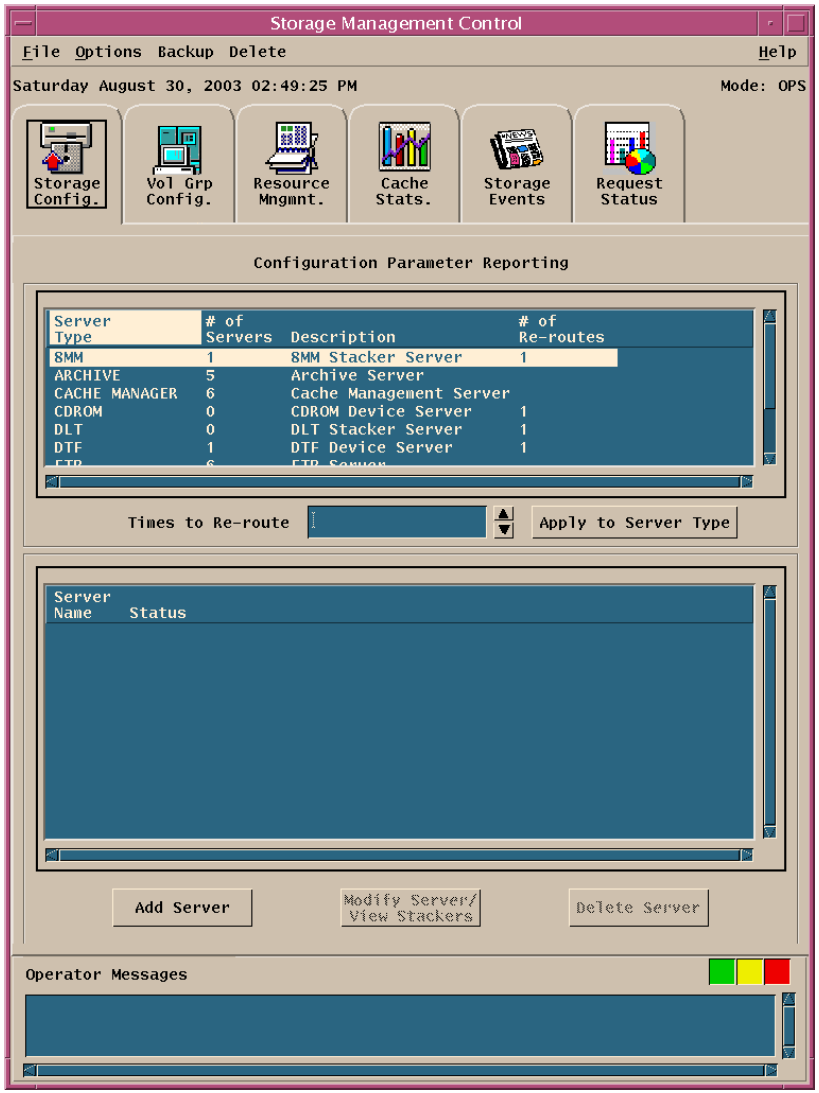
Launching the ECS Ingest and Storage Management GUIs (Cont.)

- **Software applications associated with Storage Management (Cont.):**
  - **Sybase ASE Server**
  - **Archival Management and Storage System (AMASS)**
- **Storage Management Control GUI is used in Ingest physical media operations for…**
  - **putting media IDs into the Storage Management database**
  - **marking stackers/drives/slots as being either on line or off line**

Launching the ECS Ingest and Storage Management GUIs (Cont.)

- **Procedure (Launching the Storage Management Control GUI)**
  - **Access a terminal window logged in to the Operations Workstation**
  - **Change directory to the directory containing the Storage Management Control GUI startup script**
  - **Type command to start the Storage Management Control GUI**

Storage Management Control GUI: Storage Config. Tab



Handling Cross-DAAC or Cross-Mode Ingest

- Cross-DAAC or cross-mode ingest is launched via either an order or a subscription for the desired data
  - If the data are already in the archive, an order is submitted using the EOS Data Gateway (EDG) web client
    - » An order for data can be placed via the EDG from almost anyplace by almost anyone
  - If the data are not already in the archive (if future data are to be transferred), a subscription is entered using the Subscription Service GUI
    - » A subscription is likely to be entered by User Services personnel at the Distributing DAAC
  - In either case either a subscription or order is entered in the mode (e.g., OPS) from which the data are to be transferred

Handling Cross-DAAC or Cross-Mode Ingest (Cont.)

- The subscription or order specifies (among other things)…
  - The data to be transferred
    - » Parameters of the specific granule(s) if the data are being ordered
    - » Data type and other subscription parameters (e.g., from what date/time until what date/time) if a subscription is being entered
  - Media type is FtpPush
  - Media format is typically FILEFORMAT
  - User profile ID is the user profile ID of the nominal requester (if applicable)
  - A username is specified for logging in to the ftp host at the receiving DAAC
  - A password is specified for logging in to the ftp host at the receiving DAAC

Handling Cross-DAAC or Cross-Mode Ingest (Cont.)

- The subscription or order specifies (among other things)…
  - The ftp host is the host to which the data are to be pushed at the receiving DAAC (e.g., g0acg01u.ecs.nasa.gov)
  - The ftp directory is the directory (on the ftp host) to which the data are to be pushed at the receiving DAAC/mode (e.g., /usr/ecs/OPS/CUSTOM/icl/a/data/XDAAC_Ingest/NSIDC/)
  - The e-mail address is the address for the distribution notice (DN) that is sent by Data Distribution at the transmitting DAAC/mode to the Ingest E-Mail Parser (EcInEmailGWServer) in the receiving mode at the receiving DAAC
    - » For example:

      EcInEmailGWServer_OPS@g0ins01u.ecs.nasa.gov

Handling Cross-DAAC or Cross-Mode Ingest (Cont.)

- **Assuming the processing of the subsequent acquire request(s) and processing of the distribution request(s) are successful, the following actions occur as the data are transferred:**
  - **The ftp daemon at the transmitting DAAC/mode performs the actual ftp of the files to the receiving DAAC/mode**
    - » **For example:**

      /usr/ecs/OPS/CUSTOM/icl/a/data/XDAAC_Ingest/NSIDC/ on g0acg01u.ecs.nasa.gov
  - **The Distribution Server at the transmitting DAAC/mode builds a distribution notice that the user's order has been fulfilled and sends the DN to Ingest at the receiving DAAC/mode via e-mail**

Distribution Notice Used for Cross-DAAC (NSIDC to GSFC) Ingest

```
From cmshared@n0ins01u.ecs.nasa.gov  Wed Nov  7 13:02:03 2001
Received: from n0ins01u.ecs.nasa.gov (n0ins01.nsidcb.ecs.nasa.gov [198.118.205.144])
        by g0ins01u.ecs.nasa.gov (8.8.6/8.8.4) with ESMTP
        id NAA02224 for <EcInEmailGWServer_OPS@g0ins01u.ecs.nasa.gov>; Wed, 7 Nov 2001
13:02:02 -0500 (EST)
Received: from n0dis02 (n0dis02.nsidcb.ecs.nasa.gov [198.118.205.131])
      by n0ins01u.ecs.nasa.gov (8.8.8+Sun/8.8.8) with ESMTP id LAA20693
      for <EcInEmailGWServer_OPS@g0ins01u.ecs.nasa.gov>; Wed, 7 Nov 2001 11:02:01 -0700 (MST
From: CM SHARED (CM Code Delivery) <cmshared@n0ins01.nsidcb.ecs.nasa.gov>
Received: by n0dis02 (8.8.8+Sun) id LAA02437; Wed, 7 Nov 2001 11:02:01 -0700 (MST)
Date: Wed, 7 Nov 2001 11:02:01 -0700 (MST)
Message-Id: <200111071802.LAA02437@n0dis02>
To: EcInEmailGWServer_OPS@g0ins01u.ecs.nasa.gov
Subject: ECS Notification

Thank you for using the Earth Observing System Distribution System.   For more information on your
request contact the DAAC.

Please include the data below in any correspondence with the DAAC.

The data distributed for this request can be found on the FTPHOST below in the directory specified by
FTPDIR below.
Thank You!

Support Contacts:
NSIDC User Services
449 UCB
University of Colorado
Boulder, CO  80309
e-mail: nsidc@kryos.colorado.edu
phone: 303-492-6199
fax: 303-492-2468


++++++++++

ORDERID: NONE
REQUESTID: NONE
USERSTRING: Insert of NISE granule
FINISHED: 11/07/2001 11:01:31

MEDIATYPE: FtpPush
FTPHOST: g0acg01u.ecs.nasa.gov
FTPDIR: /usr/ecs/OPS/CUSTOM/icl/a/data/XDAAC_Ingest/NSIDC/
MEDIA 1 of 1
MEDIAID:

      GRANULE:
UR:10:DsShESDTUR:UR:15:DsShSciServerUR:13:[NSC:DSSDSRV]:18:SC:NISE.001:739154
      ESDT: NISE.001

            FILENAME: NISE_SSMIF13_20011107.HDFEOS
            FILESIZE: 2224001

            FILENAME: NISE_SSMIF13_20011107.HDFEOS.met
```

Handling Cross-DAAC or Cross-Mode Ingest (Cont.)

- **Actions that occur as the data are transferred (Cont.):**
  - **The Ingest E-Mail Parser (EcInEmailGWServer) at the receiving DAAC/mode stores the DN as a text file**
    - » **For example:**

      DDIST.notify11072001130203
      in the EmailDirectory (e.g., /usr/ecs/OPS/CUSTOM/data/INS/local/InEmailGWServerPollingDirectory on g0ins01)
  - **While polling the EmailDirectory, the Ingest E-Mail Parser at the receiving DAAC/mode detects files matching the *.notify mask**
  - **The Ingest E-Mail Parser at the receiving DAAC/mode parses the Distribution Notice file**

Handling Cross-DAAC or Cross-Mode Ingest (Cont.)

- **Actions that occur as the data are transferred (Cont.):**
  - **The Ingest E-Mail Parser at the receiving DAAC/mode generates a PDR file**
    - » **For example:**

      DDIST11072001130203.PDR
    - » **When generating the PDR, the Ingest E-Mail Parser uses the ESDT, FTPHOST, FTPDIR, FILENAME, and FILESIZE fields in the Distribution Notice**
    - » **The Ingest E-Mail Parser sets the ORIGINATING_SYSTEM in the PDR to "DDIST"**
    - » **If there is an error in generating a PDR, the e-mail message (Distribution Notice) is moved to the directory specified in the FailedDirectory configuration parameter (e.g., /usr/ecs/OPS/CUSTOM/data/INS/local/InEmailGWServerFailedDirectory)**

PDR from a DN Used for Cross-DAAC (NSIDC to GSFC) Ingest

```
DAN_SEQ_NO = 100;
ORIGINATING_SYSTEM = DDIST;
EXPIRATION_TIME = ;
TOTAL_FILE_COUNT = 2;

AGGREGATE_LENGTH = 2272903;

OBJECT = FILE_GROUP;
  DATA_TYPE = NISE;
  NODE_NAME = g0acg01u.ecs.nasa.gov;
  DATA_VERSION = 001;
  OBJECT = FILE_SPEC;
    DIRECTORY_ID = "/usr/ecs/OPS/CUSTOM/icl/a/data/XDAAC_Ingest/NSIDC/";
    FILE_TYPE = "DDIST";
    FILE_ID = "NISE_SSMIF13_20011107.HDFEOS";
    FILE_SIZE = 2224001;
  END_OBJECT = FILE_SPEC;
  OBJECT = FILE_SPEC;
    DIRECTORY_ID = "/usr/ecs/OPS/CUSTOM/icl/a/data/XDAAC_Ingest/NSIDC/";
    FILE_TYPE = "METADATA";
    FILE_ID = "NISE_SSMIF13_20011107.HDFEOS.met";
    FILE_SIZE = 48902;
  END_OBJECT = FILE_SPEC;
OBJECT=XAR_ENTRY;
  XAR_INFO_COUNT=0;
END_OBJECT = XAR_ENTRY;
END_OBJECT = FILE_GROUP;
```

Handling Cross-DAAC or Cross-Mode Ingest (Cont.)

- **Actions that occur as the data are transferred (Cont.):**
  - **The Ingest E-Mail Parser at the receiving DAAC/mode copies the PDR file to the EcInPolling.DDIST polling directory**
    - » **For example:**

      /usr/ecs/OPS/CUSTOM/data/INS/local/IngestPollingDirectory on g0ins01 at the receiving DAAC/mode
  - **EcInPolling.DDIST at the receiving DAAC/mode detects files matching the *.PDR mask**
  - **EcInPolling.DDIST at the receiving DAAC/mode packages the PDR information into an Ingest Request**
  - **Ingest processing proceeds as a typical polling ingest with delivery record**

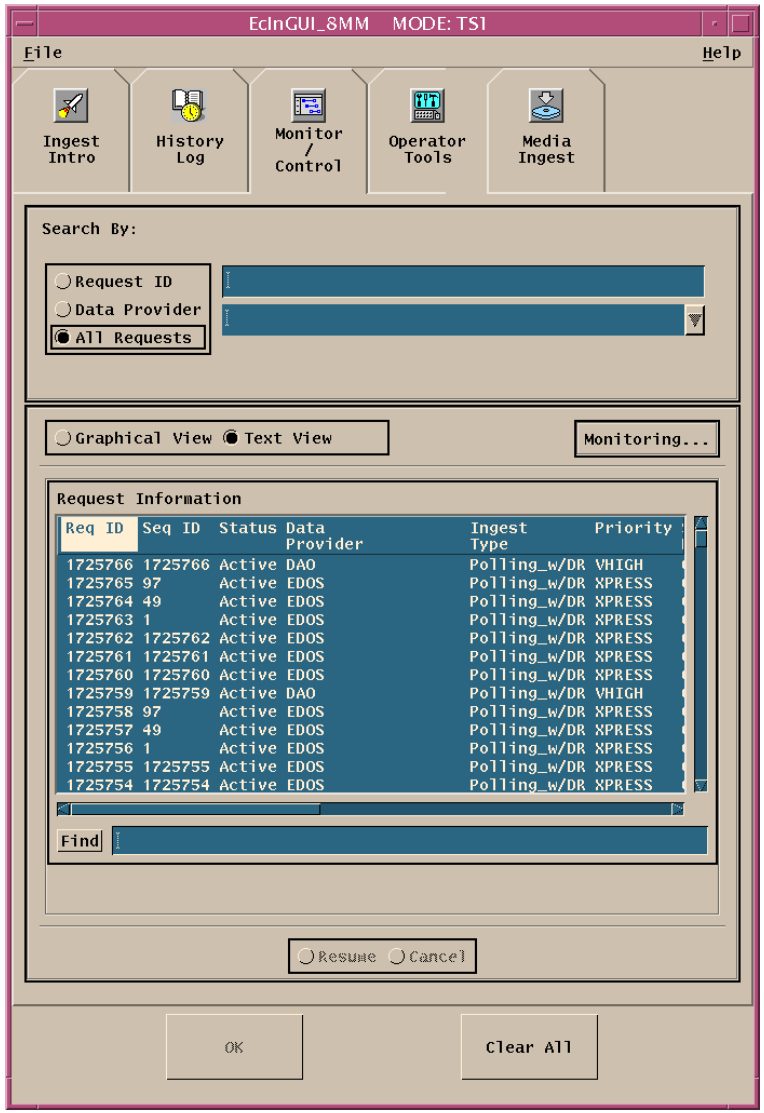
Handling Cross-DAAC or Cross-Mode Ingest (Cont.)

- **Procedure**
  - **Submit a request to User Services (at the DAAC where the data are currently available) to create an order or subscription (as applicable) to have data transferred to a different mode or DAAC**
  - **At the receiving DAAC monitor request processing to ensure that the data are received and ingested**
  - **If the data are not received as expected, contact (e.g., by telephone or e-mail) User Services at the DAAC where the order or subscription was submitted to determine the nature of the problem and have it corrected**

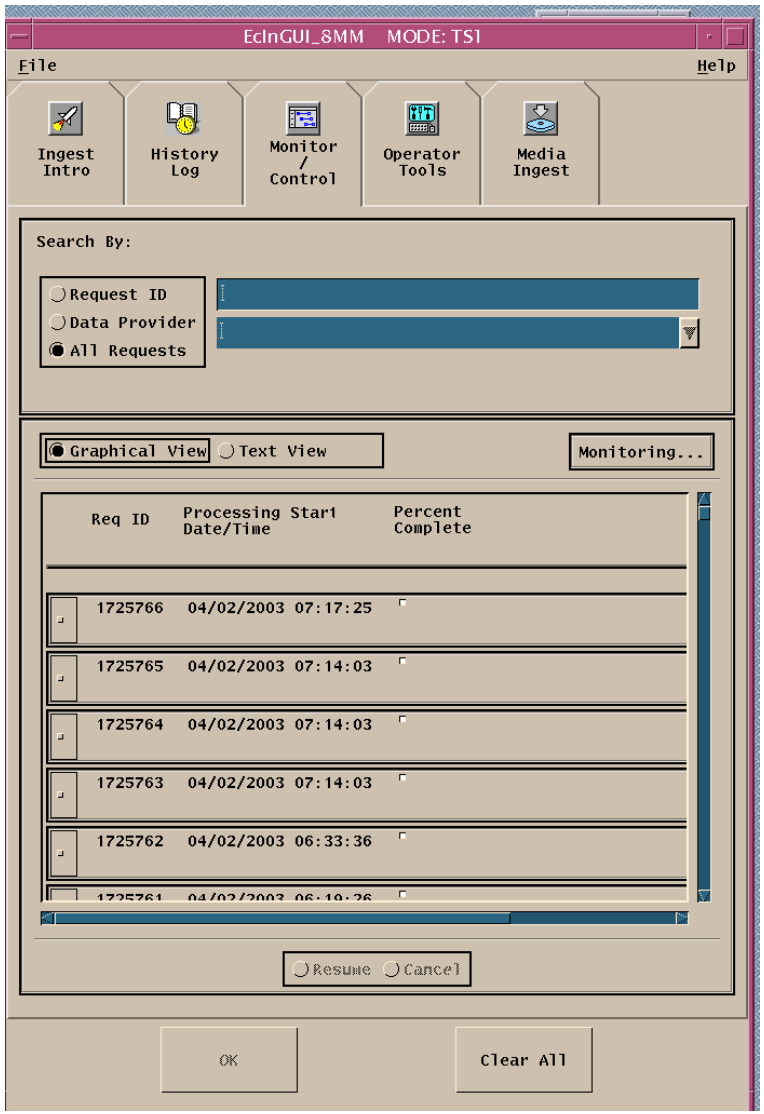
Monitoring Ingest Status

- **Assumptions:**
  - Ingest processes have been started
  - System is operating normally
  - Data are ready for ingest
  - Several PDR files have been received and logged by the system; the specific ingest processes have been assigned request IDs
- **Invoke monitoring display with "Monitoring/Controlling Ingest Requests" procedure**

Monitor/Control Tab: Text View



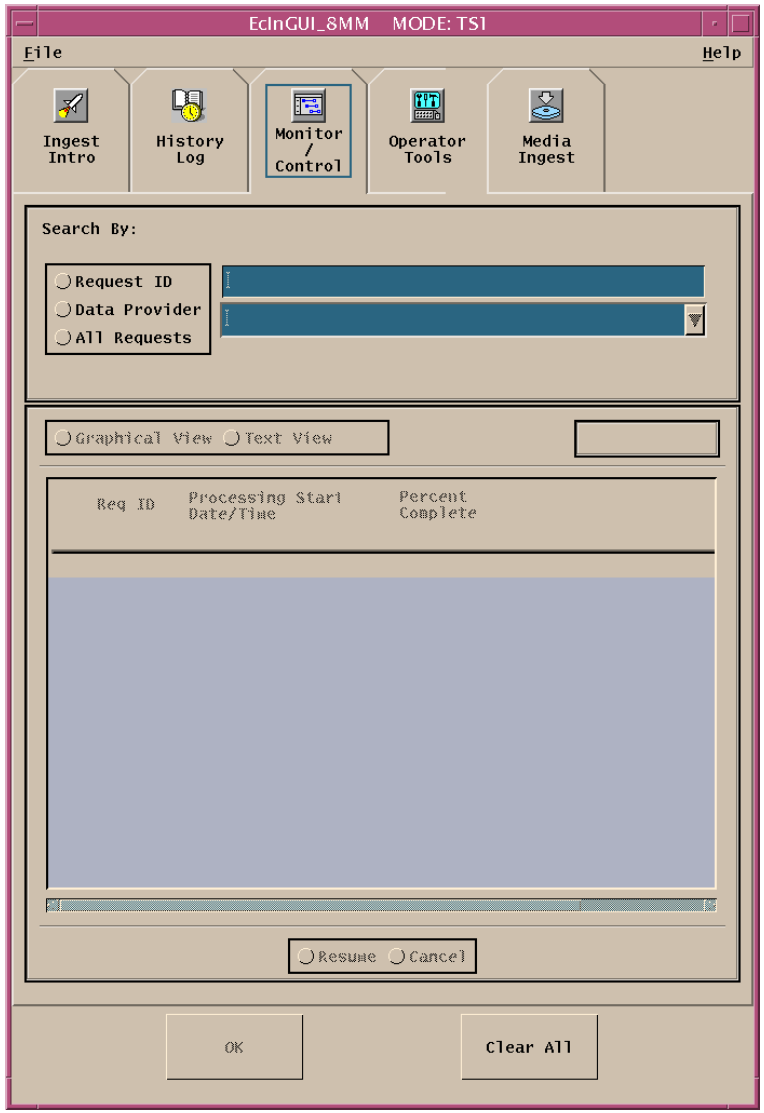
Monitor/Control Tab: Graphical View



Monitoring/Controlling Ingest Requests

- **Procedure**
  - Select the Ingest GUI Monitor/Control tab
  - Select the appropriate set of ingest requests
  - Select the type of view (i.e., graphical or text)
  - Observe ingest request processing
  - Resume/cancel requests as necessary
  - Troubleshoot ingest problems as necessary

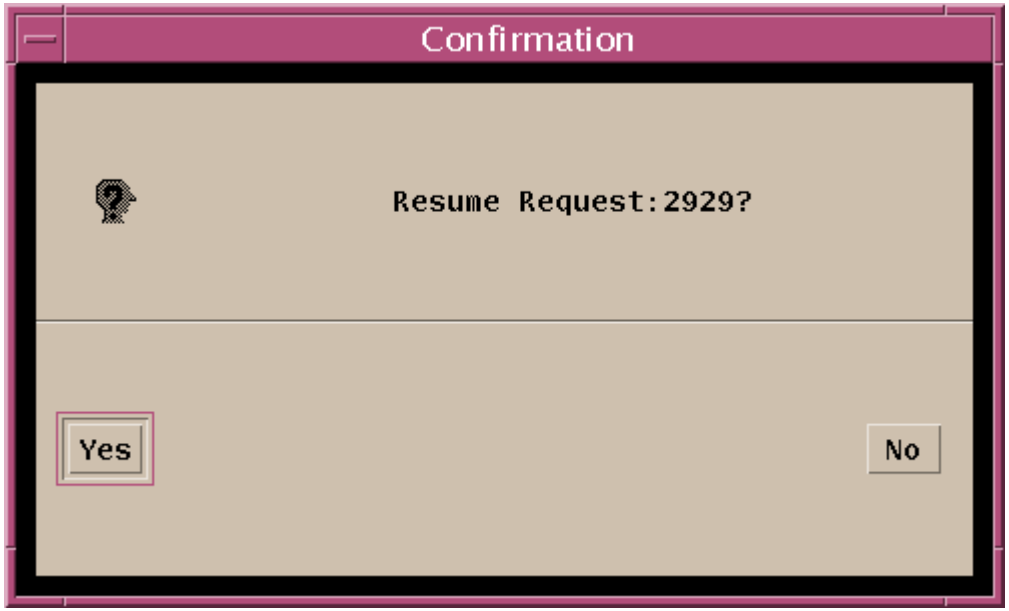
ECS Ingest GUI Monitor/Control Tab



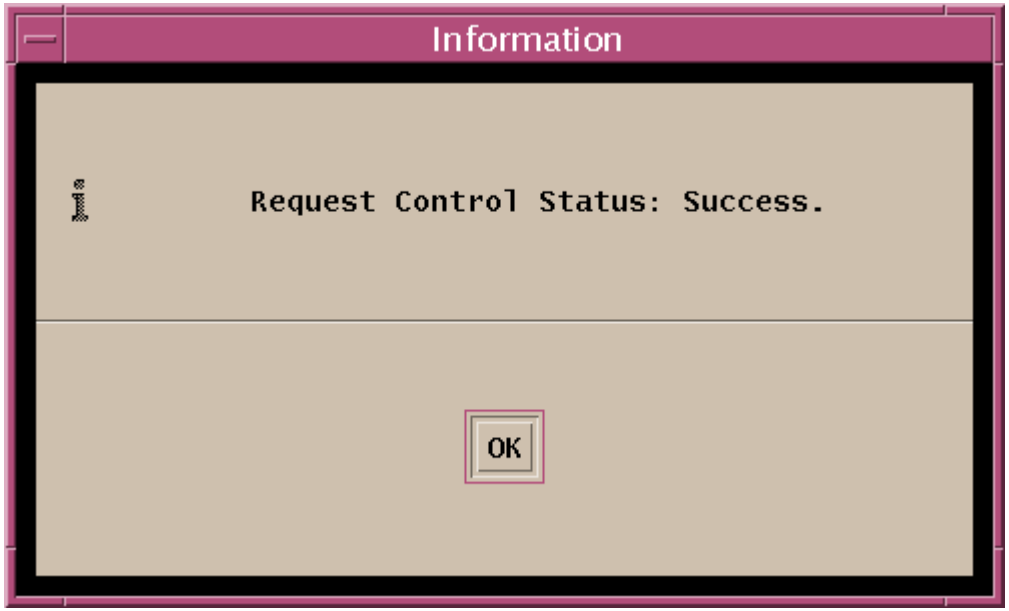
Monitoring Ingest Requests: Resuming Ingest Requests

- **Procedure**
  - Click on the row corresponding to the request or granule to be resumed on the Monitor/Control tab (Text View)
  - Click on the Resume button
  - Click on the OK button at the bottom of the GUI
  - Click on the Yes button in the Resume Request Confirmation Dialogue Box to confirm resuming the request/granule
  - Click on the OK button in the Request Control Status Information Dialogue Box

Resume Request Confirmation Dialogue Box



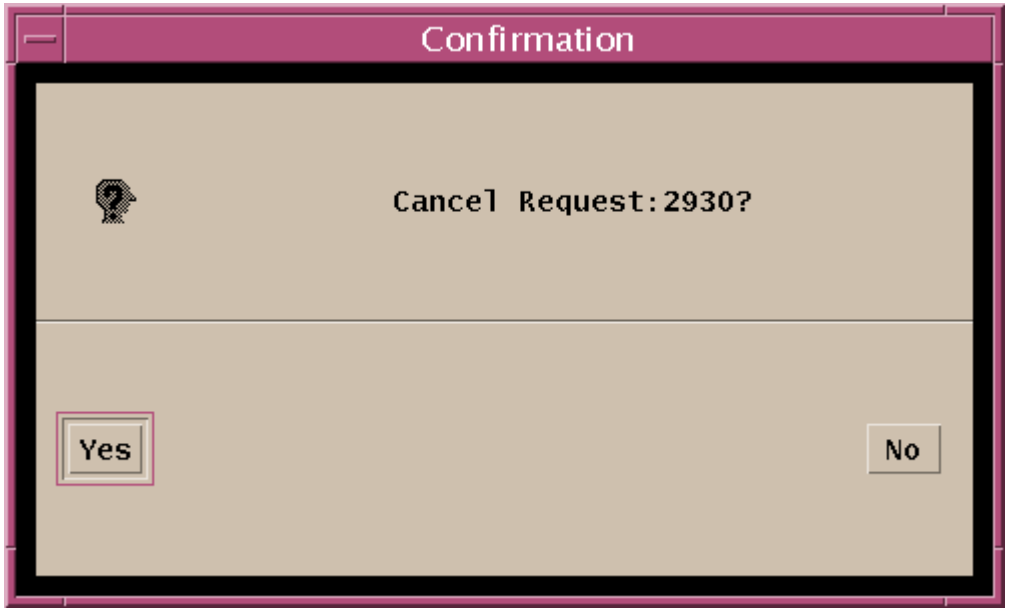
Request Control Status Information Dialogue Box



Monitoring Ingest Requests: Canceling Ingest Requests

- **Procedure**
  - Click on the row corresponding to the request or granule to be canceled on the Monitor/Control tab (Text View)
  - Click on the Cancel button
  - Click on the OK button at the bottom of the GUI
  - Click on the Yes button in the Cancel Request Confirmation Dialogue Box to confirm canceling the request/granule
  - Click on the OK button in the Request Control Status Information Dialogue Box

Cancel Request Confirmation Dialogue Box



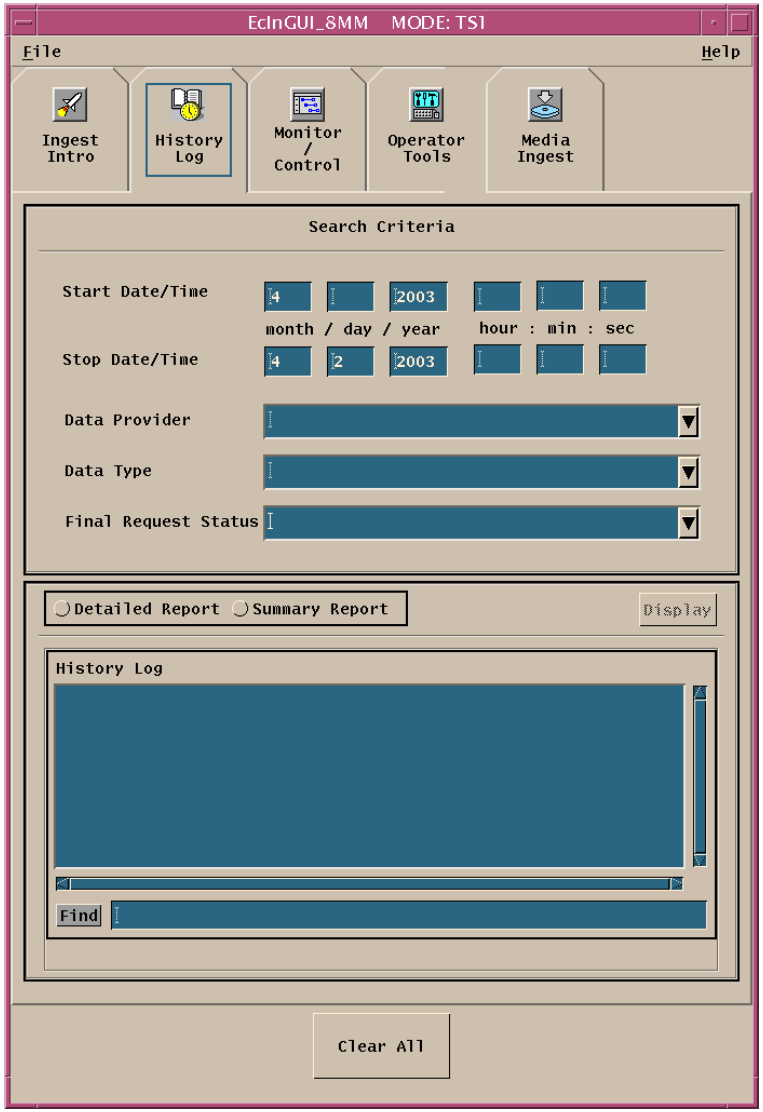
Ingest History Log

- Upon Ingest completion...
  - Notice automatically sent to data provider indicating the status of the ingested data
  - Data provider sends an acknowledgment of notice
  - Receipt of the acknowledgment logged by ECS
  - Request ID removed from the list of active requests
  - History log receives statistics on the completed transaction
- History Log search criteria
  - time period
  - data provider ID
  - data type
  - final request status

Ingest History Log (Cont.)

- **Ingest History Log formats**
  - **Detailed Report - detailed information about each completed ingest request**
  - **Summary Report - summary of ingest processing statistics, including the average and maximum time taken to perform each step in the ingest process**
    - » **Request-level Summary Report - ingest request processing statistics**
    - » **Granule-level Summary Report - ingest granule processing statistics organized by data provider and Earth Science Data Type (ESDT)**

Ingest History Log Screen



Viewing Ingest History Log

- **Procedure**
  - Select the Ingest GUI History Log tab
  - Select the search criteria
    - » time period
    - » data provider
    - » data type
    - » final request status
  - Select Detailed Report or Summary Report
  - If Summary Report, select either Request Level report or Granule Level report
  - Click on the Display button
  - Observe ingest request information displayed in the History Log/Processing Statistics field

History Log Detailed Report

EcInGUI_8MM MODE: TS1

File Help

Ingest Intro | History Log | Monitor / Control | Operator Tools | Media Ingest

Search Criteria

Start Date/Time 4 1 2003 7 47 6

month / day / year hour : min : sec

Stop Date/Time 4 2 2003 7 47 6

Data Provider

Data Type

Final Request Status

Detailed Report Summary Report Display

History Log

Req ID	Data Provider	Status	Ingest Type	Start Date
1701397	DAO	Successful	Polling_w/DR	04/01/2003
1701398	HIRDLS_SIPS	Successful	Polling_w/DR	04/01/2003
1701399	FDDAQUA	Successful	Polling_w/DR	04/01/2003
1701400	EDOS	Successful	Polling_w/DR	04/01/2003
1701401	EDOS	Successful	Polling_w/DR	04/01/2003
1701402	EDOS	Successful	Polling_w/DR	04/01/2003
1701403	EDOS	Successful	Polling_w/DR	04/01/2003

Find

Clear All

History Log Summary Report - Request Level

EclnGUI_8MM MODE: TS1

File Help

Ingest Intro | History Log | Monitor / Control | Operator Tools | Media Ingest

Search Criteria

Start Date/Time 4 1 2003 7 49 27

month / day / year hour : min : sec

Stop Date/Time 4 2 2003 7 49 27

Data Provider

Data Type

Final Request Status

Request level Granule level Display

Processing Statistics

Data Provider	Ttl Reqs	Ttl Errs	Gran Avg	Gran Max	File Avg	File Max	Size (MB) Avg	Size (MB) Max
DAO	21	0	1	1	1	1	104.634	375.745
EDOS	522	0	1	1	2	2	0.114	0.450
EMOS_HIST	15	0	1	1	2	2	4.809	4.811
FDD	1	0	15	15	15	15	15.000	15.000
FDDAQUA	5	0	1	1	1	1	4.847	4.847
GSFC-V0	58	0	1	4	1	4	31.283	104.598
HIRDLS_SIPS	15	0	1	1	2	2	0.003	0.007

Find

Go Back

Clear All

History Log Summary Report - Granule Level

EcInGUI_8MM MODE: TS1

File Help

Ingest Intro | History Log | Monitor / Control | Operator Tools | Media Ingest

Search Criteria

Start Date/Time 4 1 2003 7 51 52

month / day / year hour : min : sec

Stop Date/Time 4 2 2003 7 51 52

Data Provider

Data Type

Final Request Status

Request level Granule level Display

Processing Statistics

Data Provider	Data Type	Total Granules	Total Errors	File Avg	File Max	Size (MB) Avg	Si (M Ma
DAO	DFLAPCHM	2	0	1	1	8.348	8.
DAO	DFLAPMIS	1	0	1	1	300.620	30
DAO	DFLAXCHM	2	0	1	1	0.578	0.
DAO	DFLAXCLD	2	0	1	1	29.517	29
DAO	DFLAXENG	1	0	1	1	67.310	67
DAO	DFLAXLSM	1	0	1	1	25.316	25
DAO	DFLAXMIS	1	0	1	1	46.313	46

Find

Go Back

Clear All

Verifying the Archiving of Ingested Data

- Check the proper directory on the appropriate File and Storage Management System (FSMS) host (e.g., g0drg01)
  - Directories are identified by the type of data (e.g., aster, ceres, l7, modis) in them and correspond directly to tape volumes in the system
  - Just a matter of checking the relevant FSMS directory to determine whether the applicable files/granules have been transferred
  - Procedure does not involve the use of any archive software
  - Before starting it is essential to know what data to look for
    - » End Date(s)/Time(s) and Data Volume(s) for ingest requests shown on the ECS Ingest GUI

Verifying the Archiving of Ingested Data (Cont.)

- **Procedure**
  - Access a terminal window logged in to the appropriate FSMS Server host
  - Change directory to the directory containing the archive data
  - Perform a long listing of directory contents for a particular date (e.g., ls -la | grep 'Mar 17')
    - » It is important to limit the listing (e.g., to a particular day)
    - » If there are tens of thousands of granules in the directory, just doing a listing of the entire directory would cause serious performance problems
  - Compare End Date(s)/Time(s) and Data Volume(s) for the applicable ingest request(s) shown on the Ingest GUI with the dates/times and file sizes listed for the files in the directory

Cleaning Polling Directories

- Polling directories should be cleaned up after successful archiving to avoid running out of disk space
- Automatic clean-up should be available
- However, it may still be useful to know how to use the clean-up scripts
- Procedure
  - Access a terminal window logged in to the Operations Workstation host
  - Type command to start clean-up script
  - Type appropriate responses to clean-up script prompts

Performing Hard (Physical) Media Ingest

- **Types of Hard Media Ingest**
  - **Three (3) general methods for performing hard media ingest**
  - **Each one uses one of the following interfaces:**
    - » **ECS Ingest GUI media interface**
    - » **INGEST Media Tape Reader GUI**
    - » **UNIX commands for reading data from tape and generating PDR and signal files**

Performing Hard (Physical) Media Ingest

- Both the ECS Ingest GUI (EcInGUI) media interface and the INGEST Media Tape Reader GUI (EcInTapeReaderGUI) are used for ingesting data from either 8mm or DTF-2 tape cartridges
- The GUIs support hard media ingest in the following circumstances:
  - » 8mm tape ingest of L7 IGS browse data and metadata
  - » DTF-2 tape ingest of ASTER L1A/L1B data

NOTE: Currently the INGEST Media Tape Reader GUI is the preferred interface for media ingest of L7 IGS browse/metadata and ASTER L1A/L1B data. Eventually the ECS Ingest GUI (EcInGUI) media interface may be phased out.

Performing Hard (Physical) Media Ingest

- UNIX commands are used primarily for performing ingest of EDOS L0 replacement data from DTF-2 tape
  - » However, it is possible to use UNIX commands under other circumstances to transfer data from either 8mm or DTF-2 tapes into directories where the data can be picked up by an appropriate polling process
- In all cases of media ingest the Ingest Technician uses the ECS Ingest GUI (Monitor/Control screen) to monitor and control ingest request processing after the data have been copied from tape and an ingest request has been generated

Performing Hard (Physical) Media Ingest (Cont.)

- **Using the ECS Ingest GUI Media Interface to Perform Media Ingest**
  - **The Ingest Technician uses the ECS Ingest GUI to put the PDR and data files into the applicable directories and create the appropriate ingest request(s), which is (are) sent to the Ingest Request Manager**
  - **The Ingest Request Manager packages each request into granules and sends them to the Granule Server, which requests the Science Data Server to insert the data and/or metadata into the archive and/or catalog**

Performing Hard (Physical) Media Ingest (Cont.)

- **Labeling Tape Cartridges with Bar Codes**
  - **DAAC policy may require a bar-code label on each tape that contains data to be ingested**
  - **Labels are typically already on the tapes when received from the data provider**
    - » **However, the Ingest Technician may affix the labels to the tape cartridges if necessary**

Performing Hard (Physical) Media Ingest (Cont.)

- **Performing Media Ingest from 8mm Tape Using the ECS Ingest GUI Media Interface**
  - **Assumptions:**
    - » **The PDR/PMPDR file is available, either placed on the network by the data provider or embedded in the media**
    - » **If applicable, the contents of the PDR/PMPDR on the tape have been compared with the contents of the hardcopy version of the PDR/PMPDR and there are no discrepancies**
    - » **All applicable servers are currently running**

Performing Hard (Physical) Media Ingest (Cont.)

- **Procedure (Performing Media Ingest from 8mm Tape Using the ECS Ingest GUI Media Interface)**
  - Load the tape into a stacker as described in the procedure for Unloading and Loading Stackers
  - Launch an 8mm instance of the ECS Ingest GUI
  - Select the Ingest GUI Media Ingest tab
  - Identify the type of medium
  - Select the data provider
  - Enter the media ID
  - Identify the delivery record file location
  - Initiate and monitor the data transfer
  - Monitor request processing
  - When done, exit from the media instance of the GUI

Media Ingest Tab

EcInGUI_8MM MODE: TS1

File Help

Ingest Intro | History Log | Monitor / Control | Operator Tools | Media Ingest

Media Type: DTF Tape

Data Provider: ASTERGDS

Media ID:

Data Delivery Record File Location

- On Network
- Embedded in Media

Data Delivery Record File Name:

OK | Clear All

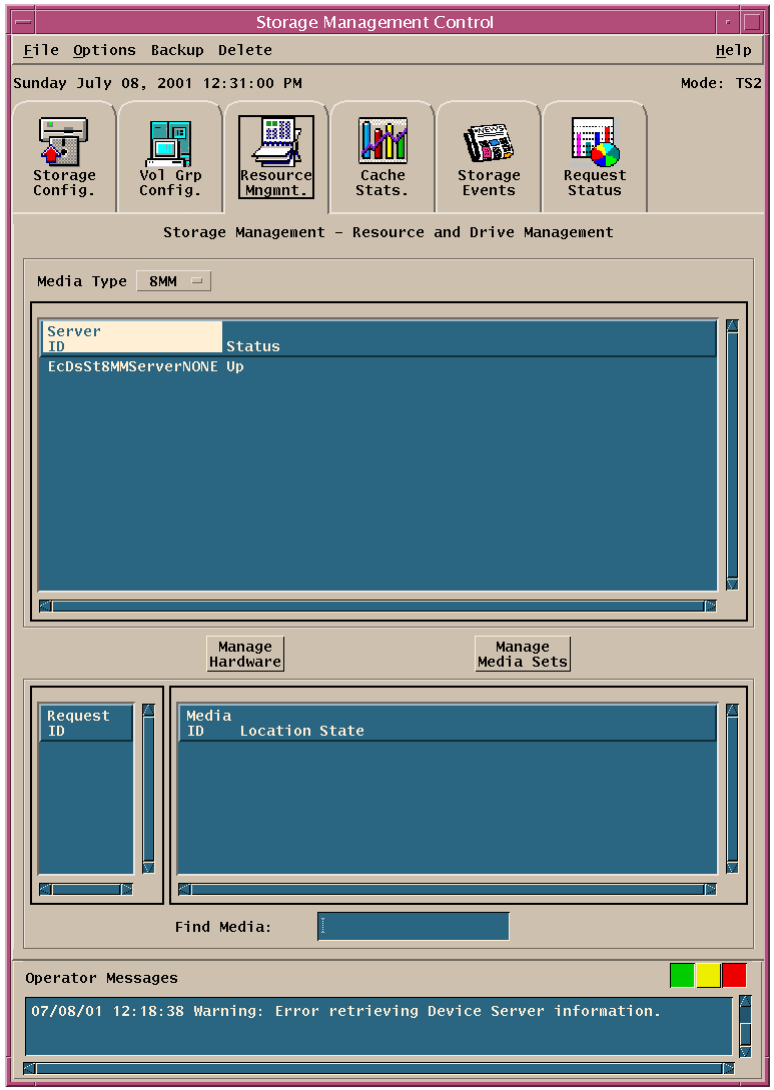
Unloading and Loading Stackers

- **Unloading and loading stackers involves the following activities:**
  - **Unloading a tape stacker**
  - **Loading a tape stacker**

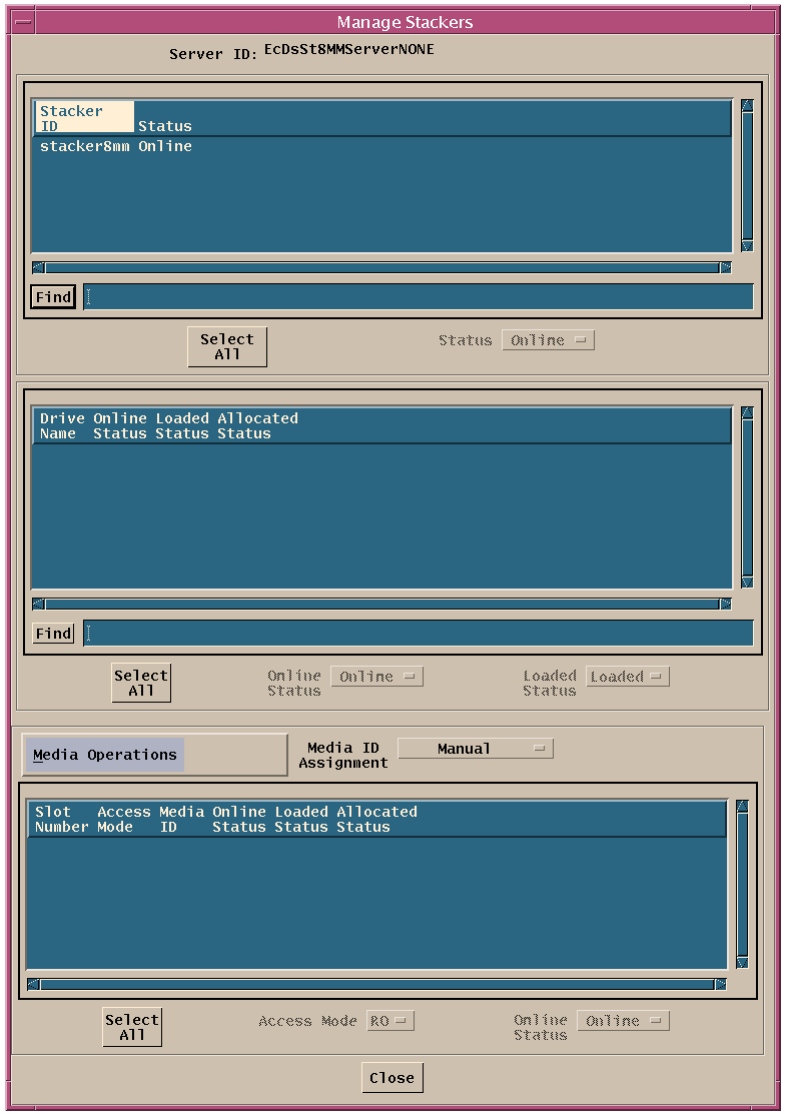
Unloading and Loading Stackers (Cont.)

- **Procedure**
  - Select the Resource Mngmnt tab of the Storage Management Control GUI
  - Select the Manage Stackers window
  - Open the stacker door
  - Remove the magazine (cartridge holder)
  - Set the write-protect switch (e.g., red tab) on each tape cartridge to be loaded
  - Load the magazine (cartridge holder) with tape(s)
  - Verify that the ”Online Status" associated with each slot number is correct
  - Enter media IDs if applicable using the Media Operations pull-down menu
  - Load the stacker with tape(s)
  - Close the Manage Stackers window

Storage Management Control GUI: Resource Mngmnt. Tab



Storage Management Control GUI: Manage Stackers Window

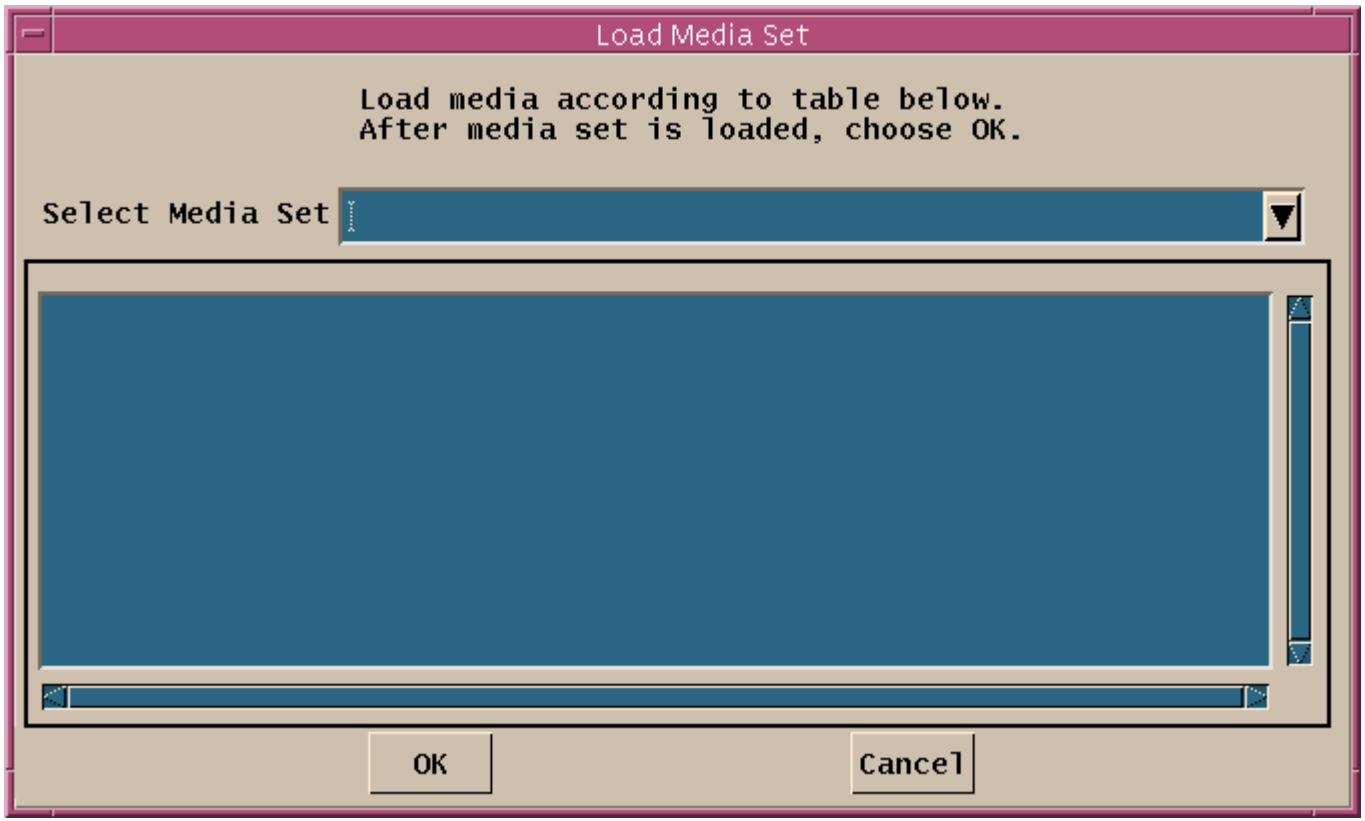


Load Media Window (Stacker Management)

Replace Media Window

Unload Media Window (Stacker Management)

Load Media Set Window



Performing Hard (Physical) Media Ingest (Cont.)

- **Procedure (Performing Media Ingest from DTF-2 Tape Using the ECS Ingest GUI Media Interface)**
  - Load the tape into a DTF-2 drive as described in the appropriate loading procedure
  - Launch a DTF instance of the ECS Ingest GUI
  - Select the Ingest GUI Media Ingest tab
  - Identify the type of medium
  - Select the data provider
  - Enter the media ID
  - Identify the delivery record file location
  - Initiate and monitor the data transfer
  - Monitor request processing
  - When done, unload the tape

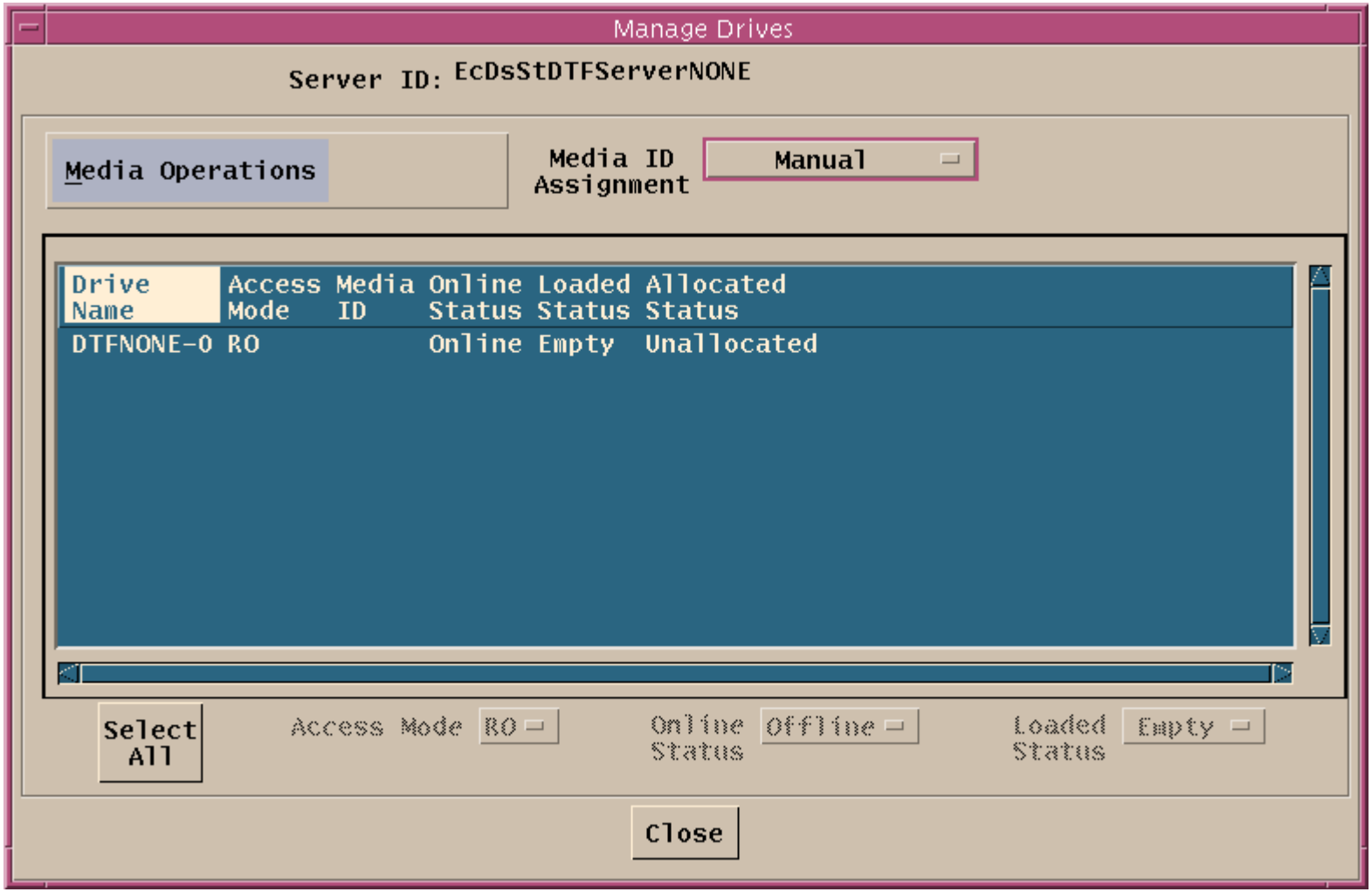
Performing Hard (Physical) Media Ingest (Cont.)

- **Loading a DTF-2 Drive for Use with the ECS Ingest GUI**
  - **A DTF-2 drive supports the reading of data from several types of cassettes, including cassettes of the following types:**
    - » **DTF-2 L [large] cassette**
    - » **DTF-2 S [small] cassette**
    - » **DTF-1 L [large] cassette**
    - » **DTF-1 S [small] cassette**
  - **Assumptions:**
    - » **All applicable servers are currently running**
    - » **The Storage Management Control GUI is running**
    - » **The DTF-2 drive is in service**

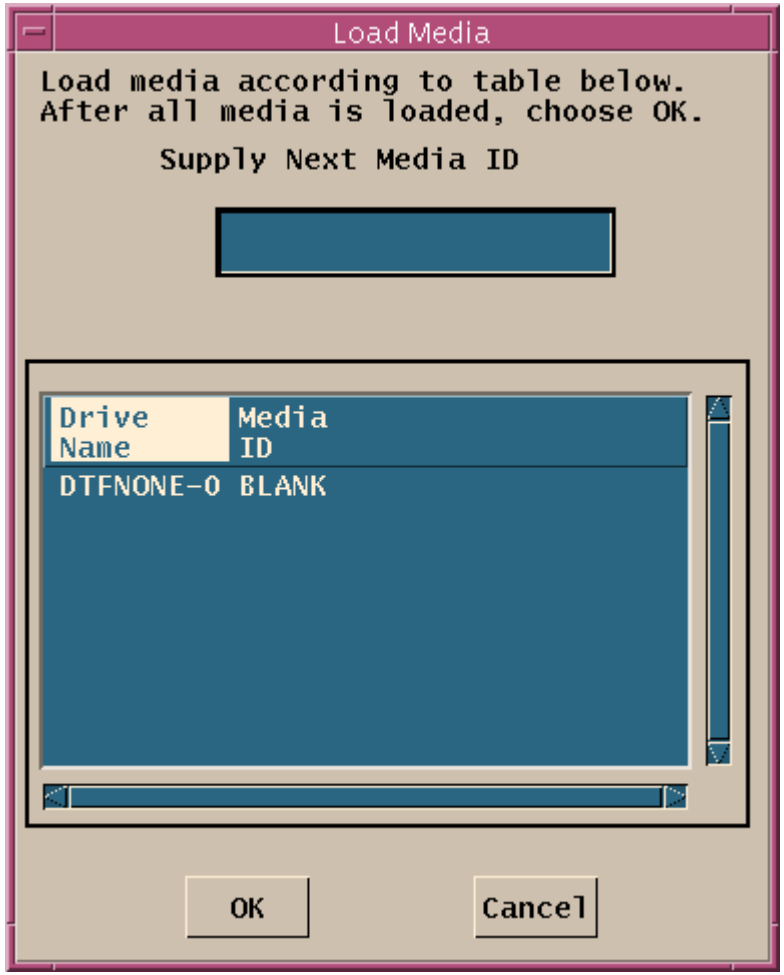
Performing Hard (Physical) Media Ingest (Cont.)

- **Procedure (Loading a DTF-2 Drive)**
  - Insert the Sony DTF tape into the cassette slot of the DTF tape drive
  - Wait for the tape to complete the loading process
  - Click on the Resource Mngmnt tab on the Storage Management Control GUI
  - Select the line corresponding to the relevant server in the Media Type window
  - Click on the Manage Hardware button
  - Select the line in the drive information window corresponding to the drive to be loaded
  - Select the Load Media option
  - Type the media ID in the Media ID field then press Return/Enter
  - Click on the OK button

Storage Management Control GUI: Manage Drives Window



Load Media Window (Drive Management)



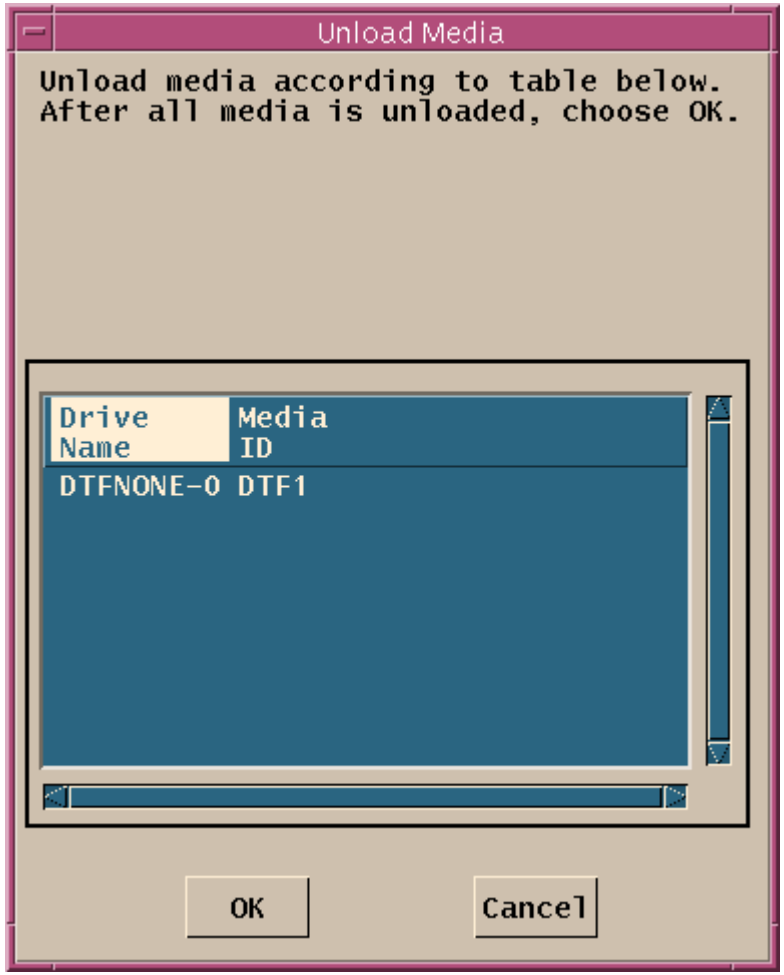
Performing Hard (Physical) Media Ingest (Cont.)

- **Unloading a DTF-2 Drive for Use with the ECS Ingest GUI**
  - **Assumptions:**
    - » **All applicable servers are currently running**
    - » **The Storage Management Control GUI is running**
    - » **The DTF-2 drive in service**
    - » **There is a DTF-2 tape in the drive**

Performing Hard (Physical) Media Ingest (Cont.)

- **Procedure (Unloading a DTF-2 Drive for Use with the ECS Ingest GUI)**
  - Push the UNLOAD button
  - After it has completed unloading, remove the DTF-2 tape cartridge from the tape drive cassette slot
  - Select the Resource Mngmnt tab on the Storage Management Control GUI
  - Select the relevant server in the Media Type window
  - Click on the Manage Hardware button
  - Select the line for to the drive to be unloaded
  - Select the Unload Media option
  - Select the line in the Unload Media window
  - Click on the Unload Media window OK button
  - Click on the Manage Drives window Close button

Unload Media Window (Drive Management)



Performing Hard (Physical) Media Ingest (Cont.)

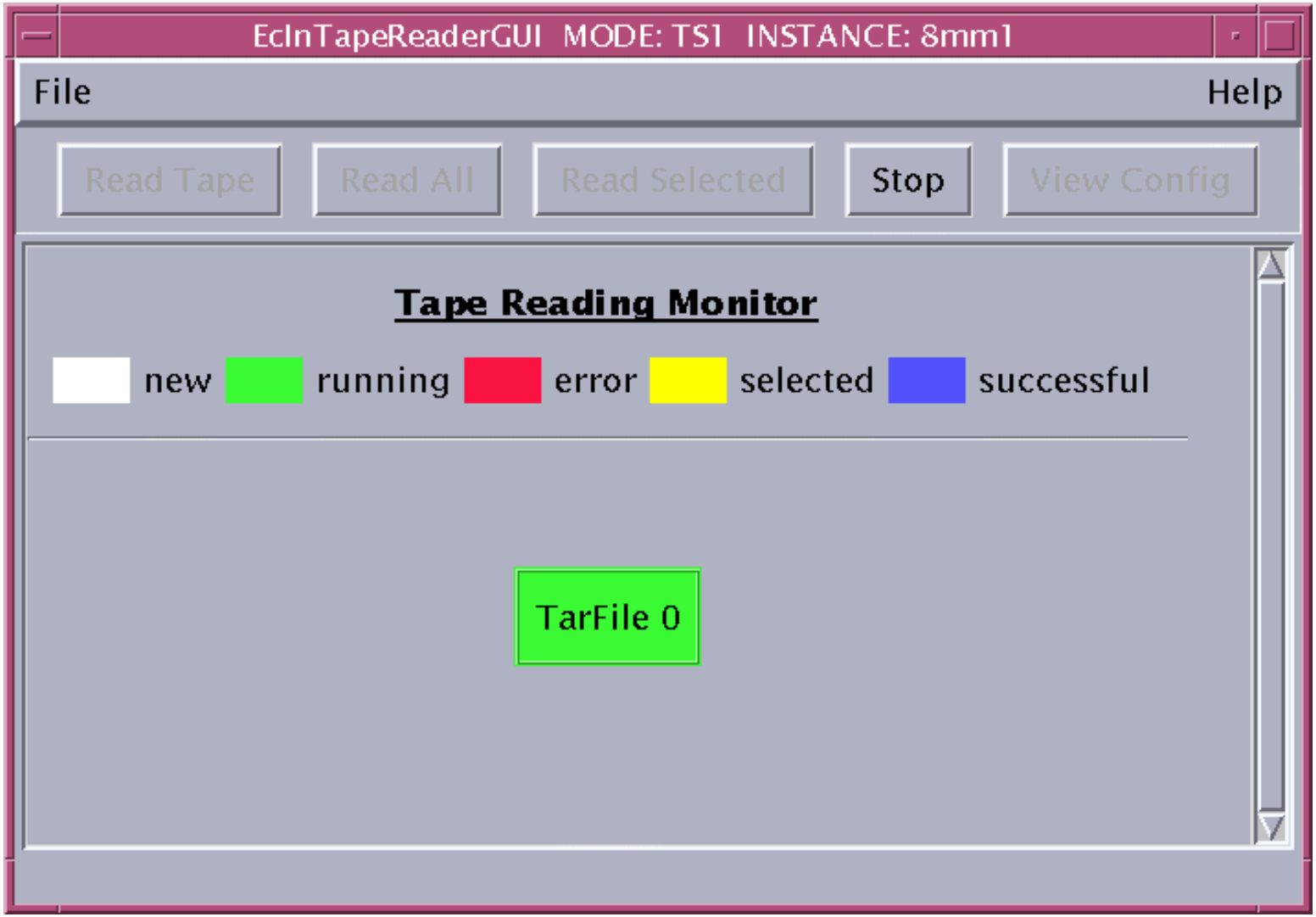
- **Using the INGEST Media Tape Reader GUI to Perform Media Ingest**
  - The Ingest Technician uses the INGEST Media Tape Reader GUI to generate the PDR(s) for the data files on the tape and put the PDR(s) and data files into the proper PDR and data directories for polling
  - The corresponding ingest polling process picks up the PDRs in the polling directory and creates the appropriate ingest request(s), which is (are) sent to the Ingest Request Manager
  - The Ingest Request Manager packages each request into granules and sends them to the Granule Server, which requests the Science Data Server to insert the data and/or metadata into the archive and/or catalog

Performing Hard (Physical) Media Ingest (Cont.)

- **Procedure (Performing Media Ingest Using the INGEST Media Tape Reader GUI)**
  - Load the tape as described in the appropriate procedure
  - Launch the INGEST Media Tape Reader GUI
  - Click on Read Tape button
  - Identify the file(s) to be ingested
  - Initiate reading of the file(s) to be ingested
  - Monitor request processing
  - When done, exit from the INGEST Media Tape Reader GUI
  - When done, unload the tape as described in the appropriate procedure

INGEST Media Tape Reader GUI: Welcome Screen

INGEST Media Tape Reader GUI: Monitor Screen



INGEST Media Tape Reader GUI: Current Configuration Screen

EcInTapeReaderGUI MODE: TS1 INSTANCE: 8mm

File Help

Read Tape | Read All | Read Selected | Stop | View Config

Welcome to INGEST Media Tape Reader

Current Configuration

MediaType:	8mm Tape
TapeDrive:	/dev/rmt/0n
DataDirectory:	/usr/ecs/TS1/CUSTOM/ddist/p2dis02/data/Media_Staging
PollingDirectory:	/usr/ecs/TS1/CUSTOM/ddist/p2dis02/data/Media_Staging/polling

Go Back

Performing Hard (Physical) Media Ingest (Cont.)

- **Procedure (Manually Loading an 8mm Tape Cartridge into a Tape Drive in an 8mm Tape Stacker)**
  - **Unlock the stacker door**
  - **Open the stacker door**
  - **Insert the tape cartridge into the cartridge slot of the appropriate tape drive**
  - **Close the stacker door**
  - **Lock the stacker door**

Performing Hard (Physical) Media Ingest (Cont.)

- **Procedure (Manually Unloading an 8mm Tape Cartridge from a Stacker)**
  - **Unlock the stacker door**
  - **Open the stacker door**
  - **Press the eject button**
  - **Remove the tape cartridge from the tape drive**

Performing Hard (Physical) Media Ingest (Cont.)

- Procedure (Unloading and Loading 8mm Tape Stackers for Sequential Mode Operation)
  - Select the Control Mode Menu on the stacker front panel
  - Select Sequential 1
  - Return to the Main Screen
  - Unlock the stacker door
  - Open the stacker door
  - Remove the magazine (cartridge holder)
  - Remove tape cartridges from the magazine
  - Insert tape cartridges in the magazine
  - Replace the magazine in the stacker
  - Close the stacker door
  - Lock the stacker door

Performing Hard (Physical) Media Ingest (Cont.)

- **Procedure (Performing DTF-2 Drive Loading)**
  - **Verify that No Tape 0H is indicated on the DTF tape drive's display window**
  - **Insert the Sony DTF tape into the cassette slot of the DTF tape drive**
  - **Wait for Loaded [00] to be indicated on the DTF tape drive's display window**

Performing Hard (Physical) Media Ingest (Cont.)

- **Procedure (Performing DTF-2 Drive Unloading)**
  - **Push the UNLOAD button on the front of the DTF tape drive**
  - **Wait for the DTF tape drive to complete the unloading process**
  - **Remove the DTF-2 tape cartridge from the tape drive cassette slot**

Performing Hard (Physical) Media Ingest (Cont.)

- **Using UNIX Commands to Perform Media Ingest**
  - Many ICDs between ECS and data providers state that tape media (especially 8mm tape) may be used as a backup method of ingest in case of emergency
  - If neither the ECS Ingest GUI nor the INGEST Media Tape Reader GUI has been set up for media ingest from a data provider that elects to supply some data on a hard medium, it is possible to ingest the data using UNIX commands
  - A special case of media ingest using UNIX commands involves the ingest of EDOS L0 replacement data from DTF-2 tape

Performing Hard (Physical) Media Ingest (Cont.)

- **Using UNIX Commands to Perform Media Ingest (Cont.)**
  - **There are two procedures related to using UNIX commands to perform media ingest:**
    - » **Performing Media Ingest Using UNIX Commands**
    - » **Performing Ingest of Data from EDOS DTF-2 Archive Tapes**
  - **In both instances the Ingest Technician uses UNIX commands to read the data from the tape, then ensures that any necessary delivery records (and signal files, if applicable) are available in the polling directory**
  - **The applicable ingest polling process picks up the delivery records (PDRs) in the polling directory and creates the appropriate ingest requests, which are sent to the Ingest Request Manager**

Performing Hard (Physical) Media Ingest (Cont.)

- **Performing Media Ingest Using UNIX Commands**
  - **If a data provider furnishes data on a hard medium for ingest with delivery record, the following three types of files should be present on the medium:**
    - » **Physical Media Product Delivery Record (PMPDR) or Product Delivery Record (PDR)**
    - » **Metadata file(s)**
    - » **Data file(s), typically in tar (tape archive) format**
  - **If a data provider furnishes data on a hard medium for ingest without delivery record, only the data file(s) will be present on the medium**
  - **Each medium should have a label and there should be a separate hardcopy document identifying the names of files contained on the medium and the order in which the files have been written**

Performing Hard (Physical) Media Ingest (Cont.)

- **Procedure (Performing Media Ingest Using UNIX Commands)**
  - **Load the medium in the appropriate drive**
  - **Access a terminal window logged in to the host (e.g., Sun Consolidation Internal Server) with access to the drive**
  - **Type tar xvf *device* to retrieve files**
  - **Use scp2 to copy files to the Ingest polling host**
  - **If applicable, use vi editor commands to generate a PDR file**
  - **If applicable, use scp2 to copy the PDR to the polling directory on the Ingest polling host**
  - **Monitor Ingest request processing**
  - **Unload the drive**
  - **Verify that the data have been inserted into the archive**

Performing Hard (Physical) Media Ingest (Cont.)

- **Performing Ingest of Data from EDOS DTF-2 Archive Tapes**
  - In response to a request (via the EOS Data Re-Order web tool) from GES DAAC Operations EDOS furnishes L0 replacement data to the DAAC on DTF-2 tapes
  - Personnel from the Level Zero Processing Facility (LZPF) hand the DTF-2 archive tape(s) containing the requested data to personnel at GES DAAC Operations
    - » DAAC Operations returns each tape to the LZPF as soon as the needed data have been archived in ECS
  - In addition to the requested PDS(s) each EDOS DTF-2 archive tape is likely to contain some additional (unneeded) PDSs
    - » The Ingest Technician and/or Archive Manager need to determine which PDS(s) on the tape should be inserted into the archive

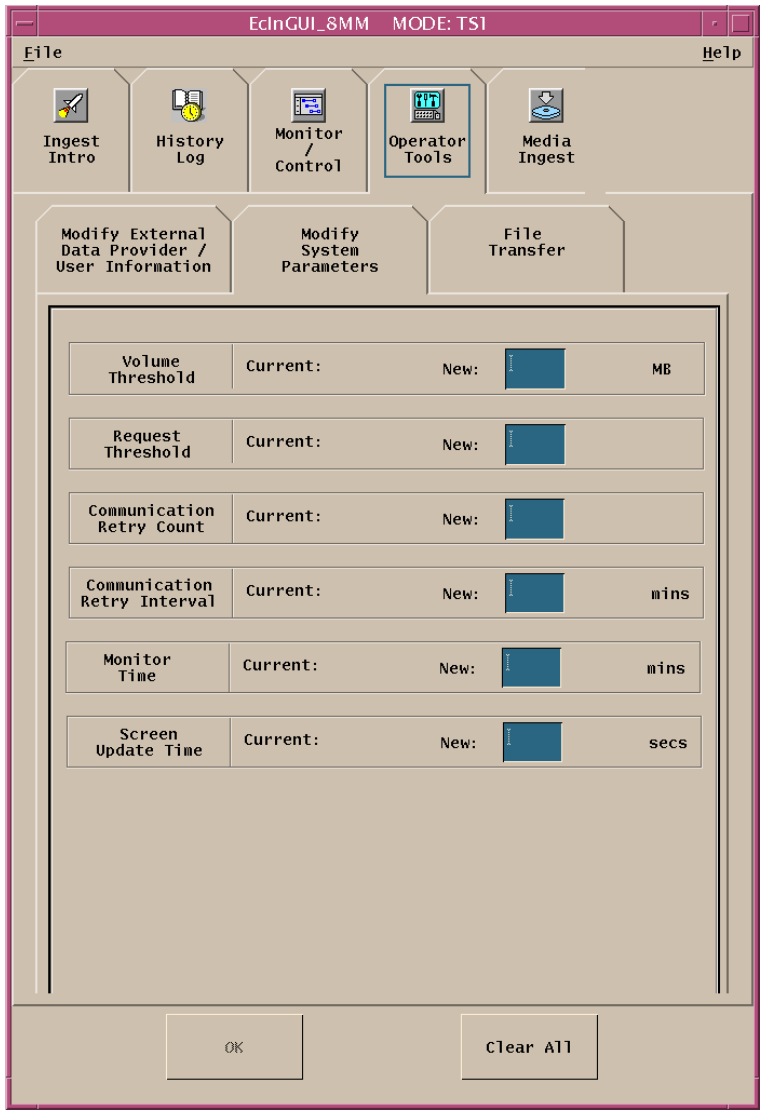
Performing Hard (Physical) Media Ingest (Cont.)

- Procedure (Performing Ingest of Data from EDOS DTF-2 Archive Tapes)
  - Load the tape containing the data to be ingested
  - Log in to the proper host (e.g., Sun Consolidation Internal Server)
  - Untar the files on the tape
  - Read the contents of the PPMUDR to identify the PDS(s) to be archived
  - Copy the files to be ingested to the pollEDOS directory
  - Create a PDR for the files to be ingested
  - Create a signal file
  - Monitor Ingest request processing
  - Unload the tape drive
  - Verify that the data have been inserted into the archive

Ingest Tunable Parameters and File Transfers

- Operator Tools Tab
  - Two GUI screens to view and set ingest thresholds
    - » Modify External Data Provider/User Information
    - » Modify System Parameters
  - One GUI screen for transferring files
    - » File Transfer

ECS Ingest GUI Operator Tools Tab



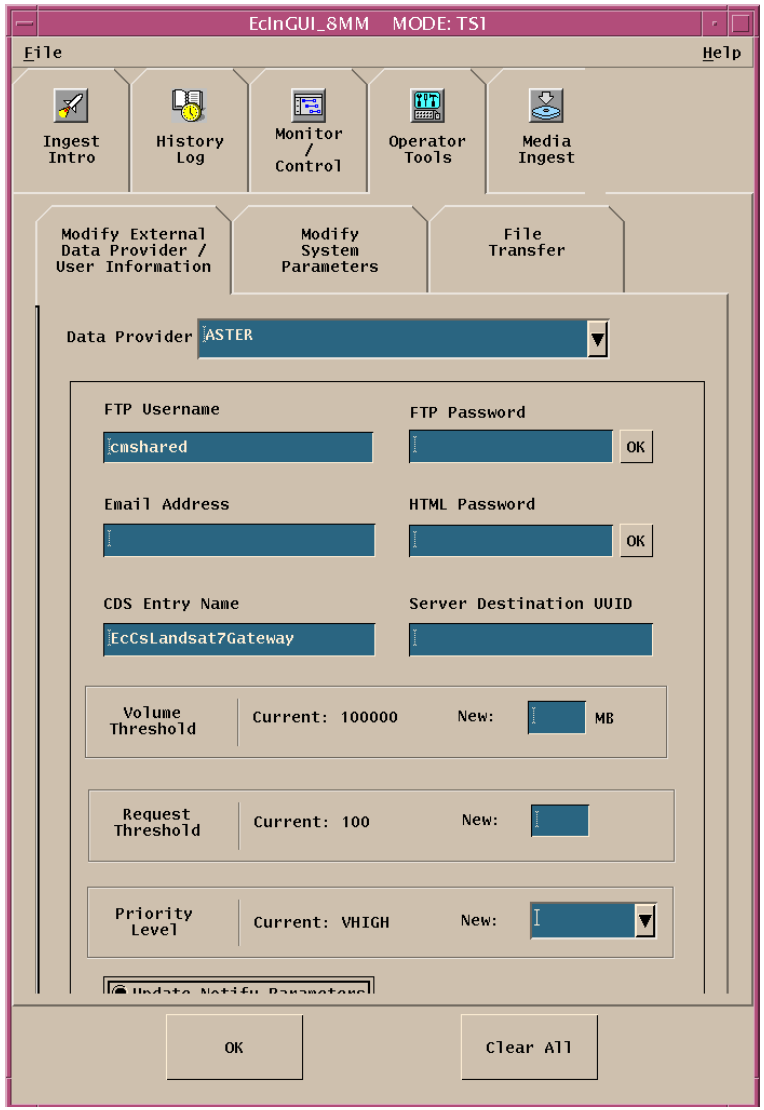
Ingest Tunable Parameters and File Transfers (Cont.)

- Data provider data/thresholds
  - FTP user name/password
  - E-mail address
  - HTML password [not used]
  - Cell Directory Service (CDS) entry name
  - Server destination Universal Unique Identifier (UUID)
  - Maximum data volume
  - Maximum number of concurrent ingest requests
  - Priority for ingest processing
  - "Notify" parameters; for example,...
    - » ftp directory
    - » ftp username/password

Ingest Tunable Parameters and File Transfers (Cont.)

- System thresholds
  - Maximum data volume to be ingested concurrently
  - Maximum number of concurrent ingest requests
  - Communication retry count
  - Communication retry interval
  - Monitor time
  - Screen update time

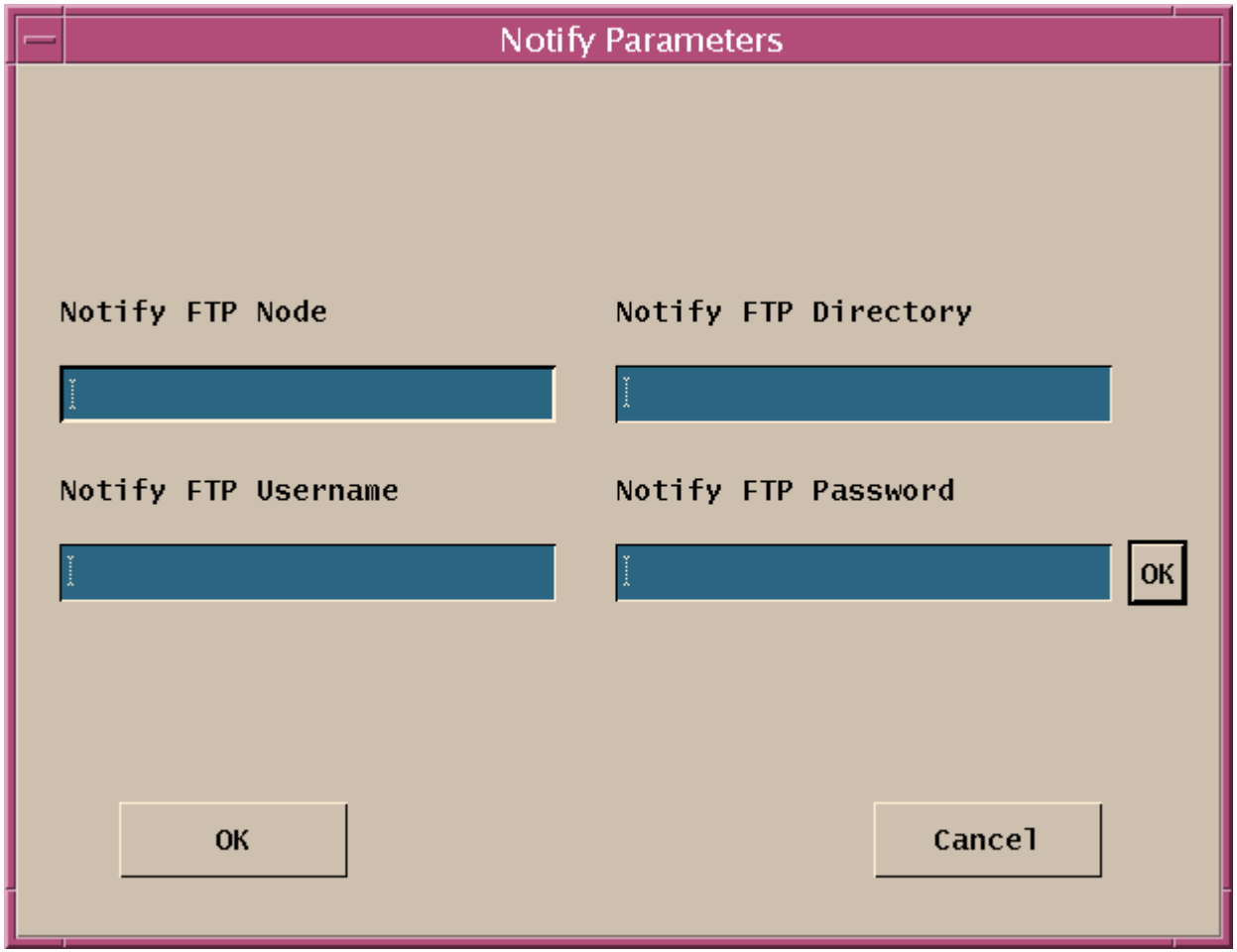
Modify Data Provider Parameters



Ingest Tunable Parameters and File Transfers (Cont.)

- **Procedure (Modifying External Data Provider Information)**
  - **Select the Ingest GUI Operator Tools: Modify External Data Provider/User Information tab**
  - **Select the data provider whose information is to be changed**
  - **Modify the data provider information as necessary**
  - **Save the changes to data provider information**

Notify Parameters Window



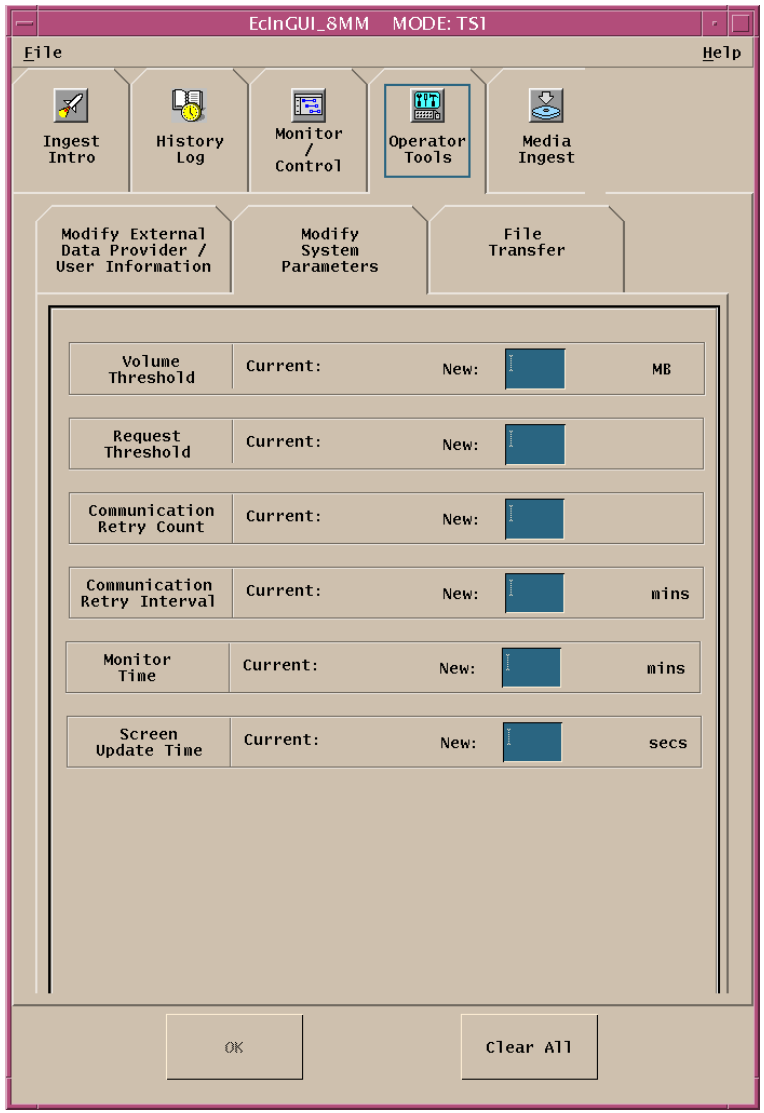
Ingest Tunable Parameters and File Transfers (Cont.)

- **Two system parameters affect communications between external data providers and ECS**
  - **Communication retry count**
    - » **The number of successive times the system tries to establish ingest communications with a data provider before registering a communications failure and moving on to the next ingest request**
  - **Communication retry interval**
    - » **The time between successive attempts to establish communication**

Ingest Tunable Parameters and File Transfers (Cont.)

- **Two system parameters may be used to set the behavior of the system according to operator preference**
  - **Monitor time**
    - » **The amount of time that information about a completed ingest transaction remains available on the Monitor/Control screen after its completion**
  - **Screen Update Time**
    - » **The amount of time between automatic data updates on the Monitor/Control screen**

Modify System Parameters



Ingest Tunable Parameters and File Transfers (Cont.)

- **Procedure (Modifying System Parameters on the Ingest GUI)**
  - **Select the Ingest GUI Operator Tools: Modify System Parameters tab**
  - **Modify Ingest operating parameters as necessary**
  - **Save the changes to Ingest operating parameters**

CAUTION

The thresholds are retrieved from the Ingest database when the Ingest Request Manager comes up

If the Granule Server thresholds are changed in the database while Ingest is running there will be a mismatch between the values in memory and the values in the database

This could cause an Ingest failure

Ingest Tunable Parameters and File Transfers (Cont.)

- **File Transfer tab**
  - **allows the Ingest Technician to transfer files**
  - **allows the Ingest Technician to build System Monitoring and Coordination Center (SMC) History Files**
    - » **SMCHeaderFile**
    - » **SMCdataFile**

Transfer Files

EcInGUI_8MM MODE: TS1

File Help

Ingest Intro | History Log | Monitor / Control | Operator Tools | Media Ingest

Modify External Data Provider / User Information | Modify System Parameters | File Transfer

Build SMC History Files | Generic File Transfer

Transfer Origin

Filter

/usr/ecs/TS1/CUSTOM/temp/INS/

Directories

:s/TS1/CUSTOM/temp/INS/.
:s/TS1/CUSTOM/temp/INS/..

Files

DataTypeLevelInfoFile27422
DataTypeLevelInfoFile2922
DataTypeLevelInfoFile6260
GraphicalViewInfoFile27422
HistRequestFile27422
RequestLevelInfoFile27422

Selection

/usr/ecs/TS1/CUSTOM/temp/INS/

OK Filter

Transfer Destination

OK

Ingest Tunable Parameters and File Transfers (Cont.)

- **Procedure (Transferring Files)**
  - Select the Ingest GUI Operator Tools: File Transfer tab
  - Select either Build SMC History Files or Generic File Transfer as appropriate
  - Select the file to be transferred
  - Enter the destination of the file to be transferred
  - Initiate and monitor the file transfer

Tuning System Configuration Parameters

- System parameters may be subject to control by Configuration Management (CM)
  - When making or requesting a change to system parameters, the CM process at the particular site must be followed (if applicable)
- Two databases where Ingest parameters can be set:
  - Configuration Registry database
  - Ingest database

Tuning System Configuration Parameters (Cont.)

- **Modifying System Parameters in the Configuration Registry Database**
  - Configuration Registry Server provides a single interface (via a Sybase server) for retrieving configuration attribute-value pairs for ECS servers from the Configuration Registry database
    - » When ECS servers are started they access the Configuration Registry database to obtain needed configuration parameters
  - Database Administrator has access to a Configuration Registry GUI for viewing and editing configuration data in the database
  - It is necessary to coordinate with the Database Administrator when changes to configuration parameters are needed
  - Changes to configuration-controlled parameters are subject to approval through the site CM process

Tuning System Configuration Parameters (Cont.)

- **Default and adjusted values assigned to system parameters vary from site to site**
  - **For guidance concerning the assignment of values to parameters included in the Configuration Registry refer to document 910-TDA-022, Custom Code Configuration Parameters for ECS**
    - » **Document is available at http://cmdm.east.hitc.com/baseline/ under "Technical Documents"**

Tuning System Configuration Parameters (Cont.)

- **Parameters whose values may be modified to enhance system functioning or performance**
  - **AppLogSize [parameter applies to all servers]**
    - » **Maximum size of the application log (ALOG) file for a particular application**
    - » **Recommended size varies considerably depending the nature of the application for which the file is being written**
  - **AppLogLevel [parameter applies to all servers]**
    - » **Level of detail provided in the ALOG file for a particular application**
    - » **Acceptable values are 0, 1, 2, or 3**
    - » **A setting of “0” provides the most data**

Tuning System Configuration Parameters (Cont.)

- **Tuning Parameters (Cont.)**
  - **DebugLevel [parameter applies to all servers]**
    - » **Level of detail provided in the debug log file for a particular application**
    - » **Normally acceptable values are 0, 1, 2, or 3**
    - » **A setting of "0" turns off logging; a setting of “3” provides a significant amount of data**
    - » **STMGT offers "enhanced" debugging based on bitmaps**
      - **Level 7 (the 4 bit) provides detailed database debugging**
      - **Level 15 (the 8 bit) frequently dumps the in-memory request queue (in the Request Manager)**
      - **Both Level 7 and Level 15 quickly create enormous log files.**

Tuning System Configuration Parameters (Cont.)

- **Tuning Parameters (Cont.)**
  - **INGEST_CONNECTION_POOL_SIZE [EcInPolling, EcInGUI, EcInReqMgr, and EcInGran parameter]**
    - » **Number of database connections - varies with the particular program connecting to the database**
    - » **Single-threaded programs (i.e., EcInGUI, EcInPolling) need one database connection only**
    - » **The number of database connections required for EcInGran depends on the maximum number of granules that can be processed at a time (for a maximum of five granules, two database connections are probably enough)**
    - » **For the EcInReqMgr there should be at least two database connections**

Tuning System Configuration Parameters (Cont.)

- Tuning Parameters (Cont.)
  - SAVEONEXIT [EcInReqMgr, EcInGran, and EcInGUI parameter]
    - » Set to "false" normally
    - » Set to “true” for debug purposes only
    - » When “true,” staging disks do not get cleaned up and the Staging Disk Server needs to be warm-started
    - » For Granule Server, when “true,” the local preprocessing disk does not get cleaned up
  - SDSRV_RETRY_INTERVAL [EcInGran parameter]
    - » Amount of time in seconds (e.g., 60) that Granule Server waits before retrying a remote procedure call (RPC) to Science Data Server (after retryable error)

Tuning System Configuration Parameters (Cont.)

- Tuning Parameters (Cont.)
  - SDSRV_RETRY_ATTEMPTS [EcInGran parameter]
    - » Number of times (e.g., 60) the Granule Server tries to make an RPC to Science Data Server when a retryable error is returned
    - » If set to one (1), then no retries are done
  - SocketLimit [EcInEmailGWServer, EcInReqMgr, EcInGran parameter]
    - » Number of connections (e.g., 200) to a server through the Hubble Space Telescope (HST) sockets middleware
    - » Too few: misses connections
    - » Too many: may adversely affect server host's memory

Tuning System Configuration Parameters (Cont.)

- Tuning Parameters (Cont.)
  - PollingTimerInterval [EcInPolling parameter]
    - » Amount of time in seconds (e.g., 120) between polling instances
    - » Specified individually for each data provider
    - » Value varies depending on rate at which each data provider sends data to EcInPolling
    - » Applies to both categories of polling (i.e., with delivery record and without delivery record)
  - CompareFileContentsFlag [EcInPolling parameter]
    - » Usually set to "no"
    - » Should be set to "yes" for any data provider that reuses the same file names for its PDRs (polling with delivery record) or for its data files (polling without delivery record) so that EcInPolling checks whether file contents have changed
    - » If set to "yes," EcInPolling processing takes longer

Tuning System Configuration Parameters (Cont.)

- Tuning Parameters (Cont.)
  - PollingDirectoryCount [EcInPolling parameter (polling without delivery record)]
    - » Number of directories to be polled by EcInPolling
    - » There must be a corresponding number of PollingDirectory, DataType, HostName, IngestFileType, and MaximumFileSize parameters
  - PollingDirectory or PollingDirectoryX [EcInPolling parameter]
    - » Path of the directory in which EcInPolling looks for new PDRs (polling with delivery record) or new data files (polling without delivery record)
    - » Used for setting the directory ID parameter in the PDRs that EcInPolling generates
    - » If the PollingDirectoryCount is "2," (polling without delivery record) there should be a PollingDirectory1 and a PollingDirectory2

Tuning System Configuration Parameters (Cont.)

- **Tuning Parameters (Cont.)**
  - **HostName or HostNameX [EcInPolling parameter (polling without delivery record)]**
    - » **Host (e.g., x0icg01) where the associated polling directory resides**
    - » **Used for setting the node name parameter in the PDRs that EcInPolling generates**
    - » **If the PollingDirectoryCount is "2," there should be a HostName1 and a HostName2 (although both may have exactly the same value)**

Tuning System Configuration Parameters (Cont.)

- **Tuning Parameters (Cont.)**
  - **DataTypeX [EcInPolling parameter (polling without delivery record)]**
    - » **Identifies the data type (e.g., AM1ATTF) associated with the corresponding polling directory**
    - » **Used for setting the data type parameter in PDRs**
    - » **If the PollingDirectoryCount is "2," (polling without delivery record) there should be a DataType1 and a DataType2 and they should have different values**
    - » **Must be a valid data type or ingest will fail**

Tuning System Configuration Parameters (Cont.)

- Tuning Parameters (Cont.)
  - IngestFileTypeX [EcInPolling parameter (polling without delivery record)]
    - » Identifies the file type (e.g., SCIENCE or DATA) associated with the corresponding polling directory
    - » Used for setting the file type parameter in the PDRs that EcInPolling generates
    - » If the PollingDirectoryCount is "2," (polling without delivery record) there should be an IngestFileType1 and an IngestFileType2 (although both may have exactly the same value)
    - » Must be a valid file type (in the Ingest database) or ingest will fail

Tuning System Configuration Parameters (Cont.)

- **Tuning Parameters (Cont.)**
  - **MaximumFileSizeX [EcInPolling parameter (polling without delivery record)]**
    - » **Specifies the maximum file size in bytes (e.g., 1000000) allowed in the corresponding polling directory**
    - » **Used for setting the file size parameter in the PDRs that EcInPolling generates**
    - » **If the PollingDirectoryCount is "2," there should be a MaximumFileSize1 and a MaximumFileSize2 (both may have exactly the same value)**
    - » **If the file size is too small, the staging disk created for ftping the files will not be big enough**

Tuning System Configuration Parameters (Cont.)

- When the value assigned to a parameter has been changed and saved in the Configuration Registry, the modified value does not take effect until the affected server has been restarted
- Example
  - Debug level for the Ingest Request Manager log has been changed from “2” to “3” in the Configuration Registry
  - Modification does not affect the recording of data in the log until after a warm restart of the Ingest Request Manager (at which time the server would read the parameters in the Configuration Registry)

Modifying System Parameters in the Ingest Database

- Some of the more important tunable parameters in the Ingest Database are included in the lesson
  - Information concerning additional tunable parameters is in the "Tunable Parameters in Databases - Descriptions" section of 910-TDA-022, Custom Code Configuration Parameters for ECS
  - The document is available at http://cmdm.east.hitc.com/baseline/ under "Technical Documents"
- When requesting a change to system parameters, the CM process at the particular site must be followed (if applicable)

Modifying System Parameters in the Ingest Database (Cont.)

- **Limits on the Number of Queued Requests and Ingest Volume**
  - **There is no direct way to set the number of queued requests**
  - **Limits on Ingest volume are managed through the following database parameters:**
    - » **TotalGranuleThreshold in the InGranuleServerInfo table**
    - » **VolumeThreshold in the InGranuleServerInfo table**
    - » **MaximumTotalRequests in the InSystemParameters table**
    - » **MaximumTotalVolume in the InSystemParameters table**
  - **Request Manager receives requests, breaks them into granules, and queues all the granules**

Modifying System Parameters in the Ingest Database (Cont.)

- **Limits on the Number of Queued Requests and Ingest Volume (Cont.)**
  - The granule queue is maintained in the Ingest database (InGranuleQueue table)
    - » The queue state of each granule and the Granule Server processing it can be determined should the Request Manager have to be restarted in response to a failure
  - If the appropriate Granule Server is not processing the maximum number of granules that it can process at a time (TotalGranuleThreshold), one or more granule(s) is (are) removed from the queue and sent to the Granule Server
  - Same action occurs if the appropriate Granule Server is not processing the maximum data volume that the Granule Server can process at a time (VolumeThreshold)

Modifying System Parameters in the Ingest Database (Cont.)

- **Limits on the Number of Queued Requests and Ingest Volume (Cont.)**
  - The Request Manager uses the VolumeThreshold and TotalGranuleThreshold parameters to control when it sends granules to each Granule Server
  - Entries in the InGranuleServerInfo database table must be set manually via interactive structured query language (isql) commands
  - If the TotalGranuleThreshold parameter is changed, the Request Manager and the appropriate Granule Server need to be restarted in order for them to see the change
  - If the VolumeThreshold parameter is changed, the Request Manager needs to be restarted
  - It is better to avoid changing either parameter while the Granule Server is in the middle of processing granules

Modifying System Parameters in the Ingest Database (Cont.)

- **Limits on the Number of Queued Requests and Ingest Volume (Cont.)**
  - There is a maximum number of requests and maximum volume that can be processed by Ingest at one time (in contrast to the Granule Server limits mentioned previously)
  - The corresponding parameters are specified in two columns in the InSystemParameters database table:
    - » MaximumTotalRequests
    - » MaximumTotalVolume
  - When a request from one of the clients (e.g., GUI or Polling) would cause one of the parameters to exceed its maximum value, the request fails and is not sent to Request Manager
  - Either parameter can be modified using the Ingest GUI Operator Tools: Modify System Parameters tab

Modifying System Parameters in the Ingest Database (Cont.)

- **Limits on the Number of Requests and Data Volume from a Data Provider**
  - **Maximum number of requests for a data provider**
    - » **MaximumRequests column in the InExternalDataProviderInfo database table**
    - » **Use the Ingest GUI to modify**
  - **Maximum data volume for a data provider**
    - » **VolumeThreshold column in the InExternalDataProviderInfo database table**
    - » **Use the Ingest GUI to modify**

Modifying System Parameters in the Ingest Database (Cont.)

- **Other Key Parameters for Ingest**
  - **MonitorTimeForCompletedRequest in the InSystemParameters database table**
    - » **Specifies the number of minutes after the request has been completed that a request remains in the database tables that get displayed on the Ingest Monitor/Control GUI window**
    - » **Use the Ingest GUI to modify**
  - **ScreenUpdateInterval parameter in the InSystemParameters table**
    - » **Specifies the number of seconds after which the GUI refreshes**
    - » **Use the Ingest GUI to modify**

Modifying System Parameters in the Ingest Database (Cont.)

- Other Key Parameters for Ingest (Cont.)
  - CommunicationRetryCount in the InSystemParameters table
    - » Specifies a number of times that a user retries a communication
    - » Use the Ingest GUI to modify
  - CommunicationRetryInterval in the InSystemParameters table
    - » Time interval in seconds between user communication retries
    - » Use the Ingest GUI to modify

Modifying System Parameters in the Ingest Database (Cont.)

- Other Key Parameters for Ingest (Cont.)
  - Number of granule servers at a DAAC
    - » Multiple granule servers can be configured
    - » Each granule server can process multiple Earth Science Data Types (ESDTs), but each ESDT can be assigned to one granule server only
    - » In order for a particular ESDT to be processed by a particular granule server the GranuleServerURKey entry for the data type in the InDataTypeTemplate table must be set to the integer representing the appropriate granule server
    - » GranuleServerURKey is the granule server ID that is mapped to a specific granule server name (GranuleServerUR)
    - » GranuleServerURKey is the primary key in the InValGranuleServerUR table

Modifying System Parameters in the Ingest Database (Cont.)

- Other Key Parameters for Ingest (Cont.)
  - Number of granule servers at a DAAC (Cont.)
    - » The GranuleServerURKey column contains the possible values that can be used in the InGranuleServerInfo and InDataTypeTemplate tables
    - » So each granule server requires an individual row in the InValGranuleServerUR table with values for GranuleServerURKey (e.g., 1, 2, 3) and GranuleServerUR (e.g., EcInGran, EcInGran0, EcInGran1)
    - » In addition each granule server requires an individual row in the InGranuleServerInfo table with values for GranuleServerURKey, TotalGranuleThreshold, and VolumeThreshold
    - » To prevent changing the mapping between GranuleServerURKey and GranuleServerUR values no changes are allowed to the values in the GranuleServerURKey column in tables InValGranuleServerUR or InGranuleServerInfo

Modifying System Parameters in the Ingest Database (Cont.)

- Other Key Parameters for Ingest (Cont.)
  - Number of granule servers at a DAAC (Cont.)
    - » However, entries can be added to both tables (InValGranuleServerUR and InGranuleServerInfo)
    - » If a new GranuleServerURKey entry is added to the InValGranuleServerUR table, in order for things to work correctly, a new entry for the GranuleServerURKey needs to be added to the InGranuleServerInfo table and a new granule server needs to be configured in order for the new table entry to be used
    - » Also, if additions are made to the InGranuleServerInfo table, the Request Manager needs to be restarted in order for it to see the changes
    - » Entries in the database tables InGranuleServerInfo, InValGranuleServerUR, or InDataTypeTemplate must be set manually via isql commands

Modifying System Parameters in the Ingest Database (Cont.)

- **Procedure**
  - Access a terminal window logged in to the Ingest Server host
  - Log in to the appropriate Ingest database using isql commands
  - Check the current contents of the relevant column/table
  - Update the relevant column/table with the new value(s)
  - Check the current contents of the relevant column/table
  - Exit from isql

Troubleshooting Ingest Problems

- **Troubleshooting:**

  process of identifying the source of problems on the basis of observed trouble symptoms

Troubleshooting Ingest Problems (Cont.)

- Problems with ingest can usually be traced to…
  - some part of the Ingest Subsystem
  - problems in other ECS subsystems, including (but not necessarily limited to):
    - » Data Server Subsystem (DSS)
    - » Communications Subsystem (CSS)
    - » System Management Subsystem (MSS)
  - mistakes in the delivery records furnished by external data providers
  - errors in transmission of the data from external data providers

Troubleshooting Ingest Problems (Cont.)

- **Fault Recovery**
  - Each request that crosses a client/server boundary is assigned a system-unique identifier referred to as an RPC ID
  - 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

Troubleshooting Ingest Problems (Cont.)

- **Fault Recovery (Cont.)**
  - The concept of reproducible RPC IDs is central to the ECS 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 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
    - » As the request is serviced, check-pointing state information may be updated in the persistent store, up to and including the request's completion status
    - » This allows the servers to resume servicing from the last check-pointed state, particularly upon resubmission from a client

Troubleshooting Ingest Problems (Cont.)

- **Fault Recovery (Cont.)**
  - **Ingest components check-point the following types of information:**
    - » **EcInGran - Granule and granule state information**
    - » **EcInReqMgr - Request state information**
    - » **EcInPolling - Request information**
    - » **EcInEmailGWServer - None**
    - » **EcInGUI - Media Ingest request information**

Troubleshooting Ingest Problems (Cont.)

- **Fault Recovery: Fault Handling**
  - **Failure events are classified according to the following three severity levels:**
    - » **Fatal error**
    - » **Retry error**
    - » **Warning**

Troubleshooting Ingest Problems (Cont.)

- **Fault Recovery: Fault Handling (Cont.)**
  - Fatal error is 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
  - Retry error is a 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
    - » For example, a tape left in a tape drive might have to be removed manually
  - Warning is provided when operations can proceed without interruption, but an unexpected circumstance was detected
    - » For example, if a client requests removal of a file but the file does not exist

Troubleshooting Ingest Problems (Cont.)

- **Fault Recovery: Fault Handling (Cont.)**
  - **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**

Troubleshooting Ingest Problems (Cont.)

- **Fault Recovery: Fault Handling (Cont.)**
  - **ECS 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 Ingest client processes are shown in the table that follows**

Troubleshooting Ingest Problems (Cont.)

Ingest Fault Handling Policies

Client Process	Fault Handling Policy
EcInGran	**Retry errors:** An error in sending a media ingest request to the Ingest Request Manager is reported to the operator and the operator can retry. Other retry errors result in the request failing. **Fatal errors:** The granule is failed. Granule failures are displayed on the Ingest GUI.
EcInReqMgr	**Retry errors:** Errors connecting to EcInGran are retried forever. Retry errors involving staging disks are retried a configurable number of times, then the request is failed. **Fatal errors:** Errors are failed immediately.
EcInGUI	**Retry errors:** Any error results in the request failing. **Fatal errors:** Any error results in the request failing.
EcInPolling	**Retry errors:** Errors are retried forever, with a delay between retries. **Fatal errors:** Errors are failed immediately, and are displayed on the Ingest GUI.
EcInEmailGWServer	**Retry errors:** N/A **Fatal errors:** E-mail that cannot be processed is moved to a failed directory, but no operator notification is provided.

Troubleshooting Ingest Problems (Cont.)

- **Fault Recovery: Client Crash and Restart**
  - When a client of an Ingest server crashes, the server (i.e., EcInGran or EcInReqMgr) continues to service the requests that were in process at the time of the client's crash
  - When a client restarts in the ECS 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
    - » EcInGUI is an exception: when EcInGUI restarts, it always informs EcDsStRequestManagerServer that it has performed a warm restart

Troubleshooting Ingest Problems (Cont.)

- **Fault Recovery: Client Crash and Restart (Cont.)**
  - When an Ingest client (e.g., EcInGran, EcInReqMgr, or EcInGUI) sends restart notification to the EcDsStRequestManagerServer, the latter calls a stored procedure to clean up the old request and staging disk 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

Troubleshooting Ingest Problems (Cont.)

- **Fault Recovery: Client Crash and Restart (Cont.)**
  - **Default server behavior in response to "warm" startup notification from a client:**
    - » **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**

Troubleshooting Ingest Problems (Cont.)

- **Fault Recovery: Client Crash and Restart (Cont.)**
  - **Default server behavior in response to "cold" startup notification from a client:**
    - » **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**

Troubleshooting Ingest Problems (Cont.)

- **Fault Recovery: 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

Troubleshooting Ingest Problems (Cont.)

- **Fault Recovery: Server Crash and Restart (Cont.)**
  - 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

Troubleshooting Ingest Problems (Cont.)

- **Fault Recovery: Server Crash and Restart (Cont.)**
  - **EcInGran-specific activities upon start/restart:**
    - » **Warm Restart: The EcInGran server automatically restarts submitted requests from the beginning. If a file has been transferred (e.g., via ftp), it does not redo the transfer of that file**
    - » **Cold Start: All granule requests are cancelled. Existing request queues are cleared**
    - » **Cold Restart: All granule requests are cancelled. Existing request queues are retained**

Troubleshooting Ingest Problems (Cont.)

- **Fault Recovery: Server Crash and Restart (Cont.)**
  - **EcInReqMgr-specific activities upon start/restart:**
    - » **Warm Restart: EcInReqMgr resynchronizes requests in progress with EcInGran, and resumes processing from the last check-pointed state**
    - » **Cold Start: All active requests are moved to the summary tables. Existing request queues are cleared**
    - » **Cold Restart: Each granule is resubmitted to the EcInGran, where it is failed. EcInReqMgr then resubmits the request to EcInGran, where it is processed as a new request. Existing request queues are retained**

Troubleshooting Ingest Problems (Cont.)

- **Fault Recovery: Server Crash and Restart (Cont.)**
  - **EcInPolling-specific activities upon start/restart:**
    - » **Warm Restart: EcInPolling resubmits requests that were in progress at the time of fault. Continues polling for remaining requests in the polling directory**
    - » **Cold Start or Cold Restart: EcInPolling cleans up files and terminates any requests that had not yet been sent to EcInReqMgr. Requests remaining in the polling directory are sent as new requests**

Troubleshooting Ingest Problems (Cont.)

- **Troubleshooting table**
  - **describes actions to be taken in response to some common ingest problems**
  - **if the problem cannot be identified and fixed without help within a reasonable period of time, call the help desk and submit a trouble ticket in accordance with site Problem Management policy**

Troubleshooting Ingest Problems (Cont.)

Symptom	Response
Unable to log in to any host (e.g., Operations Workstation, g0acs02).	Check with the Operations Controller/System Administrator to ensure that the host is "up."
GUI not displayed when the start-up script has been properly invoked.	Ensure that the DISPLAY variable was set properly. [For detailed instructions refer to the procedure for **Launching the ECS Ingest GUI** (previous section of this lesson).]
Error message associated with the Ingest GUI.	Refer to Table 3, Ingest Operator GUI User Messages (adapted from the corresponding table in 609-CD-610-003, *Release 6B Operations Tools Manual for the ECS Project*).
Message received indicating a data ingest failure.	1. Ensure (e.g., using EcCsIdPingServers) that it is possible to connect to the necessary hosts and servers (listed in Table 5). [For detailed instructions refer to the procedure for **Checking Connections to Hosts** (subsequent section of this lesson).] 2. If it is not possible to connect to any needed host(s)/server(s), notify the Operations Controller/System Administrator to check the hosts/servers and bring them back up if necessary. 3. If hosts/servers are all "up," refer to the procedure for **Recovering from a Data Ingest Failure** (subsequent section of this lesson).
Other problems.	Check the log files (e.g., EcInReqMgr.ALOG, EcInPolling.ALOG, EcInGran.ALOG, EcInGUI.ALOG) in the /usr/ecs/*MODE*/CUSTOM/logs directory of the relevant host(s) for error messages. [For detailed instructions refer to the procedure for **Checking Log Files** (subsequent section of this lesson).]

Hosts, Servers, Clients and Other Software Relevant to Ingest

HOST	SERVER/CLIENT/OTHER SOFTWARE
Ingest Server (e.g., x0icg01)	Name Server (EcCsIdNameServer) Registry Server (EcCsRegistry) FTP Server (EcDsStFtpServer) Staging Disk Server (EcDsStStagingDiskServer) Ingest Granule Server (EcInGran) Automated Polling Ingest Client Interface (EcInPolling) Ingest Request Manager (EcInReqMgr)
Operations Workstation (e.g., x0acs02)	ECS Ingest GUI (EcInGUI) Storage Management Control GUI (EcDsStmgtGui)
Access/Process Coordinators (APC) Server (e.g., x0acg01)	Archive Server (EcDsStArchiveServer) Cache Manager Server (EcDsStCacheManagerServer) (including Pull Area Manager) FTP Server (EcDsStFtpServer) Staging Disk Server (EcDsStStagingDiskServer) Automated Polling Ingest Client Interface (EcInPolling)
FSMS Server (e.g., x0drg01)	HDF EOS Server (EcDsHdfEosServer) Archive Server (EcDsStArchiveServer) Cache Manager Server (EcDsStCacheManagerServer) FTP Server (EcDsStFtpServer) Staging Disk Server (EcDsStStagingDiskServer)

Hosts, Servers, Clients and Other Software Relevant to Ingest (Cont.)

HOST	SERVER/CLIENT/OTHER SOFTWARE
Sun Consolidation External Server (e.g., x0ins01)	Data Dictionary (EcDmDictServer) Ingest E-Mail Parser (EcInEmailGWServer)
Sun Consolidation Internal Server (e.g., x0acs11)	Science Data Server (EcDsScienceDataServer) Subscription Server (EcSbSubServer) Distribution Server (EcDsDistributionServer) 8mm Server (EcDsSt8MMServer) DTF-2 Server (EcDsStDTFServer) Staging Disk Server (EcDsStStagingDiskServer) Storage Management Request Manager (EcDsStRequestManagerServer) INGEST Media Tape Reader GUI (EcInTapeReaderGUI)

Troubleshooting Ingest Problems (Cont.)

- **Procedure (Checking Connections to Hosts):**
  - **Access a terminal window logged in to the Operations Workstation host**
  - **Change directory to the utilities directory (/usr/ecs/MODE/CUSTOM/utilities)**
  - **At the command line prompt enter EcCsIdPingServers *MODE***
  - **Observe the results displayed on the screen to determine whether connections can be made with the necessary hosts and servers**
  - **Ping the servers again (EcCsIdPingServers *MODE*)**
  - **If it is not possible to connect to any needed host(s)/server(s), notify the Operations Controller/System Administrator to check the hosts/servers and bring them back up if necessary**

Troubleshooting Ingest Problems (Cont.)

- **Checking Log Files**
  - **Log files can provide indications of the following types of problems:**
    - » **Communication problems**
    - » **Database problems**
    - » **Lack of disk space**

Troubleshooting Ingest Problems (Cont.)

- **Procedure (Checking Log Files)**
  - Access a terminal window logged in to the appropriate host
  - Change directory to the directory containing the ingest log files
    - » /usr/ecs/MODE/CUSTOM/logs
  - Review log file to identify problems
    - » EcInGUI.ALOG
    - » EcInTapeReaderGUI.8mm.log
    - » EcInTapeReaderGUI.DTF1.log
    - » EcInReqMgr.ALOG
    - » EcInPolling.ALOG
    - » EcInGran.ALOG
    - » EcInEmailGWServer.ALOG
  - Respond to problems

Troubleshooting Ingest Problems (Cont.)

- Recovering from a data ingest failure
  - Operator intervention required when there is an ingest fault, or error (e.g., invalid PDR)
  - System responses to Ingest fault (error)
    - » processing of the ingest request stops
    - » message is sent to the Ingest Technician and the data provider with a brief description of the problem
  - Ingest Technician may use several sources for troubleshooting information
    - » Ingest GUI Monitor/Control screen
    - » Ingest History Log
    - » Ingest log files

Troubleshooting Ingest Problems (Cont.)

- **Procedure (Recovering from a Data Ingest Failure)**
  - **Identify the faulty ingest request**
  - **Review the information concerning the faulty ingest request**
    - » **Ingest notification files**
    - » **Log files**
  - **Perform the appropriate recovery procedure depending on the nature of the problem**

Troubleshooting Ingest Problems (Cont.)

- **Checking Ingest Notification Files (Polling with Delivery Record)**
  - **PAN (short or long): transfer errors, delivery record information discrepancies, and other file problems (if any)**
  - **PDRD (short or long): invalid delivery record**

Troubleshooting Ingest Problems (Cont.)

- **Checking Ingest Notification Files (Polling with Delivery Record): Procedure**
  - **Access a terminal window logged in to the appropriate host**
  - **Change directory to the directory containing the ingest log files**
    - » **e.g., /usr/ecs/OPS/CUSTOM/icl//x0icg01/data/remote/EDOS/Response**
  - **Review notification file to identify problems**
    - » **PAN**
    - » **PDRD**

Troubleshooting Ingest Problems (Cont.)

- **Procedure (Recovering from a Faulty PDR)**
  - Contact the data provider
    - » Report the ingest failure
    - » Discuss what has been discovered from reviewing the failure event data
    - » Determine whether the data provider will re-initiate the data ingest request with a new PDR
  - If there is an ECS Ingest process abort during file transfer, go to the procedure for Recovering from File Transfer (ftp) Error
  - If EDOS is the data provider and the File Transfer Disposition in the PAN indicates that an error occurred, send a Problem Report to EDOS to report the problem
  - If the data ingest request is to be re-initiated, monitor the subsequent ingest

Troubleshooting Ingest Problems (Cont.)

- **Regenerating Failed PDRs**
  - The Regenerate Failed PDR Tool can be used whenever a PDR fails and results in a "long PAN" message file
    - » Long PAN means that the request had more than one granule and not all granules had the same error
  - Purpose of the tool is to provide a means to generate a PDR for each failed granule in a PDR and copy the generated PDRs to an Ingest polling directory, where Ingest polling would detected them and initiate ingest of the relevant granule(s)
    - » Operations staff would not have to either manually edit the original PDR file or submit all failed granules to Ingest polling (which would create duplicate granules in the archive)

Troubleshooting Ingest Problems (Cont.)

- **Regenerating Failed PDRs: Procedure**
  - Access a terminal window logged in to the appropriate host
  - Change directory to the directory containing the Ingest utility scripts
  - Start the EcInRegenFailedPDRStart script
  - Respond to prompts
  - Perform related procedures as necessary
    - » Remove (Delete) Generated PDRs
    - » Check/Edit a PDR
    - » Check PAN Contents
    - » Check for Memory Problems
    - » Check the Polling Directory
    - » Check PAN Accessibility

Troubleshooting Ingest Problems (Cont.)

- Other ingest failures for operator intervention
  - Recovering from Exceeding the Volume Threshold
  - Recovering from Exceeding the Maximum Number of Concurrent Requests
  - Recovering from Insufficient Disk Space
  - Recovering from Exceeding the Expiration Date/Time Period
  - Recovering from File Transfer (ftp) Error
  - Recovering from Processing Errors
    - » Missing Required Metadata
    - » Unknown Data Type
    - » Template Out of Synchronization (Sync)
    - » Unavailable File Type
    - » Metadata Validation Error
    - » Missing Optional Data Files

Troubleshooting Ingest Problems (Cont.)

- Other ingest failures for operator intervention (Cont.)
  - Recovering from Failure to Store Data
  - Checking the Request Manager Server Debug Log

Troubleshooting Ingest Problems (Cont.)

- **Recovering from Failure to Store Data**
  - Successful ingest of data depends on Storage Management (STMGT) inserting the product into the archive and Science Data Server (SDSRV) inserting the associated metadata into the inventory
  - Staging disks and cache managers for the Archive server and the FTP server are also involved in the process
  - Troubleshooting failures to store data (as well as other failures) often requires the review of server or application log files

Troubleshooting Ingest Problems (Cont.)

- **Procedure (Recovering from Failure to Store Data)**
  - Check the Storage Management Request Manager Server debug log file for error messages concerning the failure to store data
  - If necessary, check the Science Data Server debug log file for error messages
  - If necessary, check the Archive Server debug log file for error messages
  - Check other server debug log(s) for error messages concerning the failure to store data
    - » EcDsStStagingDiskServerDebug.log
    - » EcDsStCacheManagerServerDebug.log
    - » EcDsStRequestManagerServerDebug.log
    - » EcSbSubServerDebug.log