

**Interface Control Document  
between the ESDIS Metrics System  
(EMS) and the Data Providers**

**Revision A**

**October 2008**

**Expires: October 2013**



National Aeronautics and  
Space Administration

Goddard Space Flight Center  
Greenbelt, Maryland

This page intentionally left blank.

**Interface Control Document between  
the ESDIS Metrics System (EMS)  
and the Data Providers**

**Reviewed by:**

---

Jeanne Behnke  
ESDIS Science Operations Office Manager  
GSFC - Code 423

---

Date

**Approved by:**

---

Hampapuram K. Ramapriyan  
ESDIS Assistant Project Manager  
GSFC - Code 423

---

Date

---

Dawn Lowe  
ESDIS Project Manager  
GSFC - Code 423

---

Date

**Goddard Space Flight Center  
Greenbelt, Maryland**

This page intentionally left blank.

## Preface

---

This document is issued and maintained by the Earth Science Data and Information System (ESDIS) Project Code 423. Proposed changes shall be submitted to the Configuration Management Office (CMO) or the ESDIS Configuration Manager, along with supportive material justifying the proposed change. Changes to this document will be made by document change notice (DCN) or by complete revision.

This document is under ESDIS configuration control. Any questions should be addressed to:

ESDIS Project Configuration Management Office  
Code 423  
Goddard Space Flight Center NASA  
Greenbelt, MD 20771

This page intentionally left blank.

## Abstract

---

The Earth Science Data and Information System (ESDIS) Metrics System (EMS) is designed to support the ESDIS project management by collecting and organizing various metrics from the Earth Observing System (EOS) Distributed Active Archive Centers (DAACs) and other Data Providers. The EMS metrics are reported to NASA HQ to evaluate performance and effectiveness of the EOSDIS. The metrics data and the analysis reports generated from EMS provide NASA managers the information needed to determine how to best apply resources to support the science community.

The EMS collects and analyzes data on the usage of products and services delivered via the Internet, stored in databases or captured in custom files. Specifically, the EMS allows the project and other users access to metrics from the Data Providers including such information as number of users, type and amount of data archived, type and amount of data distributed, and other related information. The EMS is located at the Goddard Space Flight Center (GSFC) and is operated by the ESDIS Science Operations Office (SOO).

The EMS is a combination of custom code and a commercial Web analytics package. The custom code portion, ESDIS Data Gathering and Reporting System Plus (EDGRS+) is a subsystem of the EMS that handles retrieval, preprocessing and error reporting of data sent to the EMS. Additionally this component tracks ingest, archival and user profile information collected from Data Providers. The EDGRS+ component should not be confused with the previous EDGRS system. While most of the functionality of EDGRS has migrated to EMS/EDGRS+ the backend hardware and software have been completely redesigned to improve availability, address error handling, expand reporting capabilities and reduce complexity.

Data Providers now have the ability to send the EMS standard server logs in native formats as well as custom log formats. This relieves Data Providers of the responsibility of creating and maintaining custom parsing scripts and allows for easy tracking of offline orders or other activities. Furthermore, for EOSDIS Core System (ECS) Data Providers, the EMS metric scripts will be integrated into the ECS system and maintained by ECS, thus alleviating the need for Data Providers to create and maintain scripts to provide the EMS with these data. The Commercial Off-The-Shelf (COTS) product, NetInsight®, provides the web analytic capabilities.

This Interface Control Document (ICD) between the EMS and the Data Providers defines the software interfaces between an EMS Data Provider and the EMS. Specifically, this ICD defines the data and information transmitted between a Data Provider and the EMS for metrics collection and reporting.

**Keywords:** metrics, page tags, flat files, log files, COSSL

This page intentionally left blank.

## Change Record Page

---

ISSUE	RELEASE DATE	PAGES AFFECTED	DESCRIPTION
Baseline	08/28/2006	All	CCR 423-47-01-001

ISSUE	EFFECTIVE DATE	PAGES AFFECTED	DESCRIPTION	CCR APPROVED DATE
Revision A	10/14/2008	All	CCR 423-47-01-002	09/17/2008

This page intentionally left blank.

## List of Affected Pages

Page No.	Revision	Page No.	Revision	Page No.	Revision	Page No.	Revision
Title	Revision A	4-3	Revision A	A-4	Revision A		
i	Revision A	4-4	Revision A	AB-1	Revision A		
ii	Revision A	4-5	Revision A	AB-2	Revision A		
iii	Revision A	4-6	Revision A	AB-3	Revision A		
iv	Revision A	4-7	Revision A	AB-4	Revision A		
v	Revision A	4-8	Revision A				
vi	Revision A	4-9	Revision A				
vii	Revision A	4-10	Revision A				
viii	Revision A	4-11	Revision A				
ix	Revision A	4-12	Revision A				
x	Revision A	4-13	Revision A				
xi	Revision A	4-14	Revision A				
xii	Revision A	4-15	Revision A				
xiii	Revision A	4-16	Revision A				
xiv	Revision A	4-17	Revision A				
1-1	Revision A	4-18	Revision A				
1-2	Revision A	4-19	Revision A				
2-1	Revision A	4-20	Revision A				
2-2	Revision A	4-21	Revision A				
3-1	Revision A	4-22	Revision A				
3-2	Revision A	4-23	Revision A				
3-3	Revision A	4-24	Revision A				
3-4	Revision A	4-25	Revision A				
3-5	Revision A	4-26	Revision A				
3-6	Revision A	4-27	Revision A				
3-7	Revision A	4-28	Revision A				
3-8	Revision A	5-1	Revision A				
3-9	Revision A	5-2	Revision A				
3-10	Revision A	A-1	Revision A				
4-1	Revision A	A-2	Revision A				
4-2	Revision A	A-3	Revision A				

This page intentionally left blank.

# Table of Contents

---

## Preface

## Abstract

## 1. Introduction

1.1 Identification .....	1-1
1.2 Scope .....	1-1
1.3 Purpose.....	1-2
1.4 Status and Schedule.....	1-2
1.5 Organization.....	1-2

## 2. Related Documentation

2.1 Parent Documents.....	2-1
2.2 Applicable Documents .....	2-1
2.3 Information Documents.....	2-1

## 3. Interface Overview

3.1 EMS Overview .....	3-1
3.2 EMS-Data Providers Interface Context.....	3-2
3.3 EMS User Interface Overview.....	3-6
3.4 Network Protocols and Topology.....	3-6
3.4.1 Transport Layer Protocol .....	3-6
3.4.2 Network Layer Protocols .....	3-6
3.4.2.1 Internet Protocol (IP).....	3-6
3.4.2.2 Routing .....	3-7
3.4.2.3 Physical/Datalink Protocols .....	3-7
3.4.2.4 Network Topology .....	3-7
3.4.3 Communications Protocols.....	3-7
3.4.3.1 Simple Mail Transport Protocol.....	3-7
3.4.3.2 Secure Shell.....	3-7
3.4.3.3 Rsync .....	3-7
3.5 Security .....	3-8
3.5.1 Data Transfer Accounts .....	3-8
3.5.2 Data Access Accounts .....	3-8

3.5.3 Privacy Information and Policies.....	3-8
---	-----

## **4. EMS Data Type Descriptions**

4.1 EMS Data Types .....	4-1
4.2 EMS Data Files .....	4-1
4.2.1 Data File Manifest File .....	4-2
4.2.1.1 Data File Manifest File Format.....	4-2
4.2.1.2 Data File Manifest File Naming Convention.....	4-3
4.2.1.3 Data File Manifest File Updates .....	4-4
4.2.2 Data Files Interface Mechanism.....	4-4
4.2.3 Commercial or Open Source Server Logs.....	4-5
4.2.4 Custom Logs (Media Orders, Offline Activities, Non-standard Information).....	4-5
4.2.4.1 Special Services Captured in Logs .....	4-5
4.2.5 Log Files Naming Convention .....	4-6
4.2.6 Log Files Updates.....	4-7
4.2.7 EMS Flat Files.....	4-7
4.2.7.1 Flat Files Naming Convention.....	4-8
4.2.7.2 Flat Files Updates .....	4-9
4.2.7.3 Flat Files Formats .....	4-9
4.2.7.3.1 Data Ingest Flat Files .....	4-9
4.2.7.3.2 Data Archive Flat Files .....	4-10
4.2.7.3.3 User Profile Flat Files .....	4-12
4.2.7.3.3.1 ESDIS User Type Definitions.....	4-13
4.2.7.3.4 Product Attribute Flat Files.....	4-14
4.2.7.3.4.1 Product Attribute Search Flat Files.....	4-15
4.2.7.3.4.2 Product Attribute Metadata Flat Files.....	4-15
4.3 EMS Page Tags.....	4-17
4.3.1 Page Tags Interface Mechanism.....	4-17
4.3.2 Page Tag Manifest File Format .....	4-18
4.3.3 Page Tag Manifest File Naming Convention .....	4-21
4.3.4 Page Tags Updates .....	4-21
4.4 EMS Error Recovery and Notifications.....	4-22
4.4.1 EMS Unavailable or Network Outage.....	4-22
4.4.2 Data Provider Unavailable .....	4-22
4.4.3 Error Email Notifications .....	4-23

## **5. EMS Integrated Testing**

## Appendix A. EMS Error Messages

### Abbreviations and Acronyms

#### List of Figures

Figure 3.2-1. EMS-Data Providers Context Diagram.....	3-2
Figure 4.2.3-1. EMS Recognized COSSL Formats .....	4-5
Figure 4.3.1-1. JavaScript - Basic Page Tag Example.....	4-18
Figure 4.4-1. EMS Processing Error - Email Notification Example .....	4-24
Figure 4.4-2. EMS Processing Error - ASCII Error File Example .....	4-25
Figure 4.4-3. EMS Processing Error - Missing Files Email Notification Example.....	4-27
Figure 4.4-4. EMS Unknown Search Terms Email Notification Example.....	4-28

#### List of Tables

Table 3.2-1. EMS and Data Providers Interface Overview .....	3-3
Table 3.2-2. EMS Data Flow Summary (1 of 2).....	3-4
Table 3.2-2. EMS Data Flow Summary (2 of 2).....	3-5
Table 4.2.1-1. Data File Manifest File Format .....	4-3
Table 4.2.7-1. Data Ingest Flat File Format.....	4-10
Table 4.2.7-2. Data Archive Flat Files Format .....	4-11
Table 4.2.7-3. User Profile Flat File Format.....	4-13
Table 4.2.7-4. ESDIS User Type Definitions .....	4-14
Table 4.2.7-5. Product Attribute Search Flat File Format .....	4-15
Table 4.2.7-6. Product Attribute Metadata Flat File Format .....	4-16
Table 4.3.1-1. Page Tag Manifest File.....	4-20

This page intentionally left blank.

# 1. Introduction

---

## 1.1. Identification

This Interface Control Document (ICD) documents the interfaces between the ESDIS Metrics System (EMS) and the Data Providers.

In this document, Data Providers refer to those entities that submit metrics to EMS.

## 1.2 Scope

This ICD describes the interfaces between the EMS and the Data Providers which support the exchange of information on product and service usage and data system operations. These data are in turn used by the EMS for processing and generating metrics. Data and information transmitted include the following:

1. User Profile Flat Files from the Data Provider to the EMS
2. Product Attribute Flat Files from the Data Provider to the EMS
3. Archive and Ingest Flat Files from the Data Provider to the EMS
4. Commercial/Open Source Server Logs (COSSL) or Custom Logs from the Data Provider to the EMS
5. Page Tag Data from the Data Provider to the EMS
6. Error Notification Emails from the EMS to the Data Provider.

Operations procedures and maintenance activity regarding the interfaces instantiated between each Data Provider and the EMS are beyond the scope of this document, and should be documented in Operations Agreements (OA) (e.g., User Acceptance Test (UAT) Plan) or Operations Procedures (OP) documents under the control of the Data Providers with the concurrence of the EMS. The term “Operations Agreement(s)” in this document refers to such a document, though the particular documentation name may vary from Data Provider to Data Provider. Section 2 lists documents which describe the operational interfaces between the EMS and its Data Providers.

This document does not include the EMS or the Data Providers release schedules. Information about the schedules for the implementation of the EMS or the Data Providers releases is contained in the applicable project schedules and is beyond the scope of this document.

The ESDIS Project has responsibility for the development and maintenance of this ICD. Any changes in the interface must be agreed to and assessed at the ESDIS Project Level.

In the event that the EMS or the Data Provider proposes future changes necessary to correct defects or to support enhancements, each will propose such changes to the ESDIS project.

ESDIS will collaboratively assess the changes with the EMS and the Data Providers. ESDIS approves all interface changes and incorporates the approved changes into this ICD.

### **1.3 Purpose**

This document is written to formalize the interpretation and general understanding of the interfaces between the EMS and the Data Providers. This document provides a point of mutual control of external interface definitions via the ESDIS Configuration Control Board (CCB).

### **1.4 Status and Schedule**

This is the baseline ICD for interfaces between the EMS and the Data Providers. It has been designated to be under full Government CCB control. Changes may be submitted for consideration to the ESDIS CCB at any time.

### **1.5 Organization**

Section 1 provides information regarding the identification, scope, purpose and objectives, and organization of this document.

Section 2 is a listing of the related documents.

Section 3 is an overview of the interfaces between the EMS and the Data Providers including security requirements.

Section 4 provides a description of the EMS data types and data exchange mechanisms.

Section 5 is an overview of integrated testing of the EMS.

A list of abbreviations and acronyms is also provided.

## 2. Related Documentation

---

The latest versions of all documents below should be used. The latest ESDIS Project documents can be obtained from URL: <https://cicero.eos.nasa.gov/esdis>. ESDIS documents have a document number starting with either 423 or 505.

### 2.1 Parent Documents

The parent document is the document from which this interface control document's scope and content are derived.

423-10-02                      Earth Science Data And Information System (ESDIS) Project Plan

### 2.2 Applicable Documents

The following documents are referenced within this interface control document, or are directly applicable, or contain policies or other directive matters that are binding upon the content of this volume.

GSFC-05-0059                EMS IT Security Plan  
 NPD 1382.17G                NASA Privacy Policy  
                                       EMS Service Level Agreement (SLA)  
                                       EMS Terms of Use  
                                       EMS User Acceptance Test (UAT) Plan

The latest EMS-related documents can be obtained from URLs listed below:

1. EMS Master Documentation: <http://ems.eos.nasa.gov/documents>
2. EMS Operational Downloads: <http://ems.eos.nasa.gov/downloads>
3. EMS Internal Documentation: <https://romulus.ems.eosdis.nasa.gov/wiki/doku.php>

### 2.3 Information Documents

The following documents, although not directly applicable, amplify or clarify the information presented in this document, but are not binding.

ISO 7498                      International Organization for Standardization, Basic Reference Model for Systems Interconnection

This page intentionally left blank.

## 3. Interface Overview

---

### 3.1 EMS Overview

The ESDIS Metrics System (EMS) is a system that collects and analyzes data on the usage of products and services delivered via the Internet or stored in databases (e.g. archival and ingest) of Earth Observing System (EOS) Data and Information System (EOSDIS) Distributed Active Archive Centers (DAACs) and other related Data Providers. Data stored by the EMS is password protected and viewing is restricted by username. Therefore only users with the appropriate permissions can access data, create reports and export data to standard office applications.

The EMS metrics are reported to NASA HQ to evaluate performance and effectiveness of the EOSDIS. The metrics data and the analysis reports generated from the EMS provide NASA managers the information needed to determine how to best apply resources to support the science community. The type of information to be reported on by each Data Provider includes the following general topics:

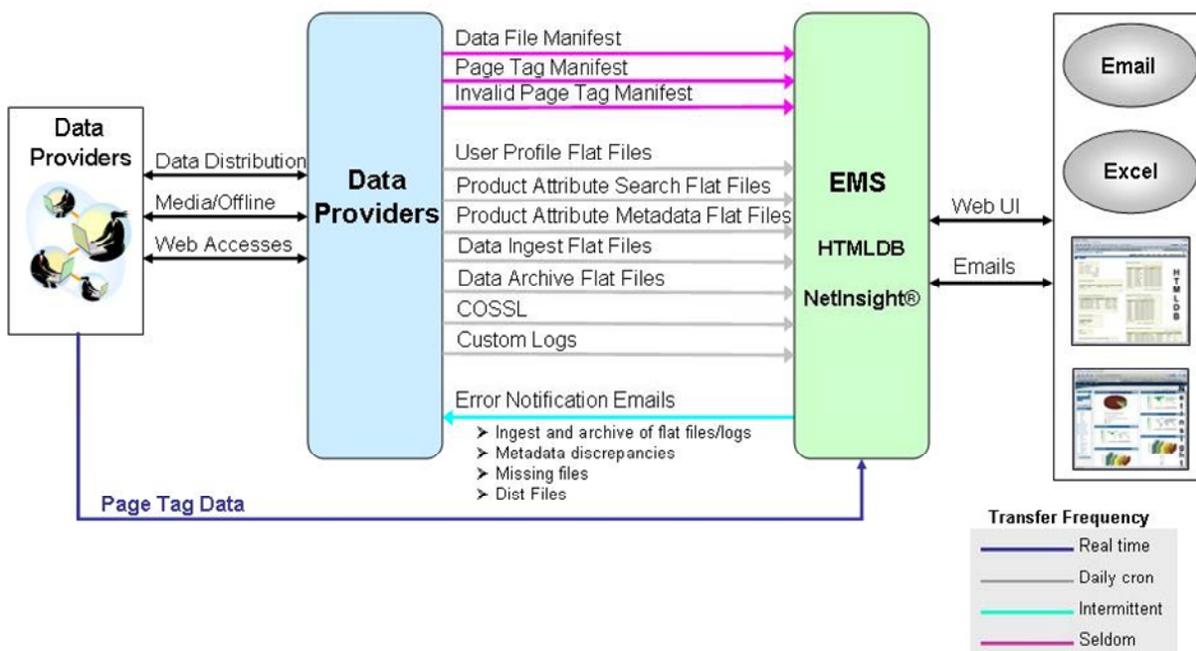
- Amount of data ingested
- Amount of data archived
- Amount of data retrieved from the archive
- Production processing metrics
- Data distributed to external users
- User and user satisfaction information.

The EMS is a combination of custom code and a commercial Web analytics package. The custom code portion, EDGRS+, allows for the retrieval, preprocessing and error reporting of metrics data. Additionally, this component tracks ingest, archival and user profile information collected from Data Providers. The COTS product, NetInsight®, provides the Web analytic capabilities. NetInsight® has a number of preformatted reports, extensive ad hoc reporting capabilities and can export data to standard office applications. The EMS is located at the Goddard Space Flight Center (GSFC) and is operated by the ESDIS Science Operations Office (SOO).

The interface between EMS and the Data Provider will not be implemented in EMS operations until the UAT Plan is signed by the Data Provider.

### 3.2 EMS-Data Providers Interface Context

Figure 3.2-1 describes data transfers between end users, Data Providers, the EMS and the EMS users. These data flows are necessary to capture metrics on product and service usage and data system operations. As shown on the right side of the diagram, metrics collected by the EMS will be made available via standard and ad hoc Web reports to all the EMS users. A username and password will be required to access this information and will limit what data users can view.



**Figure 3.2-1. EMS-Data Providers Context Diagram**

The left side of the diagram represents end user interaction with products and services delivered by the Data Providers. The dotted line from End Users to EMS shows how data on end user Web interactions are captured in ‘real-time’ by the EMS using page tagging technology. Information on end user interactions with Data Providers stored in databases (archive, ingest and user profile information) and/or transfer log files will also be collected by the EMS. This information will be sent daily or as updated to the EMS over a secure connection (rsync tunneled through ssh is the preferred method). Table 3.2-1 summarizes the data exchanges between the EMS and the Data Providers as well as identifies the associated EMS data type along with the ICD section addressing the data.

**Table 3.2-1. EMS and Data Providers Interface Overview**

Source	Dest.	EMS Data Type / ICD Section	Data	Transfer Frequency/ Volume	Transfer Mechanism
Data Provider	EMS	Data File/ Section 4.2.1	Data File Manifest	As updated by Data Provider/ < 500 KB	Excel spreadsheet via email
Data Provider	EMS	Page Tag/ Section 4.3.2	Page Tag Manifest	As updated by Data Provider/ < 500 KB	Excel spreadsheet via email
Data Provider	Data Provider Web pages	Page Tag/ Section 4.3.1	Page Tag JavaScript	N/A/ (script is less than 12 KB)	N/A
Client	EMS	Page Tag/ Section 4.3	Page Tag Data	Real time/ < 16 KB	HTTP
Data Provider	EMS	Page Tag/ Section 4.3.4	Invalid Page Tag Manifest	As updated by Data Provider / < 500 KB	Excel spreadsheet via email
Data Provider	EMS	Data File/ Section 4.2.7.3.3	User Profile Flat Files	Daily/ 0-1 MB	rsync tunnel through SSH or SFTP
Data Provider	EMS	Data File/ Section 4.2.7.3.4	Product Attribute Search Flat Files	Daily/ 0-5 MB	rsync tunnel through SSH or SFTP
Data Provider	EMS	Data File/ Section 4.2.7.3.4	Product Attribute Metadata Flat Files	Daily/ 0-5 MB	rsync tunnel through SSH or SFTP
Data Provider	EMS	Data File/ Sections 4.2.7.3.1, 4.2.7.3.2	Archive and Ingest Flat Files	Daily/ 0-150 MB	rsync tunnel through SSH or SFTP
Data Provider	EMS	Data File/ Section 4.2.3	Commercial/ Open Source Server Logs	Daily/ 500 MB	rsync tunnel through SSH or SFTP
Data Provider	EMS	Data File/ Section 4.2.4	Custom Logs	Daily/ 500 MB	rsync tunnel through SSH or SFTP
EMS	Data Provider	NA/ Section 4.4.3	Error Notification Emails	As required/ N/A	Email

\* If data is currently present, development time could be substantially less.

In the event of a missing file or if the EMS has a failure ingesting data, the EMS will deliver an error notification email to the Data Provider. The content of the emails is discussed in detail in section 4.4.

Table 3.2-2 summarizes the data flows between the EMS and Data Providers, as numbered in Figure 3.2-1. Table 3.2-2 shows three categories of data exchange interfaces; manifest file exchange, data and flat file exchange and error notification emails.

**Table 3.2-2. EMS Data Flow Summary (1 of 2)**

Item No.	Source	Dest.	Freq.	Message	Data Format	Transfer Mechanism
<b>Manifest Files (Metadata describing files or data sent from Data Providers and end users to the EMS)</b>						
1	Data Provider	EMS	As Updated	Data File Manifest – Data Providers must completed this manifest before data and flat files are sent to the EMS. This file describes characteristics of data and flat files. When updates to the Data File Manifest are necessary, Data Providers must send an updated Data File Manifest to the EMS. The revised manifest should only contain entries for valid log files.	Microsoft Excel File	Email
2	Data Provider	EMS	As Updated	Page Tag Manifest – Data Providers must complete this manifest before the EMS will capture and report web usage metrics through the NetInsight interface. The Page Tag Manifest is used to track domains where page tags have been implemented and may be optionally configured to capture special web service usage. When updates are necessary Data Providers must send an updated Page Tag Manifest in addition to an Invalid Page Tag Manifest (see next).	Microsoft Excel File	Email
3	Data Provider	EMS	As Updated	Invalid Page Tag Manifest – Data Providers must complete this manifest when they would like the EMS to stop collecting metrics from domains. This manifest can be sent in conjunction with a Page Tag Manifest or separately.	Microsoft Excel File	Email
<b>Data and Flat Files (Files containing records of distribution, ingest, archive or other transactions both online and offline)</b>						
4	Data Provider	EMS	Daily	User Profile Flat Files - User Profiles are data characterizing the type, affiliation, username and email, IP address or hostname of users. This information is primarily used to separate data distribution and web service usage into categories of production, end, QA/Testing and Data Provider internal users.	ASCII Text Flat Files. See section 4.2.7 for details.	Rsync push tunneled through SSH
5	Data Provider	EMS	Daily	Product Attribute Search Flat Files – This file offers Data Providers a mechanism to map product identifiers to search terms. Product identifiers can be associated with more than one search term and may occur multiple times within the Product Attribute Search Flat File.	ASCII Text Flat Files. See section 4.2.7 for details.	Rsync push tunneled through SSH

**Table 3.2-2. EMS Data Flow Summary (2 of 2)**

Item No.	Source	Dest.	Freq.	Message	Data Format	Transfer Mechanism
6	Data Provider	EMS	Daily	Product Attribute Metadata Flat Files – This file provides ancillary information about product identifiers. This information is joined to the Product Attribute Search Flat Files by the product field. All products in the Product Attribute Search Flat Files must be contained within the Product Attribute Metadata Flat File.	ASCII Text Flat Files. See section 4.2.7 for details.	Rsync push tunneled through SSH
7	Data Provider	EMS	Daily	Data Ingest Flat Files – This file is a summary of all granules, products, or files that have been ingested into the Data Provider’s system.	ASCII Text Flat Files. See section 4.2.7 for details.	Rsync push tunneled through SSH
8	Data Provider	EMS	Daily	Data Archive Flat Files - Archive data is a summary of all granules, products, or files that have been archived into the Data Provider’s system.	ASCII Text Flat Files. See section 4.2.7 for details.	Rsync push tunneled through SSH
9	Data Provider	EMS	Daily	Commercial / Open Source Server Logs - The EMS automatically recognizes most Commercial or Open Source Server Logs (COSSL) formats. In order for the EMS to ingest standard server logs, the type of log must be specified in the Data File Manifest File and be of a type listed in Figure 4.2.3.-1. Examples of valid COSSL types include NCSA Common/Combined (aka Apache), wu-ftp, NcFTPD, Gauntlet, etc...	ASCII Text Flat Files in the native COSSL format.	Rsync push tunneled through SSH
10	Data Provider	EMS	Daily	Custom Logs - Custom logs are not required by the EMS. Custom logs allow Data Providers to send the EMS information from customized servers/log writers, offline activities, Spreadsheets or just about anything that adheres to the custom log file requirements. For example, offline orders or media distribution could be tracked in an Excel Spreadsheet	ASCII Text Flat Files. See section 4.2.7 for details.	Rsync push tunneled through SSH
Notification Emails						
11	EMS	Data Provider	In event of a problem	See Appendix A for the error messages	ASCII Text Error File. See section 4.4 and Appendix A for details.	Email (notification) and Rsync push tunneled through SSH (error file)

### **3.3 EMS User Interface Overview**

The EMS provides the Data Providers with access to metrics and reports via a Web portal. The portal provides access to metrics collected by the EMS in addition to processing errors. The primary method for accessing web user, usage and distribution metrics will be through the NetInsight® interface. An additional Web interface, HTMLDB/APEX allows users access to archive, ingest, distribution, and error metrics. Users are required to have a unique username and password to access these resources.

### **3.4 Network Protocols and Topology**

The EMS provides inter-networking services that are based on protocols and standards corresponding to layers 1 through 4 of the Open Systems Interconnection (OSI) Reference Model. These include, respectively, the physical, datalink, network, and transport layers. The transport layer protocol provides data consistency functions. The network, datalink and physical layers play significant roles in defining external interfaces (i.e., between the EMS and non-EMS networks/systems). The EMS routers provide the physical demarcation points between the EMS networks and external networks/systems.

The EMS resides on the EOS Mission Support network (EMSn), formerly EOSDIS Backbone network (EBnet), open network at the GSFC.

#### **3.4.1 Transport Layer Protocol**

The transport layer protocol used for communications between the EMS processes and non-EMS processes is the Transmission Control Protocol (TCP). TCP is a reliable connection-oriented, end-to-end protocol designed to fit into a layered hierarchy of protocols that support multi-network applications. It provides for reliable inter-process communication between pairs of processes in host computers attached to networks within and outside of the EMS.

#### **3.4.2 Network Layer Protocols**

The network layer provides the functional and procedural means to exchange network data units (i.e., packets) between devices over network connections, both for connection-mode and connectionless-mode communications. It relieves the transport layer of any concern regarding routing and relay operations associated with network connection. Network layer protocols supported by the EMS networks include Internet Protocol (IP) plus various routing protocols.

##### **3.4.2.1 Internet Protocol (IP)**

The Internet Protocol (IP) is the network protocol that the EMS supports, based on its dominance in industry usage and wide-community support. As part of IP support, Internet Control Message Protocol (ICMP) and Address Resolution Protocol (ARP) are also supported.

### **3.4.2.2 Routing**

The EMS generally uses Routing Information Protocol (RIP) for route exchanges with external networks. Other more robust routing protocols can also be used depending on the need and Data Provider routing policies.

### **3.4.2.3 Physical/Datalink Protocols**

Gigabit ethernet cards are the standard interface used at the physical layer. The EMS typically uses standard protocols in data link protocols. The EMS utilizes several unmanaged switches to enable more systems to become network accessible: Media Access Control (MAC) address, and Logical Link Control (LLC). The MAC sublayer controls how a computer gains network access to data, and permission to transmit. The LLC controls frame synchronization, flow control and error handling.

### **3.4.2.4 Network Topology**

The EMS uses the standard twisted pair Ethernet, which is a logical bus topology.

## **3.4.3 Communications Protocols**

The EMS provides various communications services that are based on protocols and standards corresponding to the applications level of the Open Systems Interconnection (OSI) Reference Model. Some of these services are the Simple Mail Transport Protocol (SMTP), the Secure Shell Suite (SSH) with applications: ssh, secure copy (scp), secure file transfer protocol (sftp), and the rsync application. Rsync provides efficient and flexible data transfer and SSH provides encrypted connection sessions and simple data transfer. Unencrypted connection services like telnet or general ftp that transfer account information in plain text are not permitted or supported.

### **3.4.3.1 Simple Mail Transport Protocol**

All electronic mail (e-mail) message exchange is achieved through the use of Internet e-mail messages. The protocol for Internet e-mail transfer is the Simple Mail Transfer Protocol (SMTP).

### **3.4.3.2 Secure Shell**

The EMS supports secure shell for secure transfer of data. Secure shell is a suite of TCP/IP-based applications supporting encrypted remote access, file transfer and data tunneling. It uses public key authentication and/or encrypted passwords to establish connections. The EMS supports file transfer via the scp and sftp applications of the secure shell suite and via rsync over ssh. In the latter usage application, rsync provides the transport mechanism and ssh provides data encryption.

### 3.4.3.3 Rsync

The EMS uses the Rsync client tool to achieve highly flexible and efficient data transfer functionality. Rsync is recommended for use as the main data transfer application for the Data Provider clients with ssh as a wrapper to provide encryption. Applied in this way, rsync provides efficiency through compression, data integrity through checksums, and maintains the file permissions and time stamps of the transferred data. The ssh wrapper encrypts all data transferred by rsync.

## 3.5 Security

This section describes the security requirements with respect to the EMS users.

### 3.5.1 Data Transfer Accounts

The EMS Team will provide each Data Provider with an account used for transferring data to the EMS. Data Providers access their EMS data transfer account with password-less key-based authentication. Access is restricted to a Data Provider defined IP address as specified in the OA.

The Data Provider account owner is responsible for the safekeeping of the ssh-keys on local systems. If a system is compromised or the IP address changed, the Data Provider is required to contact the EMS support immediately at [ems-help@lists.nasa.gov](mailto:ems-help@lists.nasa.gov), but no later than 24 hours after the incident is reported.

Updates to the firewall rules can take from one to three weeks to be completed. Therefore, the Data Provider should notify EMS of IP address changes in advance. If the Data Provider's IP address changes and EMS is not notified in advance, the Data Provider will not be able to send EMS any log/flat files. Thus, the availability and timeliness of metric reports will be impacted.

### 3.5.2 Data Access Accounts

The EMS ad hoc and standard reports will be available via HTMLDB/APEX. To access this information, a username and password is required. The user can obtain an account by filling out an account request form at <http://ems.eos.nasa.gov/accountRequest.shtml>. The EMS Team will create the accounts for each approved user.

### 3.5.3 Privacy Information and Policies

The EMS users must comply with the NASA and ESDIS privacy information, policies, and guidelines when accessing and using the data from the EMS. The NASA Policy Directive NPD 1382.17G relates to the use, disclosure and protection of records containing information in identifiable form that is collected by NASA. Information in identifiable form is any information that directly identifies an individual by name or any other combination of elements or codes that could specifically identify an individual. In Privacy Act terminology, these data are referred to as statistical records and are maintained for statistical research or reporting purposes only. NASA GSFC ESDIS (Code 423) and the EMS users must comply with the "ESDIS Guidelines For Implementation of NASA Privacy Policy" document available at the EMS Master

Documentation URL noted in section 2.2. The ESDIS privacy guidelines are to provide a basis for handling customer information that will be consistent with U.S. regulations and NASA Privacy Policy in the context of changing environments within the ESDIS Project.

Additionally, upon access, receipt, or review of data from the EMS, the EMS user is agreeing to abide by the following ESDIS guidelines:

1. The records may only be used for the intended purpose, which is for statistical analysis leading to better understanding of NASA's Earth science user needs and improving effectiveness of NASA's Earth science programs, data product and services.
2. The records must not be used to derive a contact list for any purpose (e.g., a survey or other additional information gathering activity).
3. The records must be kept private, not accessible except by the proper people (i.e., the people doing the analysis, or with responsibility for program direction). The records must not be disclosed to anyone outside of NASA for any other purpose.
4. Results of an analysis that is intended for public disclosure or to be made accessible outside of NASA must not reveal individual information or be capable of being used to determine individual information.
5. The records must be discarded when the information is no longer needed.

In addition, the EMS Terms of Use agreement must be read and agreed to before the user is given an EMS user account. The EMS Terms of Use agreement discusses the expectations, assumptions, and rules concerning the EMS environments, data consistency, and data publishing. The agreement can be found on the EMS website.

This page intentionally left blank.

## 4. EMS Data Type Descriptions

---

### 4.1 EMS Data Types

EMS primarily collects metrics from the Data Provider in the form of data files and page tags. Data files, which consist of log files and flat/data files, are a flexible mechanism for Data Providers to supply information on archive, ingest, distribution, special services, offline ordering, media distribution and other non-Web activities to the EMS. Refer to Section 4.2 for details on data files. Page tags, described in Section 4.3, allow the EMS to collect Web usage information from Data Providers' Web sites.

### 4.2 EMS Data Files

Data files are a flexible mechanism for Data Providers to supply information on archive, ingest, distribution, special services, offline ordering, media distribution and other non Web activities to the EMS. Data Providers only need to provide the EMS with HTTP or Web server logs if these servers are used to distribute data products or metadata. The EMS accepts the following types of data files, each of which has unique requirements identified below:

- Commercial or Open Source Server Logs (COSSL)
  - A server log is a file (or several files) automatically created and maintained by a server (HTTP, FTP, SSH, etc...), of activity performed by it. Examples of COSSL include National Center for Supercomputing Applications Common/Combined (Apache), wu-ftp, and other common log formats. Each line in the log must represent one transaction. HTTP access logs are not required to be sent to the EMS unless HTTP servers are used to distribute data.
- Custom Logs
  - Custom logs allow Data Providers to specify the content and format of data delivered to the EMS. Custom logs allow Data Providers to send the EMS information from customized servers/log writers, offline activities, spreadsheets or just about anything that adheres to the custom log file requirements. Each line in the log must represent one transaction. More information on Custom Log format is available in an external document, "The EMS Custom Log Implementation Guide", located at <http://ems.eos.nasa.gov/documents>.
- Flat files
  - Flat files are created by running queries on a database or data source and having the results output to a plain text file. There is one record per line. Each record is divided into fields using '|&|' as a delimiter. The EMS accepts Data Ingest, Data Archive, User Profile, and Product Attribute flat files.

## **4.2.1 Data File Manifest File**

The EMS data files are automatically or manually created by Data Providers and then securely transmitted to the EMS. All data files should be delivered as specified in Table 3.2-1. Timing of delivery will be specified in the Data File Manifest File. The Data File Manifest File, as presented in Table 4.2.1-1, is used by Data Providers and the EMS Team to identify all data sources delivered to the EMS. If the EMS does not receive expected files, as outlined in the Data File Manifest File, or encounters errors during processing of received files, an email will be sent to the appropriate Data Provider identifying missing files and/or processing errors.

### **4.2.1.1 Data File Manifest File Format**

The Data Providers must complete a Data File Manifest File (worksheet) in order to identify all data files being delivered to the EMS. A template is available for download from <http://ems.eos.nasa.gov/downloads>. Each column must represent a single data file and follow the format described below in Table 4.2.1-1.

**Table 4.2.1-1. Data File Manifest File Format**

Category	Description (examples only)	Req'd	Data File 1	Data File 2	Data File N
Provider	Name of the Data Provider (e.g., GES, GSFCV0, GSFC S4PA, EDC, LATIS, MODAPS). Mutually agreed upon by EMS and Data Provider All uppercase characters.	Y			
Provider Category	Classification of the Data Provider (e.g., SIPS, ECS, NON-ECS). Mutually agreed upon by EMS and Data Provider.	Y			
Reference URL	Designates the base address/url of where data file resides (e.g., <a href="http://podaac.jpl.nasa.gov">http://podaac.jpl.nasa.gov</a> ). If none exists, enter "NA".	A			
Source IP	Identifies the IP address of the host(s) that files to EMS. This information is required by to allow file transfers. Each data file should have an entry (even if the source IP remains the same).	Y			
Root File Name	Designates the root name of the file in accordance with the EMS ICD file naming convention. The date does not need to be included; only the root of the file name is required (e.g., podaac_log_http.flt).	Y			
Data File Format	Designates the format of the data file and whether it is standard or custom. If standard then specify the format type. (See Figure 4.2.3-1 – e.g., Cisco PIX, Gauntlet, IBS Firewall, IIS Standard/Extended, Microsoft Proxy, etc.)	Y			
Push Frequency	Designates how often the files are sent to the EMS. The minimum push frequency is once per day for all data files.	Y			
Push Time	Time of push events based on Greenwich Mean Time (GMT).	Y			
Comments	Any comments that Data Provider can provide to assist the EMS staff in understanding the data and the process.	A			
Data File Format Details	If the data file format is custom, refer to The EMS Custom Log Implementation Guide at <a href="http://ems.eos.nasa.gov/documents">http://ems.eos.nasa.gov/documents</a> for the requirements.	Y (custom only)			

Note: 'Y' indicates mandatory data. 'A' indicates this information is required if available.

#### 4.2.1.2 Data File Manifest File Naming Convention

The name of the Data File Manifest File must be in the following format:

<YYYYMMDD>\_<Provider>\_DataFileManifest.xls

where

YYYY	designates the 4 digit year for the time the Data File Manifest File was created
MM	designates the 2 digit month, 01 through 12
DD_	designates the 2 digit day, 01 through 31, followed by an underscore
Provider_	designates the provider of the data, mutually agreed upon by EMS and Data Provider, followed by an underscore
DataFileManifest.xls	indicates that the file is the Data File Manifest worksheet

### 4.2.1.3 Data File Manifest File Updates

Each time a Data Provider adds or removes a data file, or makes any changes that impact fields within the Data File Manifest, a revised Data File Manifest must be sent to the EMS. The revised Data File Manifest must contain all valid entries for data files that will be sent to the EMS until another revision is submitted. The Data File Manifest must not contain invalid entries or error notification emails will be sent.

## 4.2.2 Data Files Interface Mechanism

Data Providers are responsible for pushing data files to the EMS as specified in Table 3.2-1. Each file Data Providers deliver to the EMS should contain one day's worth of data. In general, data within the file will correspond to the date stamp in the filename. Exceptions to this rule include Ingest data. Ingest Requests consist of multiple granules which may require additional processing time spanning more than 24 hours. In this case, the processingEndDateTime in the data file will vary from the date stamp in the filename.

The EMS Team will provide each Data Provider with an account for transferring data to the EMS. Data Providers will be able to access their EMS Data Gathering (ws1) account with password-less key-based authentication. Access to this account will be restricted to a Data Provider defined IP address as defined in the Data File Manifest (see Table 4.2.1-1). The preferred method for Data Providers to transfer data files to the EMS is an rsync client tunneled through ssh. Alternatively, Data Providers can use sftp if the rsync client is not available.

In the event Data Providers need to resend a data file because of corruption, errors, or any other reason, the original filename for that day's worth of data must be reused with the addition of a suffix of '.rev1' to show it is an updated file. If additional updates are necessary for the same day's worth of data then rev1 will be incremented for each update (e.g. rev1, rev2, rev3... revN). In other words, if a flat file describing ingest was created on 14 January 2006 the filename would be in the format of 20060114\_<Provider>\_<FileType>\_<DataSource>.flt. If the contents of this file were determined to be incorrect, the Data Provider would resend the ingest data for that day but change the filename to 20060114\_<Provider>\_<FileType>\_<DataSource>.flt.rev1.

In addition, each file is processed before it is stored in the EMS database. All IP addresses within the file are converted to hostnames by performing a DNS Lookup. This hostname replaces the IP address in the flat file and is stored in the database.

### 4.2.3 Commercial or Open Source Server Logs

The EMS automatically recognizes most Commercial or Open Source Server Logs (COSSL) formats. In order for the EMS to ingest standard server logs, the type of log must be specified in the Data File Manifest File and be of a type listed in Figure 4.2.3-1. The Data Provider COSSL distribution files must include the transfer status field. A transfer status field is enabled by default for each of the types specified in Figure 4.2.3-1. Some Data Providers may have logs in a nonstandard format or augment standard log formats with additional information that the EMS cannot read by default. If log files are in a nonstandard format, please refer to “The EMS Custom Log Implementation Guide”, located at <http://ems.eos.nasa.gov/documents>.

Cisco PIX	Microsoft Proxy	RealSystem Server (Helix)
Gauntlet	NCSA Common / Combined*	Sidewinder
IBM Firewall	Netscape Flexible	Squid
IIS Standard / Extended	Open Market Extended	W3C Extended
wu-ftp	Raptor Eagle	WatchGuard
NcFTPd		WebSite

\* Apache Logs are in NCSA format. The common / combined options are set in the Apache configuration file.

**Figure 4.2.3-1. EMS Recognized COSSL Formats**

Data providers should use page tags to collect web usage information. If log files are from mixed environments, (web and data), Data Provider should parse the log and extract only the data related to ingest, archive, or distribution. Please refer to Table 4.2.7-5 (Product Attribute Search Flat File Format) for information on how to provide this information.

### 4.2.4 Custom Logs (Media Orders, Offline Activities, Non-standard Information)

Custom logs allow Data Providers to send the EMS information from customized servers/log writers, offline activities, spreadsheets or just about anything that adheres to the custom log file requirements. For example, offline orders or media distribution could be tracked in an Excel spreadsheet; or additional information could be appended to each transaction captured in NCSA Common/Combined logs, such as special service usage or user type. The custom log file formats are defined in an external document, “The EMS Custom Log Implementation Guide”, located at <http://ems.eos.nasa.gov/documents>. The Data Provider’s custom log must include the transStatus field, which is required for all custom log formats. **Additionally, it is highly recommended that Data Providers include the product version as a field within custom distribution files. ESDIS uses this information to do version based product reports.**

#### 4.2.4.1 Special Services Captured in Logs

Data Providers are strongly encouraged to provide EMS with custom logs which capture services or other activities that fall outside of standard data delivery services. Some examples include on demand spectral or spatial subsetting, creation of data CDs/DVDs, bulk transfer of data via

external hard drive, file format conversions, among others. All services Data Providers offer users are valuable and should be tracked. This information should be provided to EMS to ensure ESDIS management is aware of these valuable contributions to the user community.

#### 4.2.5 Log Files Naming Convention

Filenames are composed of three parts:

Datestamp	designates the year, month and day that correspond to the request date of the data represented in the data. If a revised file is being sent (see below) the timestamp represents the date on which the original file was created with the .rev<1-n> file extension used to identify the file as a revision.
Root File Name	aggregation of the Provider, File Type, and Data Source components of a file name. It must be unique for each provider.
Extension	designates the type of file and the revision status.

The name of the log files must be in the following format for all the Data Providers:

$\underbrace{\langle \text{YYYYMMDD} \rangle}_{\text{Datestamp}} - \underbrace{\langle \text{Provider} \rangle \_ \langle \text{FileType} \rangle \_ \langle \text{DataSource} \rangle}_{\text{Root File Name}} \underbrace{\langle \text{flt.rev} \langle 1-n \rangle \rangle}_{\text{Extension}}$

where

YYYY	designates the 4-digit year
MM	designates the 2 digit month, 01 through 12
DD_	designates the 2 digit day, 01 through 31, followed by an underscore
Provider_	designates the provider of the data, mutually agreed upon acronym defined in the OA
FileType_	designates the type of file sent, followed by an underscore

where Type:

“DistFTP”	FTP distribution log
“DistHTTP”	HTTP distribution log
“DistMedia”	Media distribution log
“DistSCP”	Scp log
“DistSpcSrv”	Special services log
“DistFireWall”	Firewall log
“DistCustom”	Custom log
“DistFTPSUBSCRIPT”	Subscriptions accessed by FTP

“DistSCPSUBSCRIPT”	Subscriptions accessed by SFTP/SCP
“ECHOsearches”	Custom for ECHO Search Results
“ECHOOrders”	Custom for ECHO Orders

DataSource designates the source or type of system generating the log file.

where Type:

<Custom> Data Providers may populate with information needed to uniquely identify the source of a log file, including virtual server names, subsystem names. Validated against Data Source portion of the RootFileName provided in the Data File Manifest

.flt indicates the file is a plain text file

.rev<1-n> indicates the file has been resent because of errors, the number is incremented for each update starting with .rev1 (e.g. rev1, rev2, rev3... revN)

#### 4.2.6 Log Files Updates

Data Providers should supply updates, to EMS when errors are found or new information is added to log files. It is very important that the updated log file contain all records that should appear in metric reports. When a revision file is received, EMS deletes the previous data corresponding to the EMS filename and replaces it with the revised data. The naming convention for the updated log file must follow the format described above with the appended “.rev<1-n>”.

#### 4.2.7 EMS Flat Files

Each Data Provider will extract and provide the required metrics data from their respective systems in a flat file (ASCII text file) for each subsystem. The EMS flat files consist of ingest, archive, user profile, and product attribute. Flat files are intrinsically in character format. However, fields contained in the flat files should be convertible to the EMS data types defined for each file type. The metrics data contained in the flat file(s) must be in the appropriate order and format.

Data Providers will develop scripts necessary to generate and send flat files to the EMS server (ws1.ems.eosdis.nasa.gov) according to Table 3.2-1. The EMS will perform daily, automated tasks to ensure that required flat files are received from each Data Provider. If data is expected and not received from a given Data Provider, the EMS Team and Data Provider will receive automatic notification.

Once the data is delivered to the EMS, it is then loaded into the EMS database via automated scripts. All records and data that fail the load process will be sent back to the originating Data Provider with details indicating the reason for the failure.

### 4.2.7.1 Flat Files Naming Convention

Filenames are composed of three parts:

Datestamp	designates the year, month and day the content of the data file was created. If a revised file is being sent (see below) the datestamp represents the date on which the original file was created with the .rev<1-n> file extension used to identify the file as a revision.
Root File Name	aggregation of the Provider, File Type, and Data Source components of a file name. It must be unique for each provider.
Extension	designates the type of file and the revision status.

The name of the log files must be in the following format for all the Data Providers:

$\underbrace{\langle\text{YYYYMMDD}\rangle}_{\text{Datestamp}} - \underbrace{\langle\text{Provider}\rangle\_ \langle\text{FileType}\rangle\_ \langle\text{DataSource}\rangle}_{\text{Root File Name}} . \underbrace{\text{flt.rev}\langle 1-n \rangle}_{\text{Extension}}$

where

YYYY	designates the 4 digit year for the time the Data Ingest Flat File was created
MM	designates the 2 digit month, 01 through 12
DD_	designates the 2 digit day, 01 through 31, followed by an underscore
Provider_	designates the provider of the data, mutually agreed upon acronym defined in the OA, followed by an underscore.
FileType_	designates the type of flat file sent, followed by an underscore

where Type:

“Ing”	Data Ingest Flat Files (see section 4.2.7.3.1)
“Arch”	Data Archive Flat Files (see section 4.2.7.3.2)
“UsrProf”	User Profile Flat Files (see section 4.2.7.3.3)
“searchExp”	Product Attribute Search Files (see section 4.2.7.3.4.1)
“Meta”	Product Attribute Metadata Flat Files (see section 4.2.7.3.4.2)

DataSource	designates the database table name (for ECS) or data source file/table name (for Non-ECS)
.flt	indicates the file is a flat file

.rev<1-n> indicates the file has been resent because of errors, the number is incremented for each update (e.g. rev1, rev2, rev3... revN)

#### **4.2.7.2 Flat Files Updates**

Updated flat files should be sent to the EMS whenever data processing failures are encountered or data corruption is detected. The updated data flat file must only contain the corrected records. The naming convention for the updated data flat file must follow the format described above with the appended “.rev<1-n>”.

#### **4.2.7.3 Flat Files Formats**

All Data Providers must format their flat files as described in the tables below for each subsystem. There are no header records. There is one record per line. Each record is divided into fields using the delimiter “|&|”. Flat files are intrinsically in character format. However, fields contained in the flat files should be convertible to the EMS data types defined below.

##### **4.2.7.3.1 Data Ingest Flat Files**

The EMS accepts data ingest flat files from all Data Providers. The Ingest metric is the amount of data that comes into the Data Provider’s system from outside that data center over a period of time. All of the ingest data within the ingest flat file should have a processingEndDateTime that is equal to the date stamp in the filename" Needs to be amended to include exceptions to this general rule. Suggest the following: In general, data within the file will correspond to the date stamp in the filename. Exceptions to this rule include Ingest data. Ingest Requests consist of multiple granules which may require additional processing time spanning more than 24 hours. In this case, the processingEndDateTime in the data file will vary from the date stamp in the filename. Table 4.2.7-1 describes the requirements and format for the Data Ingest Flat File.

**Table 4.2.7-1. Data Ingest Flat File Format**

Field Order	Max Field Len	Field Name	Description	EMS Data Type	ECS Req'd	Non-ECS Req'd
1	40	dbID	This is the Data Provider's unique identifier for each record. A record could be a granule, product, or file.	varchar2	Y	Y
2	80	product	This is a product or dataset identifier. For ECS products it is the Shortname and for Non-ECS it is the productid. This must be identical to the product field in the Product Metadata Attribute Flat File.	varchar2	Y	Y
3	15	productVolume	Total data volume (in bytes) of a granule, product, or file to be ingested. This value must be 0 or greater.	number	Y	Y
4	20	productState	This is the state of the collection, granule, product, or file after ingest processing (i.e., "Successful", "Cancelled", etc.)	varchar2	Y	Y
5	80	externalDataProvider	This is the name of the external Data Provider from where the ingest data was sent.	varchar2	Y	Y
6	25	processingStartDateTime	This is the starting date and time of the ingest process. This value must be earlier than the current date and time and the processingEndDateTime. Format: YYYY-MM-DD HH:MMAMorPM	timestamp	Y	Y
7	25	processingEndDateTime	This is the ending date and time of the ingest process. This value must be earlier than the current date and time. Format: YYYY-MM-DD HH:MMAMorPM	timestamp	Y	Y
8	15	timeToArchive	Elapsed time (in seconds) from submission of an archive request to Data Server to receipt of completion (success or fail). This value must be 0 or greater.	number	A	A
9	15	timeToPreprocess	Elapsed time (in seconds) from start of preprocessing of a granule, product, or file to time of completion (success or fail) of preprocessing. This value must be 0 or greater.	number	A	A
10	15	timeToXfer	Elapsed time (in seconds) from start of transfer for first file in granule to time of receipt of status (success or fail) for last file in granule. This value must be 0 or greater.	number	A	A

Note: 'Y' indicates mandatory data. 'A' indicates this information is required if available. ECS refers to ECS Data Providers and Non-ECS refers to all other Data Providers.

#### 4.2.7.3.2 Data Archive Flat Files

The EMS accepts data archive flat files from all Data Providers. Archive data is a summary of all granules, products, or files that have been archived into the Data Provider's system. All of the archive data within the archive flat file should have an insertTime that is equal to the date stamp in the filename. Table 4.2.7-2 describes the requirements and format for the Data Archive Flat File.

**Table 4.2.7-2. Data Archive Flat Files Format**

Field Order	Max Field Len	Field Name	Description	EMS Data Type	ECS Req'd	Non-ECS Req'd
1	40	dbID	This is the Data Provider's unique identifier for each record. A record could be a granule, product, or file.	varchar2	Y	Y
2	80	product	This is a product or dataset identifier. For ECS products it is the Shortname and for Non-ECS it is the productid. This must be identical to the product field in the Product Metadata Attribute Flat File.	varchar2	Y	Y
3	15	productVolume	Total data volume (in bytes) of a granule, product, or file to be archived. This value must be 0 or greater.	number	Y	Y
4	15	totalFiles	The number of files per granule. This value must be 0 or greater.	number	Y	A
5	25	insertTime	The time of original insertion. This value must be earlier than the current date and time. Format: YYYY-MM-DD HH:MMAMorPM	timestamp	Y	Y
6	25	beginningDateTime	This is the beginning data observation date/time of a particular granule, product, or file. This value must be earlier than the current date and time and the endingDateTime. Format: YYYY-MM-DD HH:MMAMorPM	timestamp	Y	Y
7	25	endingDateTime	This is the ending data observation date/time of a particular granule, product, or file. This value must be earlier than the current date and time. Format: YYYY-MM-DD HH:MMAMorPM	timestamp	A	A
8	25	productionDateTime	This is the data processing date and time of a specific granule, product, or file. This value must be earlier than the current date and time. Format: YYYY-MM-DD HH:MMAMorPM	timestamp	A	A
9	100	localGranuleID	Unique identifier for a locally produced granule, product, or file.	varchar2	Y	A
10	15	versionID	Algorithm version number used to generate a granule, product, or file.	varchar2	A	A
11	1	deleteFromArchive	Flag indicating that a granule, product, or file may be deleted from archive.	char	Y	A
12	25	deleteEffectiveDate	Date on which the entry will be deleted. Format: YYYY-MM-DD HH:MMAMorPM	timestamp	A	A
13	25	lastUpdate	The time of last update. This value must be earlier than the current date and time. Format: YYYY-MM-DD HH:MMAMorPM	timestamp	A	A

Note: 'Y' indicates mandatory data. 'A' indicates this information is required if available. ECS refers to ECS Data Providers and Non-ECS refers to other Data Providers.

### 4.2.7.3.3 User Profile Flat Files

User Profiles are data characterizing the type, affiliation, username and email, IP address or hostname of users. Additionally Data Providers have the flexibility to create Data Provider specific User Types. Data Providers are required to populate the ESDIS User Type using definitions in section 4.2.7.3.4. This information is used to distinguish data distribution into categories of Production, Science Team, QA/Testing, Data Provider Internal and Public users (refer to section 4.2.7.3.4 for definitions). The EMS will categorize all users not defined in the User Profile Flat File as Public Users.

The EMS will evaluate each distribution record received from Data Providers to categorize users. Evaluations will be performed based on the available fields within each EMS Distribution Data File (COSSL, Custom and Flat Files). Prior to performing lookups all IP address will be resolved to hostnames. The order of evaluation is:

1. If a valid userID exists in the Distribution Data File and the User Profile Flat File the corresponding userType and affiliation will be assigned and the next record will be evaluated.
2. If a valid email address exists in the Distribution Data File and the User Profile Flat File the corresponding userType and affiliation will be assigned and the next record will be evaluated.
3. If a valid hostname or IP address exists in the Distribution Data File and the User Profile Flat File the corresponding userType and affiliation will be assigned and the next record will be evaluated.
4. In no match occurs the record will be assigned the Public user type and the affiliation will be populated with 'Unknown' and the next record will be evaluated.

The EMS will process user profile data from all Data Providers. These files should be provided whenever there is an update (refer to section 3.2). Table 4.2.7-3 describes the requirements and format for the User Profile Flat File.

**Table 4.2.7-3. User Profile Flat File Format**

Field Order	Max Field Len	Field Name	Description	EMS Data Type	ECS Req'd	Non-ECS Req'd
1	80	userID	This field uniquely identifies the user. If a valid user ID does not exist insert the value of the emailOrIp address field.	varchar2	Y	Y
2	1	userType	Designates type of user: 1=Production User, 2=Science Team, 3=QA/Testing User, 4=Data Provider Internal User, 5=Public User	char	Y	Y
3	255	e-mailOrIP	Designates the user's email address, IP address or hostname.	varchar2	Y	Y
4	20	affiliation	Designates the type of business or organization. Valid types areCommercial, NonProfit, Edu, State/Local Gov, FedGov, Foreign, etc.	varchar2	Y	Y
5	50	dpUserType	Designates a Data Provider defined user type. If this field is populated Data Providers must provide definitions for each type specified.	varchar2	A	A
6	150	dpUserTypeDef	Designates the definition of the Data Provider defined User Type in field 5 of the User Profile Flat File. This field is required if the dpUserType is populated.	varchar2	A	A

#### 4.2.7.3.3.1 ESDIS User Type Definitions

User type definitions have significant impacts on reports generated by the EMS. To ensure EMS can adequately support ESDIS reporting requirements Data Providers must populate the esdisUserType field as outlined in Table 4.2.7-4.

**Table 4.2.7-4 ESDIS User Type Definitions**

char	esdisUserType	Definition	Notes
1	Production	A user who is involved in the creation of standard data products and formally associated with a SIPS or other organization responsible for creation of these products.	
2	Science Team	A user formally associated with a defined EOS Science Team using data for objectives of the respective Science Team. If a Science Team user acts in dual capacities (e.g. as a Science Team User and a Public User) the user must use different credentials to access data or the user will be assigned to the Science Team User category.	It is understood that it is difficult to separate individual users into categories as indicated here unless they are diligent in use of different credentials for their different roles. This information will be collected on a best efforts basis.
3	QA/Testing	A user who is involved in operational testing and quality assurance of standard data products and/or ancillary data.	
4	Data Provider Internal	A user (human or machine) that transfers data inside the Data Providers network space for the purpose of backups, internal testing, hardware migration, temporary staging of standard products, etc...	This category will be excluded from ESDIS reports unless otherwise noted.
5	Public	A user who uses products obtained from the Data Providers for scientific or other uses. If a Data Provider does not assign user categories EMS automatically assigns all distributions to this category.	All users are assumed to be Public unless there is an entry in the User Profile Flat File. If a Public user is identified in the User Profile additional information on the user's affiliation will be available for reporting purposes.

#### 4.2.7.3.4 Product Attribute Flat Files

The product attributes are the foundation for the EMS's reporting capabilities. Product attributes are metadata such as instrument, mission, product level, and discipline describing the characteristics of a data product. This information is not natively contained within the EMS data file types (COSSL, Custom Logs and Flat Files) but can be linked or "mapped" to each record within a file because of unique patterns contained within a record. Accuracy of reports is highly dependent on providing comprehensive, consistent, and timely updates to the product attributes.

The EMS expects to receive product attributes in the form of two flat files, the Product Attribute Search and the Product Attribute Metadata. The Product Attribute Search Flat File, described in Table 4.2.7-5, contains a product identifier, an associated search term, and a field to designate a search term's associated system name. The system name in the Product Attribute Search Flat File is joined to the system name specified in the Data File Manifest File. The Product Attribute Metadata Flat File contains information describing characteristics of the product. (See Table 4.2.7-6 Product Attribute Metadata Flat File Format for details.)

#### 4.2.7.3.4.1 Product Attribute Search Flat Files

Product Attribute Search Flat Files are used to lookup products within COSSL and Custom distribution files. The Product Attribute Search Flat File contains information that links search terms with products and the data source portion of the EMS filename (e.g. Section 4.2.5). When a distribution flat file is processed, products are determined by evaluating each request path within the distribution file against search terms defined within the Product Attribute Search Flat File.

Search terms are regular expressions and may consist of a file name, directory, path, Earth Science Data Type (ESDT), Data Provider internal product ID or other information that uniquely identifies a product. Once a product has been identified by the regular expression it is joined with information from the Product Attribute Metadata Flat File to identify attributes of the product. Table 4.2.7-5 describes the requirements and format for the Product Attribute Search Flat File.

Describe how revisions to previously submitted search terms are processed.

**Table 4.2.7-5. Product Attribute Search Flat File Format**

Field Order	Max Field Len	Field Name	Description	EMS Data Type	EMS Req'd
1	80	product	This is a product or dataset identifier. For ECS products it is the Shortname and for Non-ECS it is the productid. This must be identical to the product field in the Product Metadata Attribute Flat File. The value will be converted to uppercase during processing.	varchar2	Y
2	200	searchTerm	File name, directory, path, ESDT, Data Provider internal product IDs or other information that uniquely identifies a data product as it appears in an EMS Data file. The searchTerm should not include URL query strings and associated name value pairs. searchTerms can include full strings or substrings. Values within this field are always treated as regular expressions (e.g. '+MOD1[1-9].+'). Therefore reserved grep/egrep characters should only be used when they are needed.	varchar2	Y
3	50	dataSource	Designates the name of the system, subsystem, file, table or other identifying information where the logs/flat files/metadata was created (e.g. ECSDataPool, Giovanni-A, Giovanni-B, LAADS). Data Providers have latitude in specifying the Data Source. The value will be validated against the DataSource portion of the Root File Name. See the Naming Convention sections above. <b>The combination of search term and data source must be unique because they are used as a primary key by the EMS.</b>	varchar2	Y

Note: 'Y' indicates mandatory data.

#### 4.2.7.3.4.2 Product Attribute Metadata Flat Files

The Product Attribute Metadata Flat File provides ancillary information about product identifiers. This information is joined to the Product Attribute Search Flat Files by the product field. All products in the Product Attribute Search Flat Files must be contained within the Product Attribute Metadata Flat File. A product must only appear once in the Product Attribute Metadata Flat File. Table 4.2.7-6 describes the requirements and format for the Product Attribute Metadata Flat File.

**Table 4.2.7-6. Product Attribute Metadata Flat File Format**

Field Order	Max Field Len	Field Name	Description	EMS Data Type	EMS Req'd
1	80	product	This is a product identifier. For ECS products it is the Shortname and for Non-ECS it is the productid or dataset. The value will be converted to uppercase during processing	varchar2	Y
2	1024	metaDataLongName	Identification of the long name associated with the collection or granule.	varchar2	Y
3	12	productLevel	Designates the product level values (i.e., 0, 1, 1A, 1B, 2, 3, 4).	varchar2	A
4	500	discipline	Designates the scientific area of application (i.e., Ocean, Atmosphere, Land, Cryosphere, Volcanic, Solar, Raw data, Radiance).	varchar2	A
5	80	processingCenter	Data center where this product was generated.	varchar2	Y
6	50	archiveCenter	Data center where the data product is archived. This value will be validated against the Provider portion in the Root File Name. See the Naming Convention sections above.	varchar2	Y
7	80	mission	An operation intended to carry out specific program objectives that use space and/or ground based measurement systems to provide scientific measurements (i.e., platforms, satellites, field experiments, and aerial measurements, etc.). For a multi-mission product, list all missions separated by a semi-colon (;). The primary mission should be listed first. Each mission should have 1 or more instrument associated with it. If there are multiple missions and multiple instruments, then the relationships between the missions and instruments should be defined.	varchar2	Y
8	80	instrument	An integrated collection of hardware containing one or more sensors and associated controls designed to produce data on an environment. For a multi-instrument product from one mission, list all instruments separated by a comma (.). If the product is a combined product from multi-missions involving multiple instruments, a group of the instruments from each mission should be separated by a semi-colon (;). The order of instruments should be in the same sequence as the mission field. If not applicable, enter: "N/A". (NOTE: the number of missions entered must pair evenly to number of instruments delimited by ";" i.e., if two missions entered: "mission1;mission2" then at least two instruments: "instrument1;instrument2" or "N/A;N/A" or "instrument1a,instrument1b; "instrument2a,instrument2b" etc.)	varchar2	Y
9	1	eosFlag	Flag to indicate whether the data product is an EOS or Non-EOS product. Values: E for EOS and N for Non-EOS.	char	Y
10	1	productFlag	Flag to indicate the type of product. Values: 1 = Data Product file, 2 = Instrument Ancillary file, 3 = System/Spacecraft file and 4 = External data. For non-ECS product, use the value 1.	char	Y

Note: 'Y' indicates mandatory data. 'A' indicates this information is required if available.

Ancillary Data: Data other than instrument data that is required to perform an instrument's data processing. They include orbit data, attitude data, time information, spacecraft or platform engineering data, calibration data, data quality information, and data from other instruments.

## 4.3 EMS Page Tags

Page tags allow the EMS to create a log of user interactions with Web resources in real time and provide a uniform collection mechanism across all domains and Data Providers. This method alleviates the need for Data Providers to send Hypertext Transfer Protocol (HTTP) access logs to the EMS. Page tags are implemented by inserting the same JavaScript into every HTML page to be tracked. There is a complementary ‘noscript’ tag which captures data from users with JavaScript disabled browsers. Page tags allow the EMS to sessionize user interactions across domains, so that a system-wide picture of Web usage can be obtained.

### 4.3.1 Page Tags Interface Mechanism

The required ESDIS metrics on Web usage will be derived from page tags. Therefore, Data Providers must install the EMS NetInsight® page tags on all Web pages they would like tracked. Page tags are implemented by adding a JavaScript file to the Web server(s) and inserting a page tag onto each Web page. Insertion of the tag on each Web page can be automated using server side includes, using your content management system, by doing recursive search and replace or by using Web server plug-ins.

Data are collected by “tagging” Web pages with JavaScript that causes a visitor’s browser to send information about the visitor’s visit to the EMS page tag server (<http://ws1.ems.eosdis.nasa.gov>) each time a Web page is loaded. This information, which appears in the page tag server log as a request for the page tagging image, can then be imported by the EMS for analysis. Page tags can be implemented on both development and operational servers. The EMS can filter results to allow only operations, development or both environments to be displayed. Data Providers can use this functionality to determine if the tags are working properly in development environments before promoting them to operations.

Optionally, NetInsight® page tags should be used to track on-page events, such as clicking on a link, changing a field on a form, submitting a form, and selecting an option in a drop-down list box as well. This type of page tagging is called event tagging and requires additional configuration options to be specified, and it is highly recommended. Additionally, the EMS can capture form submissions using the POST and GET methods. This tagging allows Data Providers to capture the usage of special services, but takes additional configuration and development. These configuration instructions are contained in a supplementary document, “The EMS Page Tag Implementation Guide”, located at <http://ems.eos.nasa.gov/documents>.

The EMS NetInsight® page tagging script is bundled in a ZIP file named ntpagetag.zip. The file can be downloaded from <http://ems.eos.nasa.gov/downloads>.

The ntpagetag.zip file contains these files:

- ntpagetag.js — JavaScript file that enables the EMS NetInsight® page tagging on a Web site and is stored on the Data Provider’s Web server(s).
- sample.html — file that contains sample page tags, including event tags.

- HTML\_tag\_insert.txt – Basic page tag example that can be inserted onto Web pages to be tracked.

Figure 4.3.1-1 provides an example of a basic page tag. The content of this box must be inserted into every Web page the Data Provider would like tracked using the EMS. This information will be provided to each Data Provider by the EMS.

### Example of Basic Page Tag

```
<!-- BEGIN: EMS Page Tag Notes -->
<!-- Some content Copyrighted 2004 Sane Solutions, LLC. All rights reserved. -->
<!-- Modified 02/07/2005 kjm -->
<!--Basic page tag script example-->
<!-- This information can be placed anywhere within your page before the closing BODY tag -->
<!-- YOU MUST CHANGE src="/scripts/ntpametag.js" to comply with your local implementation -->
<!-- END: EMS Page Tag Notes -->

<!-- BEGIN: EMS Page Tag -->
<script language="JavaScript">
// var NTPT_PGEXTRA = ";
// var NTPT_PGREFTOP = true;
// var NTPT_NOINITIALTAG = false;
</script>
<script language="JavaScript" src="/scripts/ntpametag.js"></script>
<!-- END: EMS Page Tag -->

<!-- BEGIN: EMS Page Tag No Script -->
<noscript>

</noscript>
<!-- END: EMS Page Tag No Script -->
```

**Figure 4.3.1-1. JavaScript – Basic Page Tag Example**

## 4.3.2 Page Tag Manifest File Format

In order for the EMS to capture and report on information collected from page tags, Data Providers need to complete the Page Tag Manifest File (Excel spreadsheet), as described in Table 4.3.1-1. The manifest has two sections. In the first section, the Data Providers identify the domain names which have been tagged with the EMS page tags. The second section is used to capture optional information and is only shown in the table as an example. This section is used to

capture information in query strings embedded in the URL from Web applications using POST and GET methods. An embedded URL is shown below:

http://<domain name>/<Program Name>?<Parameter1>=<value1>&<Parameter 2>=<value2>&<ParameterN>=<valueN>

For this optional component, the Data Provider must provide the name and a brief description in the Page Tag Manifest File for each parameter. A template of the Page Tag Manifest File is available from <http://ems.eos.nasa.gov/downloads>. Refer to “The EMS Page Tag Implementation Guide”, located at <http://ems.eos.nasa.gov/documents>, for more detail instructions.

**Table 4.3.1-1. Page Tag Manifest File**

<b>Page Tag Manifest</b>			
*Please replace the example entries below with information specific to your environment.			
<b>Domain Name Identification Section</b>			
<b>INSTRUCTIONS:</b> -Use this section to identify all domains where page tags are installed. -Metrics will only be available for domain names listed below. -Please include development and testing domains. They can be removed later. -Insert as many rows as necessary.			
Domain Name	Description		
<a href="http://www.foo.com">www.foo.com</a>	Main Site		
mirror.foo.com	Main Site mirror		
<a href="http://dev.foo.net">dev.foo.net</a>	Development site		
----- OPTIONAL -----			
<b>Web Application and Parameter Identification Section</b>			
<b>Instructions:</b> -Use this section to identify all web application URLs that should be tracked by the EMS. -The fully qualified URL is needed including the web application name. -For each parameter in the request (POST or GET), fill in the corresponding parameter section and provide a short description. -Specify the method (POST or GET). -Parameter entries are case sensitive and must be entered exactly as they appear in the URL. -If parameters are not included on this manifest, they will not be available for reporting. -Insert as many rows as necessary.			
Domain and Program Name (URL)	<a href="http://www.foo.com/cgi-bin/myApp.pl">http://www.foo.com/cgi-bin/myApp.pl</a>		
Description	Dynamically creates plots of optical thickness from a variety of atmospheric retrievals		
	Parameter	Method	Description
Parameters 1-n	Param 1	POST	...
	Param 2	GET	...
	Param N	POST	...

### 4.3.3 Page Tag Manifest File Naming Convention

The name of the Page Tag Manifest File must be in the following format for all Data Providers:

<YYYYMMDD>\_ <Provider>\_ PageTagManifest.xls

where

YYYY                      designates the 4 digit year for the time the Page Tag Manifest File was created

MM                         designates the 2 digit month, 01 through 12

DD\_                        designates the 2 digit day, 01 through 31, followed by an underscore

Provider\_                 designates the provider of the data, mutually agreed upon acronym defined in the OA, followed by an underscore

PageTagManifest.xls     indicates that the file is the Page Tag Manifest File worksheet

### 4.3.4 Page Tags Updates

Data Providers are expected to change Web servers, Web pages, applications and configurations over time. The only requirement Data Providers must adhere to during upgrades and migrations is to ensure page tags are carried over from one system to the next or are added to new systems. Because of the generic nature of basic page tag configurations, the only element that needs to be changed when migrating will be references to the JavaScript source to reflect the new operating environment. If Web application or event page tags are used, they may have to be modified for the new environment per instructions in previous sections and associated documentation.

The EMS must be notified in the event of changes in configuration of domains serving page tags or changes to parameters captured by page tags. Each time a change in configuration occurs, Data Providers should notify the EMS Team two weeks prior to migration so accounts can be modified and metrics can continue to be captured. To facilitate this process, Data Providers must provide the EMS with two manifests. The first file, Page Tag Manifest, will contain the new configuration including carryovers from the previous manifest. The second file, Invalid Page Tag Manifest, will contain domains and parameters that should no longer be captured by the EMS. A template of this document is available from <http://ems.eos.nasa.gov/downloads>.

The name of the Invalid Page Tag Manifest file must be in the following format for all Data Providers:

<YYYYMMDD>\_ < Provider>\_ InvalidPageTagManifest.xls

where

YYYY                      designates the 4 digit year for the time the Invalid Page Tag Manifest File was created

MM                         designates the 2 digit month, 01 through 12

DD\_                        designates the 2 digit day, 01 through 31, followed by an underscore

Provider\_ designates the provider of the data, mutually agreed upon acronym defined in the Operations Agreements between the EMS and the Data Providers, followed by an underscore.

InvalidPageTagManifest.xls indicates that the file is the Invalid Page Tag Manifest worksheet

## **4.4 EMS Error Recovery and Notifications**

The EMS is designed to handle network outages, hardware failure, and other sources of downtime (Note: The planned downtime is outlined in the EMS SLA document). The error types and recovery options are presented in sections 4.4.1 and 4.4.2 respectfully. The EMS implements failover servers to maximize hardware uptime and is located on the EMSn to minimize network downtime. Furthermore, the transfer method (Rsync wrapped with ssh) allows for graceful handling of system and/or network outages. Rsync accomplishes this by comparing the source and destination directory content and only sending files that are new or have been modified to the EMS servers. Thus, if Data Providers continue to produce data files per the requirements, once a connection to the EMS can be maintained, all necessary data will be transmitted.

If the Data Providers opt out of the preferred transfer method (Rsync wrapped with ssh), they are responsible for ensuring the necessary files are transmitted to the EMS.

In the event of a missing file or if the EMS has a failure ingesting data, the EMS will deliver an error notification email to the Data Provider. The content of the emails is presented in section 4.4.3. The EMS team will also e-mail notifications/requests in situations where the product is new and has not been defined and for known or unresolved request paths.

### **4.4.1 EMS Unavailable or Network Outage**

In the event of a multi-day EMS outage, Data Providers should continue creating and trying to transmit data to the EMS. Files for each day should be created and an attempt to send these files to EMS should be made (multiple attempts per day are also fine). After the EMS has recovered, it will be prepared to receive all Data Provider data files and process them. Rsync will make sure that the appropriate files are delivered to the EMS.

### **4.4.2 Data Provider Unavailable**

In the event a Data Provider experiences an outage or prolonged downtime, it is the Data Provider responsibility to recover from this situation, including populating data files and transmitting them to the EMS in the format specified in the ICD. However, if the Data Provider is experiencing prolonged downtime that impacts the distribution, archive, ingest or other systems it is unlikely that metrics for these systems would be available. In that case those metrics would not be required.

The wide variety of configurations and operating environments make it impossible for the EMS to recommend the best approach for recovering from Data Providers downtime. However, one suggestion to gracefully and automatically recover from these situations would be to create a file tracking table or flat file. When Data Providers initiate the process for creating and sending data

to the EMS this table or file can be used to determine the last successful export. Then using the current system time and the last successful export date, files between these two dates can be created dynamically.

#### 4.4.3 Error Email Notifications

The EMS Data Flat File Processing Error emails will include the following:

1. Date
2. Provider
3. File Type
4. Filename
5. Number of records failed
6. Total number of records
7. HTML link to the directory where the error file exists (ASCII Format). The error file will contain:
  - a) Date/Time
  - b) Provider
  - c) Filename
  - d) Header (field names)
  - e) Data Record Received
  - f) Error Message (s) – See Appendix A

Figure 4.4-1 provides a sample EMS error email notification. Figure 4.4-2 presents a sample error file created by the EMS.

From: [ems-qa-bounces@lists.nasa.gov](mailto:ems-qa-bounces@lists.nasa.gov)  
[<mailto:ems-qa-bounces@lists.nasa.gov>] On Behalf Of OPS@ems.eos.nasa.gov  
Sent: Friday, June 27, 2008 8:42 AM  
To: data provider (primary contact email address); (secondary email address)  
Cc: [ems-qa@lists.nasa.gov](mailto:ems-qa@lists.nasa.gov)  
Subject: [ems-qa] EMS Process Error - <Filetype>

Report created on <MON-DD-YYYY HH:MI AM/PM>

Environment: <Env>  
Provider: <Provider>  
Filetype: <Filetype>  
Filename: <Filename>  
Number of Failed Records: <n>  
Total Number of Records: <n>

Please visit

<https://ems.eos.nasa.gov/emslogs/error/<filename>.txt>

to review details about the records the EMS was not able to ingest.

For your convenience, the EMS has compiled all bad records into a single file. You can download this file from the following location: <https://ems.eos.nasa.gov/emslogs/error/<filename>.flt.bad>

If these records do not need to be ingested by the EMS, please reply to this email [ems-help@lists.nasa.gov](mailto:ems-help@lists.nasa.gov) stating no further action is needed.

If these records need to be reprocessed, please correct the errors identified on the page above, update the file name according to the ICD naming conventions, and reply to this email stating the file will be resent.

Thanks,  
EMS Team

Please note: Due to a known issue with Oracle's email functionality, the url above may have missing periods. If you are unable to retrieve your file(s), please check the url or contact [ems-help@lists.nasa.gov](mailto:ems-help@lists.nasa.gov)

#### **Figure 4.4-1. EMS Processing Error – Email Notification Example**

```

Add screenshots for notification
EMS Execution Date/Time: 12-JUN-06 10.10.36.000000 AM
Provider: NATTEST
FileName: 20060531_NATTEST_searchExp_File.flt

Header:
product|searchTerms|dataSource

Data Record Received:
|mySearchTerm2|mySysName2

Error Message:
product field is empty

-----
Data Record Received:
MOBCTP4irkeifdh4hfu4j3uidj4jdfu3jddj3jddj4jfhwkiqaksxcnwkj2341298uysfsln23
4kjsnvskjcvh98234kjnsdfkjnsckj234iuh|mySearchTerm4|mySysName4

Error Message:
product field length is greater than 80

-----
Data Record Received:
MYBCIR_E3||mySysName5

Error Message:
searchTerm field is empty

-----
Data Record Received:
||

Error Message:
product field is empty
searchTerm field is empty
dataSource field is empty

-----
Data Record Received:
MYBCTT_D3|mySearchTerm6|

Error Message:
dataSource field is empty

-----

```

**Figure 4.4-2. EMS Processing Error – ASCII Error File Example**

The File(s) Not Received Error email will include the following:

1. Date
2. Provider
3. File Type
4. Number of received files
5. Number of expected files
6. Names of missing files

Figure 4.4-3 displays the EMS Processing Error – Missing Files Email example.

From: OPS@ems.eos.nasa.gov  
To: data provider (Primary Contact Email Address) (Secondary Contact Email Address)  
Cc: ems-qa@lists.nasa.gov  
Subject: [ems-qa] EMS(env) – (data provider) Missing Files Report

---

(data provider) Missing Files Report as of (MON-DD-YYYY) 8:00 AM.

The following daily files have not been received:  
(file list)

Please notify EMS at [ems-help@lists.nasa.gov](mailto:ems-help@lists.nasa.gov) when the files listed above have been sent.  
Thank you,  
EMS Team

**Figure 4.4-3. EMS Processing Error – Missing Files Email Notification Example**

The EMS Unknown Search Terms email will include the following:

1. Date
2. Environment
3. List of files

Figure 4.4-4 displays the EMS Unknown Search Terms Email example.

From: OPS@ems.eos.nasa.gov  
To: data provider (Primary Contact Email Address); (Secondary Contact Email Address)  
Cc: ems-qa@lists.nasa.gov  
Subject: [ems-qa] EMS Unknown Search Terms

---

Report created on <MON-DD-YYYY HH:MI AM/PM>

Environment: <env>  
All Filenames with Unknown Searchterms

<filenames>

For a detailed report, please visit the following link:  
[https://ems.eos.nasa.gov/emslogs/error/<YYYYMMDD>\\_<provider>MissingSearchTerms.csv](https://ems.eos.nasa.gov/emslogs/error/<YYYYMMDD>_<provider>MissingSearchTerms.csv)

To resolve, please send EMS an updated searchExp flat file, and resend the distribution files listed above, appending the filename with the next revision, (.flt.rev#).

For a copy of the most recent searchExp flat file EMS has processed, please visit:  
[https://ems.eos.nasa.gov/emslogs/error/<YYYYMMDD>\\_<provider>\\_searchExp\\_<datasource>.flt<rev> if applicable](https://ems.eos.nasa.gov/emslogs/error/<YYYYMMDD>_<provider>_searchExp_<datasource>.flt<rev> if applicable)

For assistance with this notification or any other questions, please contact [ems-help@lists.nasa.gov](mailto:ems-help@lists.nasa.gov)

**Figure 4.4-4. EMS Unknown Search Terms Email Notification Example**

## 5. EMS Integrated Testing

---

Data Providers must contact ([ems-help@lists.nasa.gov](mailto:ems-help@lists.nasa.gov)) each time there is a modification to systems or configurations that will impact the EMS ability to capture metrics. To ensure all data streams (data files and page tags) will continue to function in the new operating environment the Data Provider and the EMS Team will follow the initialization and testing steps outlined in the OA. Briefly these steps are:

- The Data Provider will modify code at their location and send the resulting output to the EMS.
- The EMS will ingest the new data and make it available from the standard EMS user interface.
- This process will continue until all issues are resolved and the EMS Team and the Data Provider POC agree metrics are being captured correctly.

This page intentionally left blank.

## Appendix A. EMS Error Messages

Filetype: Ing		
Field	Requirement	Error Message
dbID	ECS Required field =Y Non-ECS Required field =Y Max field length =20 Datatype (varchar2)	“dbID field is empty” “dbID field is empty” “dbID field length is greater than 20” n/a
product	ECS Required field =Y Non-ECS Required field =Y Max field length =80 Datatype =varchar2	“product field is empty” “product field is empty” “product field length is greater than 80” n/a
productVolume	ECS Required field =Y  Non-ECS Required field =Y Max field length =15 Datatype =number	“productVolume field is empty or is not a numeric value” “productVolume field is empty or is not a numeric value” n/a “productVolume field is empty” “productVolume field is not numeric or is a negative value”
productState	ECS Required field =Y Non-ECS Required field =Y Max field length =20 Datatype =varchar2	“productState field is empty” “productState field is empty” “productState field length is greater than 20” n/a
externalDataProvider	ECS Required field =Y Non-ECS Required field =Y Max field length =20 Datatype =varchar2	“externalDataProvider field is empty” “externalDataProvider field is empty” “externalDataProvider field length is greater than 20” n/a
processingStartDateTime	ECS Required field =Y  Non-ECS Required field =Y Max field length =25 Datatype =timestamp	“processingStartDateTime field is empty or is not a valid date” “processingStartDateTime field is empty or is not a valid date” n/a “processingStartDateTime field is empty or is not a valid date” “processingStartDateTime field is greater than the current date”
processingEndDateTime	ECS Required field =Y Non-ECS Required field =Y Max field length =25 Datatype =timestamp	“processingEndDateTime field is empty or is not a valid date” “processingEndDateTime field is empty or is not a valid date” n/a “processingEndDateTime field is empty or is not a valid date” “processingEndDateTime field is greater than the current date” “processingStartDateTime field is greater than the processingEndDateTime field
timeToArchive	ECS Required field =A Non-ECS Required field =A Max field length =15 Datatype =number	n/a n/a n/a If provided: “timeToArchive field is not numeric or is a negative value”
timeToPreprocess	ECS Required field =A Non-ECS Required field =A Max field length =15 Datatype =number	n/a n/a n/a If provided: “timeToPreprocess field is not numeric or is a negative value”
timeToXfer	ECS Required field =A Non-ECS Required field =A Max field length =15 Datatype =number	n/a n/a n/a If provided: “timeToXfer field is not numeric or is a negative value”

Filetype: Arch		
Field	Requirement	Error Message
dbID	ECS Required field =Y Non-ECS Required field =Y Max field length =20 Datatype =varchar2	“dbID field is empty” “dbID field is empty” “dbID field length is greater than 20” n/a
product	ECS Required field =Y Non-ECS Required field =Y Max field length =80 Datatype =varchar2	“product field is empty” “product field is empty” “product field length is greater than 80” n/a
productVolume	ECS Required field =Y Non-ECS Required field =Y Max field length =15 Datatype =number	“productVolume field is empty or is not a numeric value” “productVolume field is empty or is not a numeric value” n/a “productVolume field is empty” “productVolume field is not numeric or is a negative value”
totalFiles	ECS Required field =Y Non-ECS Required field =A Max field length =15 Datatype =number	n/a “totalFiles field is empty or is not a numeric value” n/a “totalFiles field is empty” “totalFiles field is not numeric or is a negative value”
insertTime	ECS Required field =Y Non-ECS Required field =Y Max field length =25 Datatype =timestamp	“insertTime field is empty or is not a valid date” “insertTime field is empty or is not a valid date” n/a “insertTime field is empty or is not a valid date” “insertTime field is greater than the current date”
beginningDateTime	ECS Required field =Y Non-ECS Required field =A Max field length =25 Datatype =timestamp	If provided “beginningDateTime field is not a valid date” If provided “beginningDateTime field is not a valid date” n/a If provided “beginningDateTime field is not a valid date” “beginningDateTime field is greater than the current date”
endingDateTime	ECS Required field =A Non-ECS Required field =A Max field length =25 Datatype =timestamp	If provided: “endingDateTime field is not a valid date” If provided: “endingDateTime field is not a valid date” n/a If provided: “endingDateTime field is not a valid date” If provided: “endingDateTime field is greater than the current date” If provided: “beginningDateTime field is greater than the endingDateTime field”
productionDateTime	ECS Required field =A Non-ECS Required field =A Max field length =25 Datatype =timestamp	If provided: “productionDateTime field is not a valid date” If provided: “productionDateTime field is not a valid date” n/a If provided: “productionDateTime field is not a valid date” If provided: “productionDateTime field is greater than the current date”
localGranuleID	ECS Required field =Y Non-ECS Required field =A Max field length =100 Datatype =varchar2	“localGranuleID field is empty” n/a “localGranuleID field length is greater than 100” n/a
versionID	ECS Required field =A Non-ECS Required field =A Max field length =15 Datatype =varchar2	If provided: “versionID field length is greater than 15” n/a If provided: “versionID field length is greater than 15” n/a
deleteFromArchive	ECS Required field =Y Non-ECS Required field =A Max field length =1 Datatype =char	“deleteFromArchive field is empty” n/a “deleteFromArchive field length is greater than 1” n/a

Filetype: Arch		
Field	Requirement	Error Message
deleteEffectiveDate	ECS Required field =A Non-ECS Required field =A Max field length =25 Datatype =timestamp	If provided: "deleteEffectiveDate field is not a valid date" n/a n/a If provided: "deleteEffectiveDate field is not a valid date"
lastUpdate	ECS Required field =A Non-ECS Required field =A Max field length =25 Datatype =timestamp	If provided: "lastUpdate field is not a valid date" n/a n/a If provided: "lastUpdate field is not a valid date"

Filetype: UsrProf		
Field	Requirement	Error Message
userID	ECS Required field =Y Non-ECS Required field =Y Max field length =80 Datatype =varchar2	"userID field is empty" "userID field is empty" "userID field length is greater than 80" n/a
userType	ECS Required field =Y Non-ECS Required field =Y Max field length =1  Datatype =char	"userType field is empty" "userType field is empty" "userType field length is greater than 1" "userType field must be a numeric value between 1 and 4" n/a
emailOrIP	ECS Required field =Y Non-ECS Required field =Y Max field length =255 Datatype =varchar2	"emailOrIP field is empty" "emailOrIP field is empty" "emailOrIP field length is greater than 255" n/a
affiliation	ECS Required field =Y Non-ECS Required field =Y Max field length =20 Datatype =varchar2	"affiliation field is empty" "affiliation field is empty" "affiliation field length is greater than 20" n/a
dpUserType	ECS Required field =A Non-ECS Required field =A Max field length =50 Datatype =varchar2	n/a n/a "dpUserType field length is greater than 50" n/a
dpUserTypeDef	ECS Required field =A Non-ECS Required field =A Max field length =150 Datatype =varchar2	n/a n/a "dpUserTypeDef field length is greater than 50" n/a

Filetype: searchExp		
Field	Requirement	Error Message
product	EMS Required field =Y Max field length =80 Datatype =varchar2	"product field is empty" "product field length is greater than 80" n/a
searchTerm	EMS Required field =Y Max field length =200 Datatype =varchar2	"searchTerm field is empty" "searchTerm field length is greater than 200" n/a
dataSource	EMS Required field =Y Max field length =50 Datatype =varchar2	"dataSource field is empty" "dataSource field length is greater than 50" n/a

Filetype: Meta		
Field	Requirement	Error Message
product	EMS Required field =Y Max field length =80 Datatype =varchar2	"product field is empty" "product field length is greater than 80" n/a
metaDataLongName	EMS Required field =Y Max field length =1024 Datatype =varchar2	"metadatalongname field is empty" "metadatalongname field length is greater than 1024" n/a
productLevel	EMS Required field =A Max field length =12 Datatype =varchar2	n/a If provided: "productLevel field length is greater than 12" n/a
discipline	EMS Required field =A Max field length =500 Datatype =varchar2	n/a If provided: "discipline field length is greater than 500" n/a
processingCenter	EMS Required field =Y Max field length =80 Datatype =varchar2	"processingCenter fields is empty" "processingCenter field length is greater than 80" n/a
archiveCenter	EMS Required field =Y Max field length =50 Datatype =varchar2	"archiveCenter fields is empty" "archiveCenter field length is greater than 50" n/a
mission	EMS Required field =Y Max field length =80 Datatype =varchar2	"mission field is empty" "mission field length is greater than 80" n/a
instrument	EMS Required field =Y Max field length =80 Datatype =varchar2	"instrument field is empty" "instrument field length is greater than 80" n/a
eosFlag	EMS Required field =Y Max field length =1  Datatype (char)	"eosFlag field is empty" "eosFlag field length is greater than 1" "eosFlag field does not contain a value of E or N" n/a
productFlag	EMS Required field =Y Max field length =1  Datatype (char)	"productFlag field is empty" "productFlag field length is greater than 1" "productFlag field must be a value between 1 and 4" n/a

## Abbreviations and Acronyms

---

ARP	Address Resolution Protocol
ASCII	American Standard Code for Information Interchange
CCB	Configuration Control Board
CCR	Configuration Change Request
CMO	Configuration Management Office
CDRL	Contract Data Requirements List
COSSL	Commercial or Open Source Server Logs
COTS	Commercial Off-The Shelf
DAAC	Distributed Active Archive Center
DCN	Document Change Notice
EBnet	EOSDIS Backbone Network
ECS	EOSDIS Core System
EDGRS	ESDIS Data Gathering and Reporting System
EMS	ESDIS Metrics System
EMSn	EOS Mission Support network
EOS	Earth Observing System
EOSDIS	EOS Data and Information System
ESDIS	Earth Science Data and Information System
ESDT	Earth Science Data Type
FTP	File Transfer Protocol
GSFC	Goddard Space Flight Center
HTTP	Hypertext Transfer Protocol

ICD	Interface Control Document
ICMP	Internet Control Message Protocol
IP	Internet Protocol
ISO	International Standards Organization
LAN	Local Area Network
LATIS	Langley TRMM Information System
LGID	Local Granule Identifier
LLC	Logical Link Control
MAC	Media Access Control
MB	Megabyte
NASA	National Aeronautics and Space Administration
NCSA	National Center for Supercomputing Applications
OA	Operations Agreements
OP	Operations Procedures
OSI	Open Systems Interconnection
POC	Point of Contact
QA	Quality Assurance
RIP	Routing Information Protocol
SA	System Administrator
scp	Secure Copy
sftp	Secure File Transfer Protocol
SIPS	Science Investigator-led Processing System
SLA	Service Level Agreement
SMTP	Simple Mail Transport Protocol
SOO	Science Operations Office

SSH	Secure Shell
TCP	Transmission Control Protocol
TCP/IP	Transmission Control Protocol/Internet Protocol
URL	Universal Resource Locator

This page intentionally left blank.