Science Data Transfer Protocol (SDTP)

Status of this Memo

This memo describes a data transfer protocol convention for the NASA Earth Science Data Systems (ESDS) community. Distribution of this memo is unlimited.

Change Explanation

None.

Copyright Notice

Suggested Citation

Ho, E., & Voorhies, S. (2023). Science Data Transfer Protocol. NASA Earth Science Data and Information System Standards Coordination Office. https//doi.org/10.5067/DOC/ESCO/ESDS-RFC-045v1.

Abstract

This document establishes the Science Data Transfer Protocol (SDTP) as an ESDS standard interface mechanism. The SDTP is used for electronic transfer of data and related information. This RFC document provides a brief introduction, with the full SDTP specification detailed in Section 3 of the *Science Data Transfer Protocol (SDTP) Interface Control Document (ICD)* [1].

Table of Contents

1 Introduction

As File Transfer Protocol (FTP) and Secure FTP (SFTP) are being phased out, a new method for transferring files between Earth Science Data and Information System (ESDIS) elements is needed. The purpose of the Science Data Transfer Protocol (SDTP) is to provide an up-to-date mechanism for this data transfer using Hypertext Transfer Protocol (HTTP) over Transport Layer Security (TLS) (HTTPS) methods between a data provider and a data subscriber. For instance, this protocol would be used to transfer data from a Science Investigator-led Processing System (SIPS) to a Distributed Active Archive Center (DAAC).

1.1 Background

SDTP was developed to be used as a standard protocol for exchange of data between ESDIS elements. These elements may be in the NASA Earth Science Cloud, at a NASA data center (including an on-premises cloud), or at another science data provider or subscriber.

The primary SDTP users are ESDIS data providers and subscribers that have sustained data flows and currently use the Polling with Delivery Record (PDR) protocol, had planned to use the PDR protocol in the future, or expect to establish new sustained data flows. ESDIS data providers are typically science data producers, e.g. SIPS, Science Data System (SDS), Earth Observing System (EOS) Data and Operations System (EDOS). ESDIS data subscribers are primarily DAACs. Note that sometimes the roles are reversed with a DAAC providing input data to a science data producer. Also, some science data producers may subscribe to data produced by other producers (e.g. Visible Infrared Imaging Radiometer Suite (VIIRS) L1 (Level 1) and Land SIPS flow to the VIIRS Atmos. SIPS). This protocol could also be used for sustained data flows between DAACs and other (outside) data providers or subscribers.

1.2 Evidence of Implementation

At present, SDTP is already in use by a number of organizations, including:

- University of Wisconsin, Atmosphere SIPS
- Global Imagery Browse Services (GIBS)
- Goddard Earth Sciences Data and Information Services Center (GES DISC)
- Ice, Cloud, and land Elevation Satellite (ICESat-2 SIPS)
- Land Processes Distributed Active Archive Center (LP DAAC)
- MODAPS and Land SIPS
- Level-1 Atmosphere Archive and Distribution System (LAADS)
- National Snow and Ice Data Center (NSIDC)

Earth Observing System (EOS) Data and Operations System (EDOS) and NASA Sentinel Gateway (NSG) are also in the process of migrating to SDTP.

2 Overview of the SDTP System

2.1 Key Characteristics

The defining characteristics of the SDTP are as follows:

- File lists and files are pulled by the subscriber.
- SDTP is designed to work for any file format/type, e.g. science data, science metadata, browse imagery.
- A file list may contain many files to minimize the overhead of obtaining new file lists.
- Subscriber acknowledges each successful file transfer.
- File list is a JavaScript Object Notation (JSON) object.
- Required fields in the file list are minimized.
- A set of commonly used *tags* is defined.
- Additional *tags* can be added as needed for each provider/subscriber pair.
- X.509 certificates are used for authentication.
- Focused on transferring files between the provider and subscriber.
- Not focused on transferring metadata or other information about the file.
- Standard way to group related files.
- HTTPS responses are used to indicate status (no other status message responses are needed).
- Handling of non-common errors is done *out-of-band*, via email or other means.
- File size is only limited by the HTTPS protocol.

2.2 Prerequisites

Both the provider and subscriber of the SDTP agree on the following:

- Provider URL for the HTTPS file transfer.
- Certificate Authority for authentication.
- Subscriber certificate Distinguished Name.
- A set of *tags* and corresponding valid values.
- A set of *extra* fields and corresponding valid values.
- Values for a set of parameters that control the interface, such as, the maximum number of files in a list.
- Points of contact.

This information is documented in the Operations Agreement (OA) between the provider and subscriber.

Tags and their valid values are used to control the data that is transferred from a provider to a specific subscriber. For example, the OA may limit the value of the Earth Science Data Type (ESDT) *tag* to just the products produced by a SIPS (provider) that are meant to flow to a specific DAAC (subscriber).

2.3 Data Flow Overview

The interface only supports a pull mode of operation where the subscriber pulls the file lists and file contents from the provider.

The first step is for the subscriber and provider to use a certificate to establish a secure HTTPS connection. Then, to begin the file transfer, the subscriber requests a list of files from the provider using a set of *tags*. The subscriber then pulls all of the files in the list and acknowledges receipt of each file.

Information on obtaining certificates can be found on the SDTP wiki page.

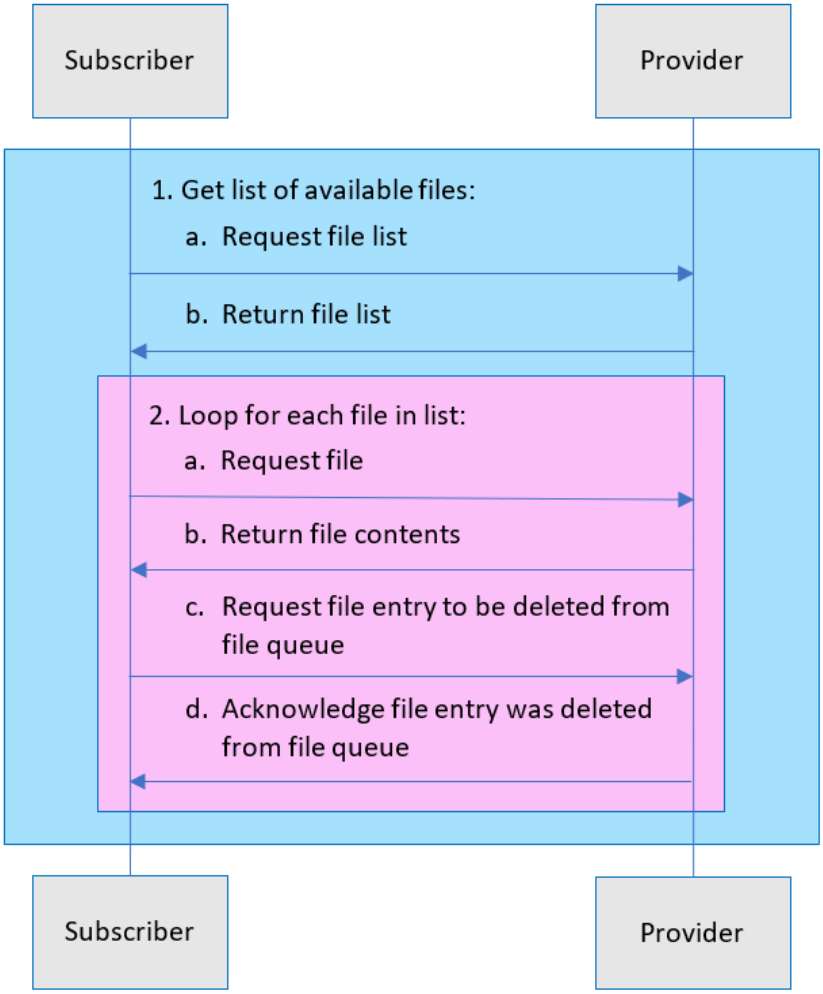

Figure 1 Data Flow

The provider stages a file to be transferred to a subscriber and assigns a unique *fileid* to the file (not shown). The provider adds the file to a queue of files to be transferred. Each file entry in the queue has a set of *tags* and values.

The individual steps that are involved in the data transfer (Fig. 1) are:

1. To obtain a list of available files:
    a. The subscriber performs a HTTPS *GET* with a set of *tags* and values.
    b. The provider returns a JSON object containing a list of files ready to be transferred. Only files for the subscriber that have a matching set of *tags* and values are included in the list.
2. For each file in the list:
    a. The subscriber requests the file contents by performing a HTTPS *GET* containing the *fileid*.
    b. The provider returns the file contents.
    c. The subscriber acknowledges the file transfer by performing a HTTPS *DELETE* containing the *fileid* of the file.
    d. The provider removes the file from the queue for the subscriber and returns a HTTPS *Success* status.

Most file transfer errors are handled using standard HTTPS error codes. Other errors, such as problems with the file list contents, are handled by email communication between the subscriber and provider.

3 Summary

SDTP is a common set of commands over HTTPS that meets all of NASA's security requirements for data transfer processes and has been developed to replace the PDR mechanism.

4 References

Normative References

[1] The SDTP ICD can be found on the Configuration Management EOSDIS Tool (COMET) at https://ops1-cm.ems.eosdis.nasa.gov/ by searching for document number 423-ICD-027. A login account is needed to access COMET.

[2] SDTP Wiki Page,

https://wiki.earthdata.nasa.gov/display/SDTP/SDTP+Home

[3] Ordering a Certificate,

https://wiki.earthdata.nasa.gov/display/SDTP/SDTP+Home#SDTPHome-OrderingaCertificate

[4] The *ICD between the EOSDIS Core System (ECS) and the Science Investigator-led Processing Systems (SIPS) Volume 0 Interface Mechanisms* can be found on COMET (https://ops1-cm.ems.eosdis.nasa.gov/) by searching for document number 423-41-57.

[5] The *ICD between EDOS and the ESDIS Science Data Segment (SDS) SIPS for the Joint Polar Satellite System (JPSS)* can be found on COMET (https://ops1-cm.ems.eosdis.nasa.gov/) by searching for document number 423-ICD-010.

Informative References

[6] Science Investigator-led Processing Systems (SIPS),

https://earthdata.nasa.gov/about/science-investigator-led-processing-systems

[7] EOSDIS Distributed Active Archive Centers (DAACs),

https://earthdata.nasa.gov/about/daacs

[8] SDTP reference implementation and documentation can be found at: https://gitlab.modaps.eosdis.nasa.gov/infrastructure/APS/containers/sdtp.

5 Authors' Addresses

ESDIS Standards Office staff

esco-staff@lists.nasa.gov

Appendix A

Glossary of acronyms

Acronym	Description
COMET:	Configuration Management EOSDIS Tool
DAAC:	Distributed Active Archive Center
ECS:	EOSDIS Core System
EDOS:	Earth Observing System (EOS) Data and Operations System
EOS:	Earth Observing System
EOSDIS:	Earth Observing System Data and Information System
ESDIS:	Earth Science Data and Information System
ESDS:	Earth Science Data Systems
ESDT:	Earth Science Data Type
FTP:	File Transfer Protocol
GES DISC:	Goddard Earth Sciences Data and Information Services Center
GIBS:	Global Imagery Browse Services
HTTP:	Hypertext Transfer Protocol

HTTPS:	Hypertext Transfer Protocol Secure
ICD:	Interface Control Document
ICESat-2:	Ice, Cloud, and land Elevation Satellite-2
JPSS:	Joint Polar Satellite System
JSON:	JavaScript Object Notation
L1:	Level 1
LAADS:	Level-1 Atmosphere Archive and Distribution System
LP DAAC:	Land Processes Distributed Active Archive Center
MODAPS:	MODIS Adaptive Processing System
MODIS:	Moderate Resolution Imaging Spectroradiometer
NSG:	NASA Sentinel Gateway
NSIDC:	National Snow and Ice Data Center
OA:	Operations Agreement
PDR:	Polling with Delivery Record
RFC:	Request for Comment
SDS:	Science Data Segment
SDTP:	Science Data Transfer Protocol
SFTP:	Secure File Transfer Protocol
SIPS:	Science Investigator-led Processing Systems
TLS:	Transport Layer Security
URL:	Uniform Resource Locator
VIIRS:	Visible Infrared Imaging Radiometer Suite