THE NPOESS PREPARATORY PROJECT SCIENCE DATA SEGMENT (SDS) DATA DEPOSITORY AND DISTRIBUTION ELEMENT (SD3E) SYSTEM ARCHITECTURE

Evelyn L. Ho, Robert J. Schweiss

NASA Goddard Space Flight Center, Code 586 and 581
Greenbelt, MD, USA
evelyn.l.ho@nasa.gov, robert.j.schweiss@nasa.gov

ABSTRACT

The National Polar-orbiting Operational Environmental Satellite System (NPOESS) Preparatory Project (NPP) Science Data Segment (SDS) will make daily data requests for approximately six terabytes of NPP science products for each of its six environmental assessment elements from the operational data providers. As a result, issues associated with duplicate data requests, data transfers of large volumes of diverse products, and data transfer failures raised concerns with respect to the network traffic and bandwidth consumption. The NPP SDS Data Depository and Distribution Element (SD3E) was developed to provide a mechanism for efficient data exchange, alleviate duplicate network traffic, and reduce operational costs.

Index Terms— National Polar-orbiting Operational Environmental Satellite System (NPOESS); NPOESS Preparatory Project (NPP) Science Data Segment; data depository and distribution; data exchange; data broker

1. INTRODUCTION

The NPP Science Data Segment (SDS) is intended to enable Climate Analysis Research Systems (CARS) development that will focus on the following areas: Atmosphere Composition, Climate Change, Carbon/Ecosystems, Solid Earth, Weather, and Water/Energy Cycle [1]. The SDS is composed of five Product Evaluation and Analysis Tool Elements (PEATEs), one for each of the following disciplines: Atmosphere, Land, Ocean, Ozone, Sounder, and more recently Earth Radiation. The other elements of the SDS include the NPP Instrument Calibration Science Element (NICSE), a Project Science Office Element (PSOE), an Integration and Test System Element (I&TSE), and the SDS Data Depository and Distribution Element (SD3E). This paper explains the architecture of the SD3E.

2. SDS SYSTEM OVERVIEW

The NPP SDS Data Depository and Distribution Element (SD3E) serves as the central data broker that temporarily stores data in a 32-day rolling cache for retrieval by the environmental data elements. Figure 1 provides a context diagram of the system and its external interfaces. These environmental elements, also known as the PEATEs, include the following disciplines: Atmosphere, Land, Ocean, Ozone, and Sounder [2]. The PEATEs are the primary recipients of xDRs, i.e., Raw Data Records (RDRs), Science Data Records (SDRs), Environmental Data Records (EDRs), & Temperature Data Records (TDRs), Intermediate Products (IPs), ancillary/auxiliary data, calibration products, algorithms, and associated source code. Additionally, the SD3E makes data available to the I&TSE to demonstrate algorithm and calibration enhancements, to diagnose science data quality, and to regenerate intermediate products, if necessary. The NICSE primarily receives the calibration look-up tables and software to assess and validate pre-launch and post-launch radiometric and geometric calibration and characterization of the Visible Infrared Imaging Radiometer Suite (VIIRS) instrument data. Finally, the PSOE provides overall management, coordination and science direction to the SDS elements for data evaluation and assessment.

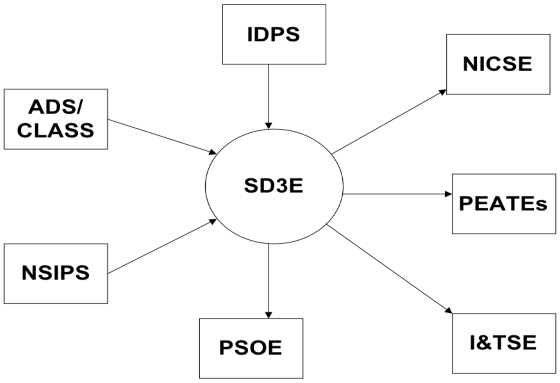

Figure 1. SD3E Context Dataflow Diagram

The three major data providers of NPP products include the National Environment Satellite, Data, and Information Service (NESDIS) Interface Data Processing Segment (IDPS), the Archive and Distribution Segment (ADS), more commonly known as the NOAA Comprehensive Large Array-data Stewardship System (CLASS) [3], and the NPP Science Investigator-led Processing System (NSIPS). The NESDIS IDPS is the primary provider for all Instrument RDRs. This comes to about 540 gigabyte (GB) per day. The NOAA ADS/CLASS nominally provides, nearly all instrument, SDR, EDRs, Delivered Intermediate Products (DIPs), and ancillary/auxiliary data. Additionally, the ADS/CLASS provides the calibration products and operational algorithms and source code. This comes to approximately 4,848 GB of data per day. Finally, the NSIPS will provide the Retained Intermediate Products (RIPs), for a daily total of approximately 1,590 GB.

3. SD3E SOFTWARE ARCHITECTURE

The major software components of the SD3E are developed using Perl toolkits and scripts consisting of about 22,200 Source Lines of Code (SLOC). The Web interface for data requests and Application Programming Interface (API) to the NESDIS IDPS uses Java, and consists of about 5,300 SLOC. The open source Apache HTTP Server is used as the Web server and PostgreSQL (also open source) is used for the database and consists of about 1,838 SLOC.

3.1. Software Reuse

The SD3E system adapted designs and software from the Moderate Resolution Imagining Spectrometer (MODIS) Data Processing System (MODAPS) and from the Ozone Monitoring Instrument (OMI) Data Processing System (OMIDAPS)[4]. Building on the experiences of MODAPS's distribution and archival methods, the SD3E adapted the concept of the job scheduler and the use of multiple central processing units (CPUs) to ingest and verify data products and to distribute tasks/processes to utilize resources from other computers efficiently. From the OMIDAPS, the SD3E reused database wrappers written for PostgreSQL and customized those wrapper functions for our use (e.g., database insert, update, delete, select).

3.2. File Transfer Protocol (FTP) Directory Structure

The three external data segments, IDPS, ADS/CLASS, and NSIPS, interface with the SDS PEATEs/NICSE via anonymous FTP with internet protocol (IP) restriction to push data products and requests to an inbound location specified by the SD3E. Items pushed to SD3E's inbound location can only be written and not read.

For data retrieval, the directory structure is organized into three different categories. All three structures house products for a maximum of 32 days. Once a file passes integrity checks, soft links are created to point to the products stored on disk. The three file hierarchies include the NPP_Products, NPP_Closed, and DailyIngest. The PEATEs/NICSE can use one or more of the three directories to anonymously FTP pull their desired data products. Each directory serves a different purpose. Both the NPP_Products and NPP_Closed directories are grouped by instrument data capture date and product type. The NPP_Products directory provides a means of traversing the tree for older products. The NPP_Closed, similar to the NPP_Products in structure, indicates to the PEATEs/NICSE that no more data is expected to arrive for the data day; the data day is complete. The DailyIngest directory groups the ingested products by Eastern Local Time and by element (PEATEs/NICSE). This directory provides a one-stop location for data retrieval of the most newly ingested products, regardless of the instrument data date.

3.3. SD3E Software

The SD3E software is divided into seven major components: the operator, the scheduler, the database, the interface controller, the ingest controller, the data storage, and the housekeeping functions. See Figure 2 for the SD3E Software Diagram.

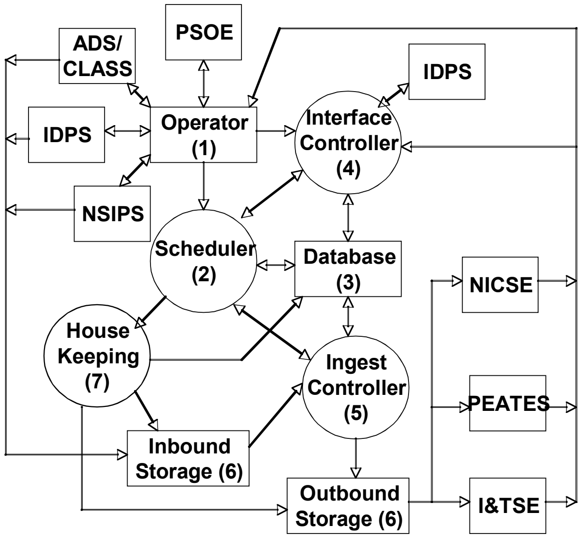

Figure 2. SD3E Software Diagram

1) The Operator: The operator, a person available five days a week, eight hours a day, serves as the primary interface between the SD3E and the external data providers and the PEATEs/NICSE. Simple graphical user interfaces, written in Perl and Java, provide the operator with tools to monitor the processes running and to troubleshoot system issues. Additionally, the operator manually generates and submits data requests to the IDPS, ADS/CLASS, and NSIPS for non-nominal products using respective Web interfaces.

2) The Scheduler: This persistent task, i.e., a daemon, orchestrates and coordinates all system activities. It controls the scheduling and execution of the Controller (Ingest and Interface) tasks. It assigns tasks to the appropriate host based on available resources. It also schedules tasks to be

executed at a scheduled time, such as removing files older than 32 days. Additionally, the Scheduler handles the cleanup of processes terminating under abnormal conditions and logs status and error messages.

3) The Database: The relational database, PostgreSQL, is the primary mechanism for data accounting, task coordination, task monitoring, task communication, and resource tracking. All of the system's components can connect and access the database. It provides information regarding the products requested, the requestor of a product, the integrity of the product, missing products, and status of the products. It also tracks information regarding system resources and queues tasks for the Scheduler to execute.

4) The Interface Controller: Two major components make up the interface controller – the interface to the external segments and the interface to the PEATEs/NICSE. The PEATEs/NICSE initiate events by submitting either a subscription (a standing order) or an ad-hoc request either using the SD3E Web interface or the machine-to-machine interface. The machine-to-machine interface, which encapsulates handshaking, provides Extensible Markup Language (XML) users the ability to automate their data request procedures. The PEATE/NICSE would create the XML request and anonymously FTP push the request to the SD3E inbound location. The SD3E polls the inbound request directory every 15 minutes (a configurable number). When a request becomes available, the interface controller validates the syntax of the request and updates the product definition table in the database. The second mechanism, the Web interface, provides a simplified, straightforward, graphical method for data ordering. The interface controller minimizes duplicate data product requests by coalescing the multiple requests. It records a data product's identifier, product type, and aggregation format in the product definition table by tracking each requester's request for a product. As a result, the product definition table summarizes all of the PEATEs/NICSE data requests for each product. The subscription is then generated and submitted to the appropriate data provider. This mechanism reduces the data request and transfer load and bandwidth usage from approximately 40.8 terabyte (TB) down to approximately 6.8 TB, a worst-case scenario. It also reduces the number of data requests/interfaces to the external data provider from six interfaces (the PEATEs/NICSE) to one (the SD3E). If the interface controller receives an ad-hoc request, it will determine the time the product was last retrieved and the product type. If the product is an RDR and has been less than 24 hours since initially ingested, it will submit a request to the IDPS. If more than 24 hours has passed since the initial receipt of the RDR, then the request is submitted to ADS/CLASS. All other requests for products are submitted to ADS/CLASS, regardless of when it was received by the SD3E.

5) The Ingest Controller: The ingest controller temporarily stages, ingests, verifies, and stores requested products. The products pushed to the SD3E incoming FTP directories from the IDPS, NSIPS, and the ADS/CLASS will be stored by default on the same disk as the SD3E uses to store the files. Additionally, in the event that there are disk problems or insufficient disk space, the soft links to the incoming directories for product delivery can be modified to point to a different disk. This change is transparent to the external data segments pushing data to the SD3E. Each external data segment will push products to an inbound location specified by the SD3E as a part of the subscription request. Every 15 minutes (a configurable number), the ingest controller checks for the availability of a data delivery notification (DN) from either the IDPS or ADS/CLASS. The DN indicates that the product has completed transfer. The task verifies the checksum (for products from IDPS and NSIPS) and the digital signature (for products from ADS/CLASS). Once validated, the file is stored and soft links are created to link the file from the disk to the NPP_Products directory under the appropriate instrument data date and product type. For the DailyIngest directory, the file is soft linked to the PEATE/NICSE who requested the product. The product location and status is updated in the appropriate database tables. If the ingest controller determines that there are missing files or data integrity failures, it notifies the interface controller by placing the anomalous product's information in the re-order database table for the interface controller to automatically re-order. To check for missing data products, the ingest controller expects a certain number of products for each product type, based on the time duration of the product, per orbit. If it finds that there are missing products, those products are flagged and reported on the Web site. If there are no missing products and the expected number of files have been received for the product type, the files are soft linked to the NPP_Closed directory. The ingest controller will track the number of times a product has been re-ordered. If a product has been reordered more than twice, then the product is flagged as invalid, the status is made available on the Web site, and this task will no longer request the interface controller to re-order the product.

6) The Data Storage: The data storage consists of two separate areas – an inbound and an outbound directory. Both directories use anonymous FTP for data pushes to the inbound and for data pulls from the outbound area. For security reasons, the inbound area is IP restricted. The outbound area provides a 32-day rolling data cache of xDR products and 7 days of IP storage. The five most recent versions of the algorithms and source codes, calibration products, and ancillary/auxiliary product are retained. The inbound area is sized to store approximately 3 days worth of data products.

7) The Housekeeping: The housekeeping tasks are scheduled for execution either on a daily basis or on an as-needed basis. These tasks perform cleanup of products older than the specified time window of either 32 days for xDRs and 7 days for IPs, database integrity checks, and data consistency checks between files on the physical disk to files recorded in the database. Additionally, a report generation capability is available to provide management with resource and data transfer volumes and statistics.

4. SD3E HARDWARE ARCHITECTURE

The current SD3E development hardware consists of two Dell PowerEdge 2850 servers running Mandriva Linux. Each of the servers includes two Intel Xeon 3.0 GHz processors, 4 GB of memory, and a 73 GB hard drive. A 12 TB file system is Network File System (NFS) mounted to the primary server.

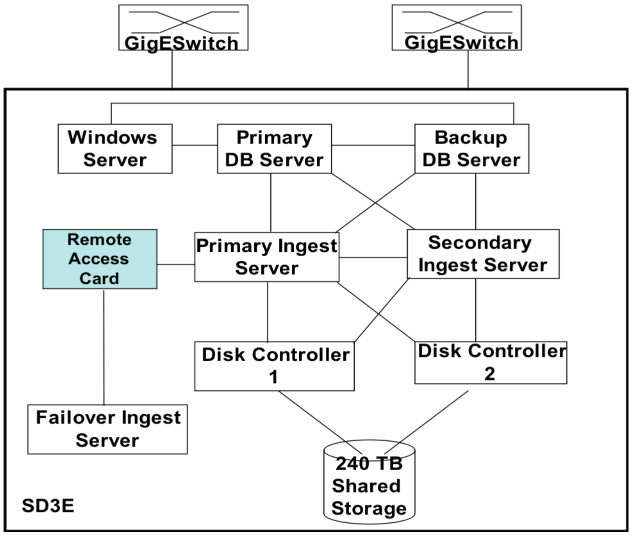

Figure 3. SD3E "At-Launch" Hardware Diagram

The "at-launch" configuration, shown in Figure 3, will be three Dell PowerEdge 2950 servers with dual processors. One of the three servers will serve as the primary ingest server, one will serve as the secondary ingest server, and the third server will provide system failover capabilities. The primary server will run all of the major SD3E software tasks. For example, it will run the FTP server, the Scheduler, the Ingest Controller, and the Interface Controller. The secondary server provides computing resources to the primary server. In the event that the primary server goes down, the failover server will automatically take over the job of the primary server, with minor modifications and configurations to the database, and continue processing. One of the Dell PowerEdge 2850 provides database failover capabilities. The SD3E uses Pgpool-II for database replication. If the primary database fails, the secondary database will assume responsibility as the primary database and continue the ingest procedures. This Dell PowerEdge 2850 will also remain as the SD3E development and test platform. A Dell PowerEdge R300 Windows 2003 Server will connect to the database and primary server. The Windows 2003 Server houses the Apache server for the Web interface and also will have network access (all necessary ports opened) to interface with the IDPS data ordering system using their Java APIs. The system includes 240 TB of shared storage and two disk controllers, from Data Direct Networks (DDN) Federal, to house the 32 days of data. Redundant Array of Independent Disks (RAID) 6 disks are utilized since they provide for high data fault tolerance and protects against multiple storage drive failures. The 240 TB of storage includes a 10 percent margin.

5. CONCLUSION

The SD3E system design incorporates the use of open source software and is able to recognize areas where system automation is possible. This resulted in the reduction in operator intervention and operator time. Developing the system to be flexible and modular provides the capability to integrate new functionality with minimal effort, thus reducing future maintenance time and costs. The use of commodity hardware provides the flexibility to purchase what is needed with the possibility to expand (e.g., CPU, memory, disk) at a later time. Moreover, the development of the central interface substantially reduced the projected data transfer volume, alleviating some of the network bandwidth usage and traffic, and optimized the data ordering mechanism. The data ordering mechanism eliminates duplicate product requests and thus reduces the volume of data transfer.

ACKNOWLEDGMENT

We thank the members of the NPP SDS, SD3E, and PEATEs team members for their contributions to this work.

REFERENCES

[1] Asrar, Ghassem, Level I Requirements for the NPOESS Preparatory Project (NPP), Signed November 3, 2003
[2] Schweiss, R., Ho, E., Ullman, R., Samadi, S., The NPOESS Prepartory Project Science Data Segment: Brief Overview, IEEE Geoscience and Remote Sensing Society, 2006
[3] NOAA Satellite and Information Service, National Environment Satellte, Data, and Information Service (NESDIS) CLASS, Retrieved May 16, 2006, from http://www.class.ngdc.noaa.gov/saa/products/welcome
[4] Tilmes, C., Linda, M., Fleig, A., 2004, Development of two Science Investigator-led Processing Systems (SIPS) for NASA's Earth Observation System (EOS), IEEE Geoscience and Remote Sensing Society, 2004