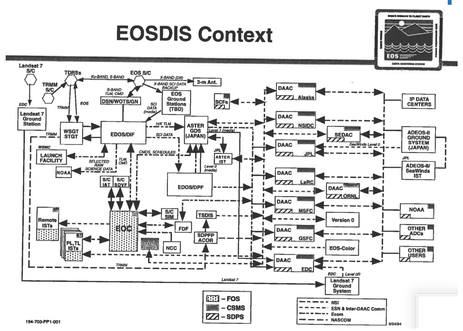
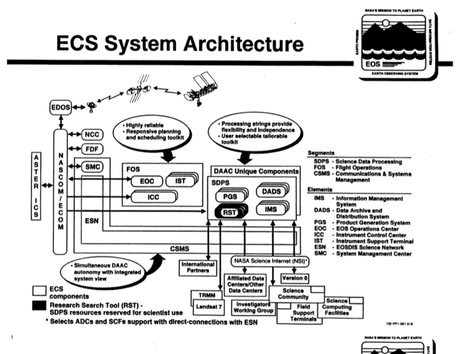
IN51F-0695
Fall AGU Meeting
December 9-13, 2019
San Francisco, CA

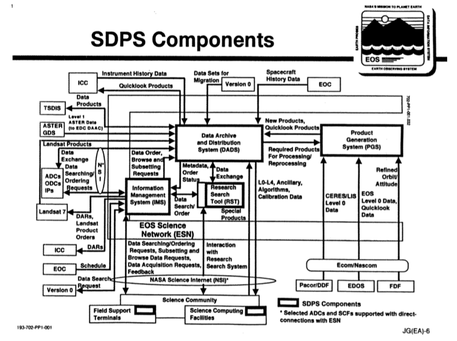
Past and Future Operations Concepts of NASA's Earth Science Data and Information System

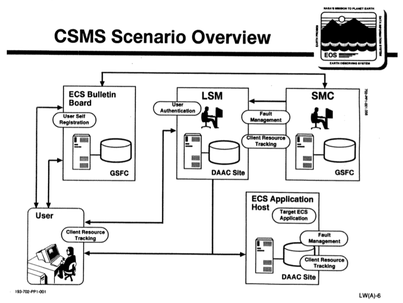
John F Moses[1], Francis Lindsay[1], Jeanne Behnke[1]
[1] NASA Goddard Space Flight Center, Greenbelt, MD, United States.

EOSDIS Architecture and Operations Concept 1990's









EOSDIS Core System defined 3 Segments

- **Flight Operations**: manages and controls EOS platforms and instruments.
- **Science Data Processing**: science data processing, archival and distribution elements and a data information management system (IMS). Distributed Active Archive Centers (DAACs) host the SDPS elements a distributed portion of the IMS. Toolkits & Application Programming Interfaces (APIs) for DAAC-unique extensions and integration into Science Computing Facilities.
- **Communication and System Management**: overall management and operation of ground system resources. Includes the System Management Center (SMC) and the EOSDIS Science Network (ESN) at GSFC. DAACs host a distributed portion of SMC and ESN.

Science Data Processing Segment

- **Data Archive and Distribution System:** receives and stores science data, ancillary data. Updates metadata with inventory and quality assurance. Receives orders from IMS, updates inventory, generates accounting and request status.
- **Product Generation System**: routine processing and reprocessing to standard products.
- **Information Management System**: provides user information on EOS data and documentation, accesses databases throughout ECS, processing status, science data products, and mission descriptions and schedules. Tools for the user to enter a data acquisition requests (DAR); the IMS sends the DAR to the EOC for scheduling and tracks status.
- **Research Search Tool**: accesses portion of ECS resources for new search, filtering, subsetting, special product generation. User interfaces with ECS via APIs.

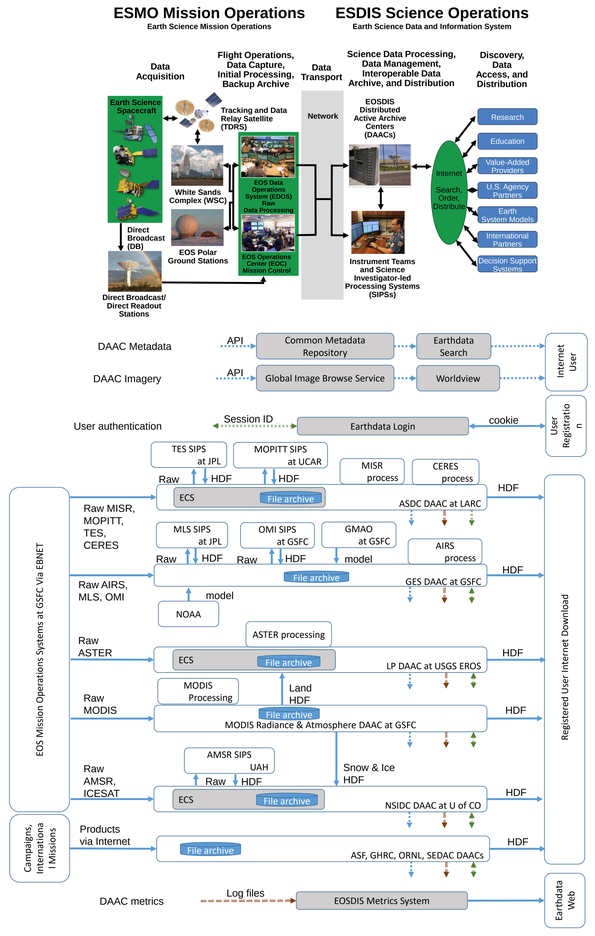
DAAC operations

- Each DAAC performs data management and processing for a specific set of Earth science data; data product generation, on-demand processing, algorithm integration, data archiving and distribution. Standard products, browse, subsets, summaries are created during routine and on-demand processing in response to user requests.
- DAACs control data quality, integration and testing of algorithms, storage, software support and provision of user access to data and related information. DAACs provide science consulting support and software tools to EOS users to allow them to perform algorithm updates, quality assessment and test product reviews. DAAC operations are 24 hours x 7 days a week.
- Product generation, processing & reprocessing, data ingest and distribution are scheduled, monitored, controlled, reviewed, managed by the DAAC operations staff using ECS. The staff performs product distribution, archive maintenance, systems support and data backup and recovery. High level direction for science data processing provided by the NASA Science Operations Manager.
- Technical Assistance Group assists in providing science community access to and use of ECS. TAG assists users in registration, general inquiry, product requests, logistics and user accounting, product acquisition and tracking, and aids them in use and development of metadata, subsets of data, numerical methods and tools, vector and parallel processor techniques, visualization and graphic tools, analysis tools expert systems, data formats and computing techniques. TAG also support the scientist in testing and integrating their product generation algorithms into the PGS.
- System Maintenance and Operations management was to be established at each of the ECS sites to plan, budget, configure and control resources, finance, personnel, logistics, property, security, and training. M&O will maintain hardware, software, firmware, make system upgrades and perform sustaining engineering for system improvements.

Communication and System Management

Each ECS element manages configuration, billing and accounting, performance measurement and security, SMC interfaces with all elements for monitoring, coordination, guidelines, directives and control. SMC authorizes user and operator access. End-to-end fault, configuration, performance, and security data are shared between all segments. DAACs host a distributed portion.

EOSDIS As-Built



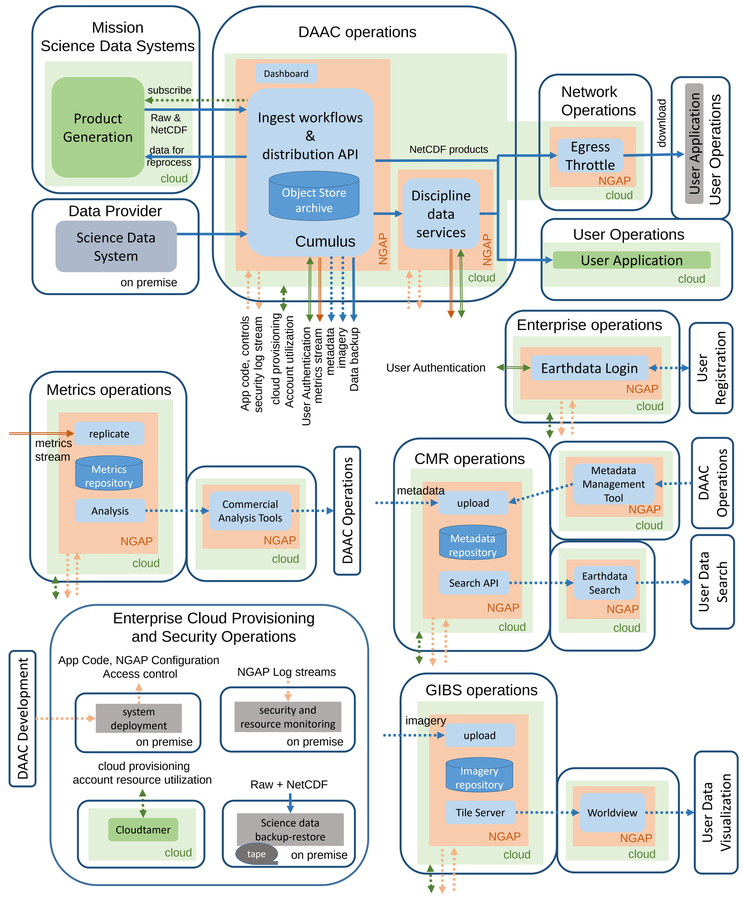
Earth Observing System Context

- Flight mission operations separated from Science Data Systems for EOS platform command and control, maneuvers, instrument planning & operations, spacecraft housekeeping, attitude/orbit determination, science data capture
  - Independent mission/instrument operations enabled by standardized flight products managed by interface control documents (ICD) and sent to science data systems via secure network
  - Coordination and planning for all instruments/platforms managed through unique interfaces with each mission science team. Instruments operate full duty cycle, ASTER is only instrument supporting prioritized science research acquisition requests.

EOS Data and Information System separated elements

- Archive and distribution systems developed and operated by DAACs
  - EOSDIS Core System (**ECS**) designed by NASA for selected complex/high volume instruments; implementation limited to three DAACs.
  - Each DAAC built and operates custom software systems and data services for discipline-unique collections, ancillary data, campaigns and missions, other agency and international satellite data
  - Each DAAC and host institution selected system hardware, provided network and security
- **Science Investigator-led processing systems (SIPS)**: separated from ECS, evolved from PGS.
  - Custom designed to efficiently integrate science algorithms and process specific instrument data.
  - Each SIPS and host institution selects system hardware, provides network and security
  - Interfaces managed by ICD and use an EOSDIS-wide standardized scripted protocol
- Enterprise user registration, search, visualization, metrics and toolkit support
  - **ECS Toolkit**: Basic tools for integration of science algorithms into production systems, including a common product format (HDF) and physical parameter and earth location metadata
    - Widely adopted prior to launch and continued use by all EOS science algorithm developers
    - Enabled cross discipline search and interoperability among diverse instrument measurements
  - **Earthdata Search** and **Common Metadata Repository (CMR): Evolved from** Research Search Tool and IMS. Based on unified EOSDIS metadata model custom built and operated by NASA to achieve microsecond search response for internet users.
  - **Worldview** and **Global Image Browse Service (GIBS)** based on creating standard browse imagery from science data and custom build to achieve rapid display of high resolution imagery
  - **Earthdata Login**: custom configured using industry standard libraries for user email registration prior to data download. Evolved from SMC.
  - **EOSDIS Metrics System**: Custom configured commercial tools for collecting formatted DAAC system logs per ICD and for periodic analysis and annual reporting. Evolved from SMC.

EOSDIS Data Collections in-the-cloud Target Architecture



- **NASA General Application Platform (NGAP)**: provides a cloud-based platform with access and security monitoring for DAAC applications.
  - Deployed with applications code to operate on a commercial cloud environment.
  - Enterprise operations manage cloud accounts, resource utilization and access monitoring, security, ingress/egress, networks for application owners.
- **Cumulus**: an application framework for data ingest and distribution utilizing object-based storage archives. Configured for each data collection to perform acquisition, ingest, validation, backup, processing, metadata harvesting & creation, publication to CMR, Image creation for GIBS data distribution and metrics reporting. DAAC Operators use a dashboard to monitor multiple Cumulus automated workflows, discipline specific data services and cloud utilization and workflow status metrics.
  - **EOSDIS Product Delivery Record/Product Availability Notice (PDR/PAN)** legacy interface protocol used for acquisition of science data from on-premise facilities
  - **EOSDIS Cloud Notification Message** utilized for Cloud-to-Cloud interface for Mission Science Data Systems or Product Generation Systems. Application subscribes with vender service to receive data availability notices, and on receipt pulls data from object storage archive using cloud service.
  - Application Programming Interfaces (APIs) are utilized by Cumulus Ingest and Distribution workflows to interface with Enterprise services Earthdata Login, CMR, GIBS, Cloud Metrics and science data backup systems.
- **DAACs** develop data collection ingest workflows utilizing Cumulus and data service applications in development 'sandbox' accounts. Developers integrate and test custom changes to Cumulus through System Integration and Test accounts, and compete end-user testing through User Acceptance Test accounts. DAACs release Cumulus and Data Services in production accounts. Data Services are separate from Cumulus to maintain minimum resources for ingest workflows. Egress throttle responds to ESDIS established limits and DAAC configured controls to manage cloud network utilization.
- **Mission Science Data Systems** perform instrument calibration and science product generation**,** scale cloud resources and retrieve raw mission data from the DAAC as needed for reprocessing campaigns.
- **Enterprise Operations** expand to add cloud vendor provisioning and resource monitoring, security access controls and monitoring through NGAP, cloud egress controls through Network Operations and backup of irreplaceable mission science data.

www.nasa.gov