CMR Virtual Products

New Solutions for Enabling Discovery of User-Centric Virtual Data Products in NASA's Common Metadata Repository

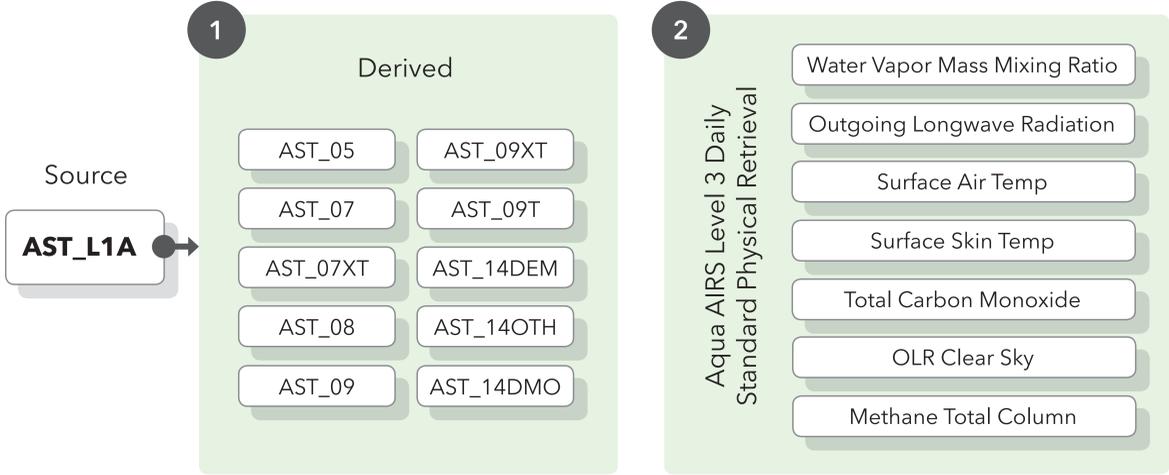**Dan Pilone** EED-2 / EARTHDATA

dan@element84.com

Jason Gilman | Katie Baynes | Dana Shum

IN33B-1804

DERIVED NASA PRODUCTS

In addition to standard data products, NASA's EOSDIS has the ability to create dynamic derived products when requested by the user. Called OnDemand Products, these products share a common source but are the result of additional processing or user requested extraction (subsetting) of specific pieces of information.

1 Derived

Source

AST_L1A →

AST_05	AST_09XT
AST_07	AST_09T
AST_07XT	AST_14DEM
AST_08	AST_14OTH
AST_09	AST_14DMO

2 Aqua AIRS Level 3 Daily Standard Physical Retrieval

- Water Vapor Mass Mixing Ratio
- Outgoing Longwave Radiation
- Surface Air Temp
- Surface Skin Temp
- Total Carbon Monoxide
- OLR Clear Sky
- Methane Total Column

1 For example, the ASTER L1A product is used to create multiple OnDemand products including Level 2 Surface Emissivity, Surface Reflectance, and Surface Temperature.

2 In other cases many different physical variables are packaged into a standard product but not all are applicable for a given user. Aqua AIRS Level 3 Data is a great example of a broad range of measured parameters packaged into a single product. Subsetting tools like OPeNDAP can be used to extract the desired parameters, but require effort on the user's part.
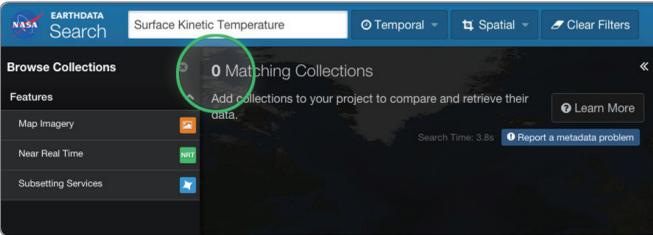
USER CHALLENGES

While derived and custom data products are tremendously valuable to end users, discovery of these products is difficult and inconsistent explicitly because of their dynamic nature.

EOSDIS data discovery is based around product metadata. OnDemand products typically have limited or no discovery level metadata that tools can leverage to help users find and access the data. Improving this situation either means authoring more metadata, or…

THE COMMON METADATA REPOSITORY

The Common Metadata Repository (CMR) builds on the work done by ECHO and the GCMD to provide a unified, authoritative repository for Earth Science metadata.

The CMR enables Earth Science applications to provide end users with nearly immediate access and interactivity across massive stores of Earth Science data by providing high performance, standards compliant, temporal, spatial, and faceted search of the associated metadata.
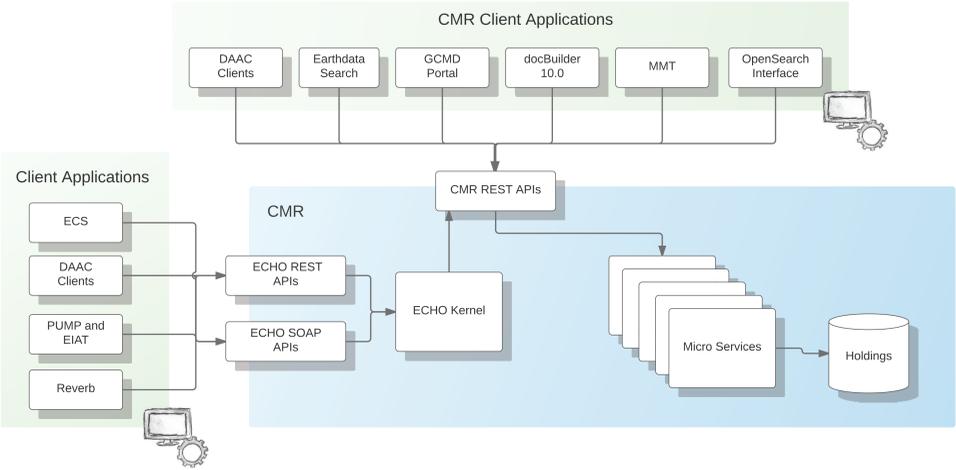
VIRTUAL PRODUCTS

The CMR's microservice architecture and asynchronous ingest pipeline provide hooks to capture incoming source data products. Rather than create derived product metadata by hand, a configurable metadata generation service was developed and attached to the CMR. This metadata generator can watch the CMR ingest flow and dynamically create and maintain metadata for derived products.

While the actual product may not exist on disk, metadata fully describing the collections and granules exists in the CMR and as far as client applications are concerned, represents real products available to users. The CMR can transparently unwrap virtual product requests for compatibility with existing subsetting services while users can request derived products that meet their unique needs, such as an AIRS derived methane product.

Looking forward, the CMR virtual product capability can be used to create bundled products of related data, complex transformations, or even allow for on demand generation of processed data in real-time, avoiding the need for long term storage of reproducible products.

AGU_1215_DP2