ARCHIVE MANAGEMENT OF NASA EARTH OBSERVATION DATA TO SUPPORT CLOUD ANALYSIS

Christopher Lynnes, Kathleen Baynes, Mark McInerney

National Aeronautics and Space Administration

ABSTRACT

NASA collects, processes and distributes petabytes of Earth Observation (EO) data from satellites, aircraft, in situ instruments and model output, with an order of magnitude increase expected by 2024. Cloud-based web object storage (WOS) of these data can simplify the execution of such an increase. More importantly, it can also facilitate user analysis of those volumes by making the data available to the massively parallel computing power in the cloud. However, storing EO data in cloud WOS has a ripple effect throughout the NASA archive system with unexpected challenges and opportunities. One challenge is modifying data servicing software (such as Web Coverage Service servers) to access and subset data that are no longer on a directly accessible file system, but rather in cloud WOS. Opportunities include refactoring of the archive software to a cloud-native architecture; virtualizing data products by computing on demand; and reorganizing data to be more analysis-friendly.

Index Terms— Cloud computing, archive, architecture, data analytics

1. INTRODUCTION

NASA collects and processes increasingly large volumes of Earth Observation (EO) data from satellites, aircraft, in situ instruments and model output. NASA's Earth Observing System Data and Information System (EOSDIS) is responsible for archiving the data and distributing them to a variety of end user communities, including science researchers and applied science users [1]. EOSDIS EO data archives comprise 12 Distributed Active Archive Centers at a variety of locations in the United States who save the data mostly on on-premise disk arrays, with some tape storage. These archives are knitted together by a Common Metadata Repository of metadata at the data collection and file level, which allows a search client to search across all 12 DAACs using a single database.

Since the turn of the century, the data volume archived in has increased 400-fold, to approximately 25 PB in 2017. An additional order of magnitude increase is expected by the year 2024 (Fig 1). Just as important, EOSDIS distributes an annual data volume that is of the same order of magnitude as its cumulative archive volume.

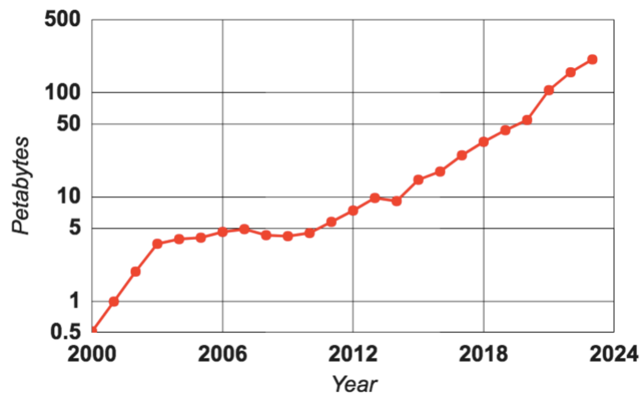

FIGURE 1. HISTORICAL AND PROJECTED CUMULATIVE ARCHIVE VOLUME IN EOSDIS. (YEARS RUN FROM OCTOBER TO SEPTEMBER.)

Cloud-based storage simplifies the ramp-up to handle such large volume increases. It obviates the need to specify and procure large amounts of hardware, plus many of the ancillary activities required, such as allocating (or build out) precious raised floor space, tracking property items, upgrading cooling systems, and upgrading internal networks. Also, the diversity of the community served by major cloud vendors has led to a variety of storage options with different latency, throughput and access options balanced against the respective costs.

However, while modest cost savings may be achievable by using cloud storage over on-premise storage, the real potential arises in the proximity of enormous computing power "next to" the cloud storage. In theory, science researchers using the data could now apply data-parallel processing to analyze data volumes that would simply be too big to download and analyze with their own hardware. Another potential advantage is that having so many data collections in one virtual "place" lends itself to more multi-data-collection studies; these are a particular feature of Earth Observation studies which often meld data from multiple sources based on satellites, aircraft, in situ and model outputs. In reality, the data are often physically separate in cloud storage, but the high-bandwidth interconnects within clouds mitigate this distance.

2. ARCHIVING EO DATA IN CLOUD STORAGE

For the above reasons, NASA is exploring a variety of prototypes using public cloud to archive and distribute EO data. Several of the prototypes use Web Object Storage (WOS) for data archiving in the form of Amazon Web Services Simple Scalable Storage (AWS S3). The prototypes identify the business and operational implications of archiving data in the cloud, as well as demonstrating some of the potential benefits from cloud-based archives.

The core prototype in this suite, named Cumulus, is developing a science archive hosted in the public cloud. Rather than lift and shift an existing archive software system from within EOSDIS, a conscious decision was made to employ cloud-native architecture and services in the prototype. This enabled an architecture centered on AWS Lambda functions triggered by the arrival of data notices, and orchestrated through Step Function workflows. As a result, the workflow aspect is handled largely by cloud-provided services, with the result that most of the custom code is focused on the "business logic", in this case the ingest and processing of different EO science products. The ideal would be to have custom code only for the specific business logic, with cloud services supplying the software infrastructure.

One impact of storing the data into S3 is the egress cost of distributing data out of the cloud, which is exacerbated by the short-term uncertainty of user-requested egress. A traffic rate shaper can mitigate this, taking care to not impact user access unduly. On the other hand, transferring data from WOS to a compute node in the same cloud region does not incur egress cost, incentivizing users to make the paradigm shift toward analyzing data in place (or nearby), rather than downloading to a local machine.

This also incentivizes the archive to offer a number of data reduction services to aid the user in preprocessing the data and decreasing the volume that might be transferred out of the region. Currently, EOSDIS offers several subsetting services, particularly based on the Open-Source Project for a Network Data Access Protocol (OPeNDAP) [2]. Other services to support custom subsetting, regridding, reprojection, quality screening and mosaicking are offered for certain data products. Most of these services are designed to run on a host with attached storage in the form of a POSIX filesystem. Instead, Web Object Storage offers data through the Hypertext Transfer Protocol. Thus, one of the prototypes employs a form of OPeNDAP server that can serve data in Web Object Storage.

Another novel aspect of archiving data in the cloud is that costs accrue so long as the data rest in storage, which can add up to significant expenditures over time. One approach to address this is to virtualize some of the high-volume data and produce them on demand. This strategy was previously employed in serving MODIS Calibrated Radiance data during the transition of EOSDIS from mostly tape to mostly disk in the mid 2000's. However, once disk prices had dropped enough to afford to put the MODIS Calibrated Radiance on disk, this strategy was largely phased out, due to the relatively large latency of on-demand production to simply serving from disk. However, while large on demand requests may be painfully slow for big requests on current computer systems, the ability to access hundreds or thousands of compute nodes at once in the cloud could conceivably shrink the response time to be almost indistinguishable from serving the data from storage. At that point, it becomes a tradeoff between the cost of compute cycles needed to make virtual products vs. the storage cost.

The broad ecosystem of cloud services to fulfill common functions provides another opportunity. In the course of prototyping, we can refactor the archiving software system to use off the shelf services, such as queues, databases, and workflow support services to dramatically reduce the code base.

3. SUPPORTING CLOUD ANALYTICS

Ultimately, the "killer app" for archiving in the cloud is to support analytics using the massively parallel capabilities offered by cloud computing. This has a particularly wide variety of solutions being explored in the community. Most of them involve sharding data across a large number of nodes to enable parallel computing. The sharding solutions include highly distributed databases (e.g., Cassandra [3]), highly distributed filesystems (e.g., Hadoop File System [4]) or in some cases simply dividing data up amongst many WOS buckets. These can be paired with an equally varied set of computational technologies (e.g., Spark [5]). Cloud prototypes are underway to develop end-to-end demonstrations of such systems, with three main aims: (1) to determine feasibility and operability; (2) to demonstrate to the science community what can be accomplished with cloud computing near the data; and (3) to determine the possible impacts on the archive architecture.

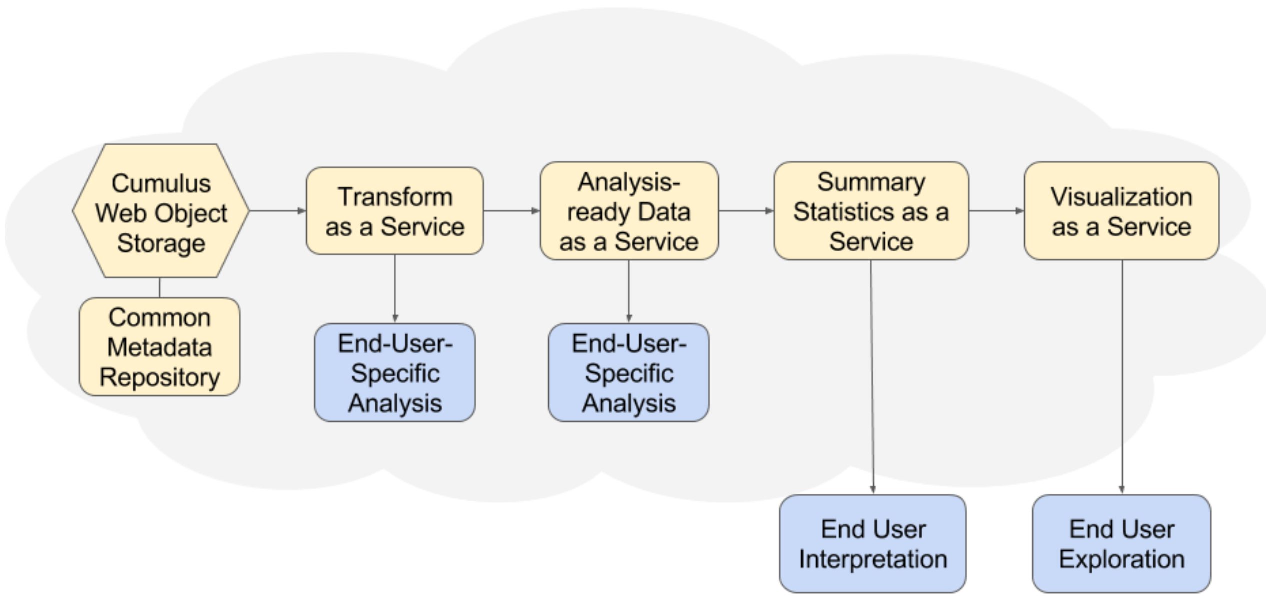

FIGURE 2. ABSTRACTED PIPELINE FOR DATA ANALYTICS IN THE CLOUD.

One common factor among most of these analysis technologies is that they usually require reorganizing and reformatting the data in order to store them in a highly distributed database or filesystem. These forms of analytics-optimized data storage typically have different performance characteristics when paired with appropriate corresponding analytics frameworks. Unfortunately, it is not yet clear if there is a universal optimum combination of analytics optimized storage and analytics frameworks with respect to cost and speed. The optimum may depend on the data characteristics, the analysis algorithm, and the user's specific use case, say, data exploration vs. in-depth analysis. Therefore, we are developing an architectural concept that abstracts the main steps in the analysis process, presenting them to the world as services: this will provide a common framework that can accommodate different components for different combinations of data and use cases (Fig. 2).

A typical end-to-end analysis begins with extraction of the necessary data variables for the spatial and temporal Region of Interest from the Cumulus Web Object Storage. A Common Metadata Repository stores the essential metadata that allow us to generalize this process to work with many types of data in EOSDIS. This is followed by optional data transformation steps, which may include quality filtering, regridding, and/or aggregation over time, space or variables. This corresponds to the "Transform" of the common Extract-Transform-Load process in analysis pipelines. The data are then stored in an analytics-optimized storage framework such as Parquet, HDFS, or Cassandra. The next step provides a set of summary statistics that are commonly used in the EOSDIS user community, usually involving an averaging over latitude, longitude, or time. Based on the remaining dimensions in the data, this is followed by visualization. This overall flow is exemplified by the Geospatial Interactive Online Visualization ANd aNalysis Infrastructure (Giovanni) [6], a current EOSDIS on-premise tool currently being ported to the cloud. Giovanni serves over 1700 data variables to a user base measured in the tens of thousands.

The abstracted architectural concept for archive-proximal analysis in Fig. 2 follows the cloud computing pattern of exposing each key element as a service. This produces several salutary effects. First and foremost, it makes for an open system, one that allows a variety of analysis system developers to plug into the system at any step in the process. Similarly, it can serve an even wider diversity of users than the monolithic on-premise analysis solutions. Interdisciplinary users, educational users, and applications users can work with the visualizations that provide data exploration capabilities with little user effort. On the other hand, research scientists who create and use built-to-purpose analysis can gain value from the data preprocessing and reorganization available via the Transform-as-a-Service and Analysis-Ready Data as a Service. Analysis Ready Data have been promoted by the Committee on Earth Observing Satellites as "*are satellite data that have been processed to a minimum set of requirements and organized into a form that allows immediate analysis with a minimum of additional user effort and interoperability both through time and with other datasets"[7]*. Note also that the data volume generally is smaller on the right side of the pipeline, with summary statistics and their visualizations usually representing a small fraction of the original volume from which they were generated. Thus, it is still important for egress charge reasons for end users to be able to easily apply their own analyses on transformed and analysis-ready data within the cloud. On the other hand, there is little penalty to distributing summary statistics and visualizations to users outside the cloud.

One challenge of this architecture is that archives are hesitant to abandon the data format as received from the provider, implying that they will likely manage two or more copies in different forms. However, it may be cost prohibitive to keep the reorganized version on fast, expensive storage needed for high performance indefinitely. This is particularly the case when a clear winner in price per performance for different data storage technologies is still up in the air. This implies that strategies and mechanisms will be needed for deciding which data to make available in the analytics optimized form, and for how long. These strategies need to be flexible enough to adapt to the ever-changing cost and capabilities on offer by commercial cloud providers.

4. CONCLUSIONS

The heightened interest in Big Data in the larger business community has spawned an increase in off-the-shelf services that are useful for managing and processing data. Managing Big Data in Earth Observation archives can benefit from adopting many of the resultant capabilities. There are many challenges in pivoting from storing data on on-premise hardware to storing them in the cloud. However, there are at least as many opportunities to leverage the co-location of massive processing power near enormous storage resources in order to perform science analysis on larger datasets than ever before, as well as faster than ever before. Recognizing the significant (but sometimes subtle) differences in cloud archive management continues to drive prototype development in NASA Earth Science systems to explore the opportunities and mitigate the risks inherent in such a major evolutionary change in archive architecture.

5. REFERENCES

[1] H.K. Ramapriyan, R. Pfister, and B. Weinstein, “An Overview of the EOS Data Distribution Systems”, *Land Remote Sensing and Global Environmental Change,* Springer, New York, pp. 183-202, 2010.

[2] P. Cornillon, J. Gallagher, and T. Sgouros. “OPeNDAP: accessing data in a distributed, heterogeneous environment,” *Data Science Journal* **2**, pp. 164–174, 2003.

[3] A. Lakshman and P. Malik. 2010. “Cassandra: a decentralized structured storage system,” *SIGOPS Oper. Syst. Rev.* **44**, 2, pp. 35-40, 2010.

[4] K. Shvachko, H. Kuang, S. Radia and R. Chansler, "The Hadoop Distributed File System," *2010 IEEE 26th Symposium on Mass Storage Systems and Technologies*, Incline Village, NV, pp. 1-10, 2010.

[5] J. G. Shanahan and L. Dai. “Large Scale Distributed Data Science using Apache Spark.” *Proc. 21st ACM SIGKDD Intl. Conf. on Knowledge Discovery and Data Mining* (KDD '15). ACM, New York, pp. 2323-2324, 2015.

[6] Liu, Z., and J. Acker. “Giovanni: The bridge between data and science, *Eos, 98,* doi:10.1029/2017EO079299. 2017.

[7] CEOS Analysis Ready Data for Land (CARD4L) Description Document, accessed at:
http://ceos.org/document_management/Meetings/Plenary/30/Documents/5.5_CEOS-CARD4L-Description_v.22.docx