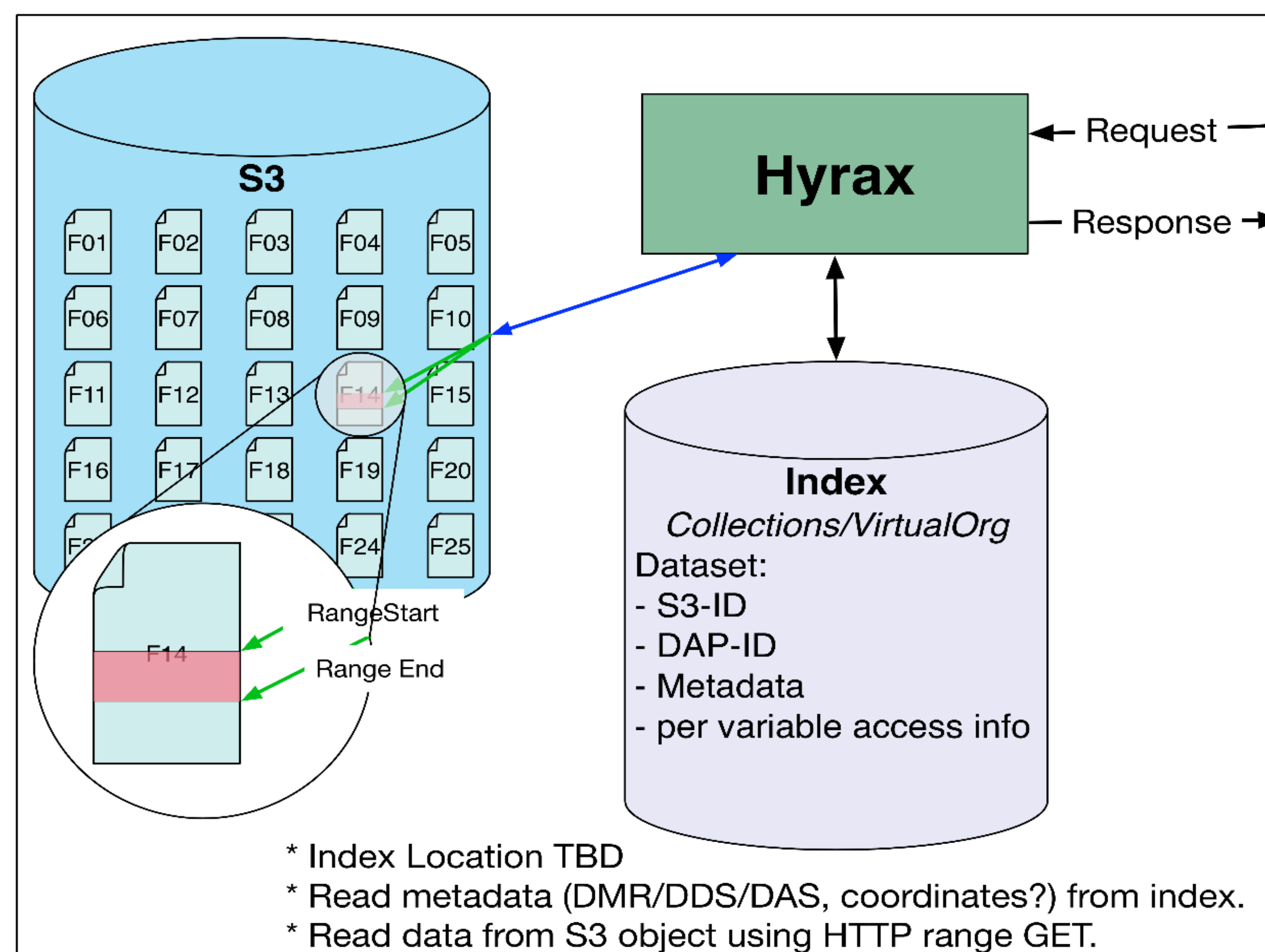


Forming Aggregations using Virtual Sharding: Lessons Learned from Simple Scalable Storage (S3)

James Gallagher, Nathan Potter, Kodi Neumiller
OPeNDAP, Inc.

About Hyrax Support for S3

- Hyrax is a data server that supports access to scientific data using the Data Access Protocol (DAP)
- The data server can read from files, other web services, databases and Web Object Stores (WOS) such as Amazon Simple Storage Service (S3) and Google Cloud Store (GCS).
- Support S3 and GCS uses Virtual Sharding; this enables subsetting files/objects directly from either S3 or GCS.
- The Virtual Sharding technique also enables fast aggregation of data spanning many discrete files/objects.



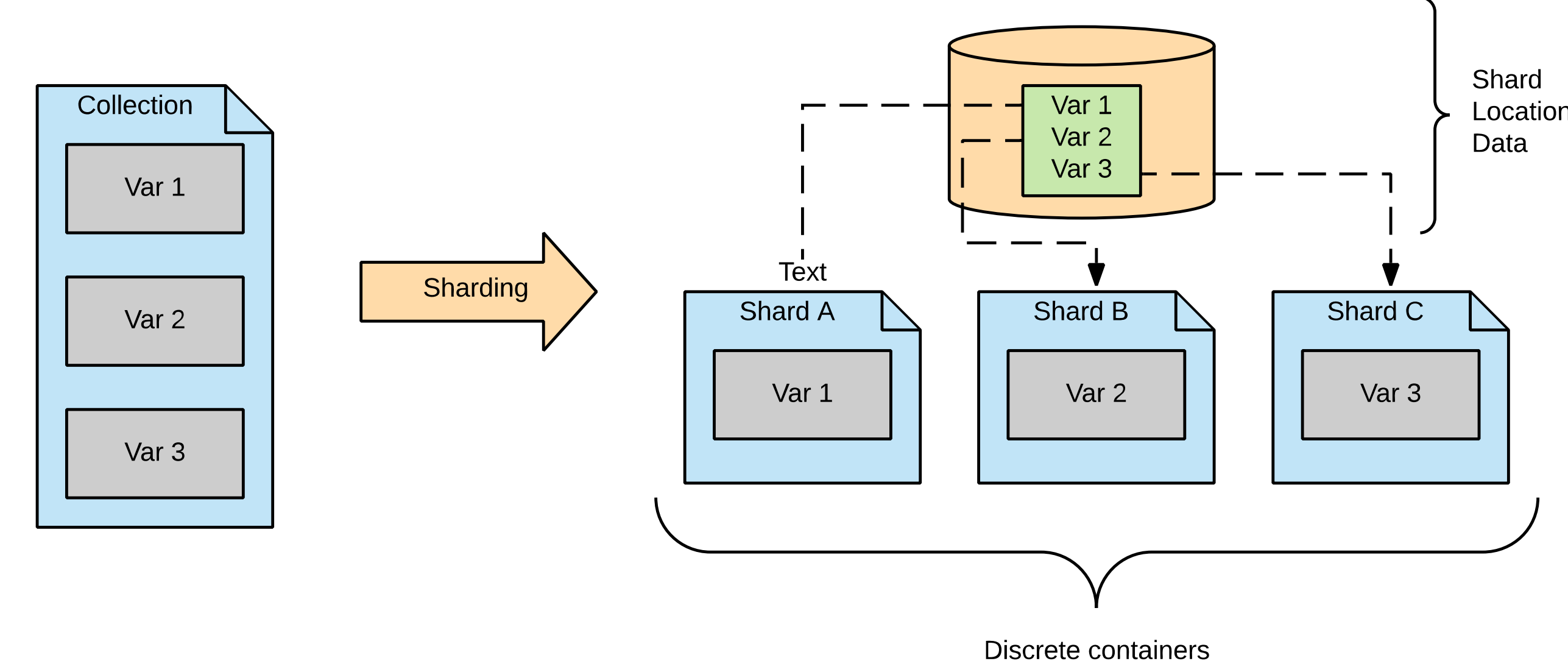
Direct Subsetting from S3 using Sharding

- Hyrax uses **HTTP Range Get** operations to perform Virtual Sharding.
- This enables Hyrax to serve subsets of data without transferring the entire file/object.
- This same technique can be used with local data as well, enabling significant performance improvements when format API libraries are slow to open/read data.
- The modular design of Hyrax enables other groups to build new tools with this (open source) software.

About *Sharding* and *Virtual Sharding*

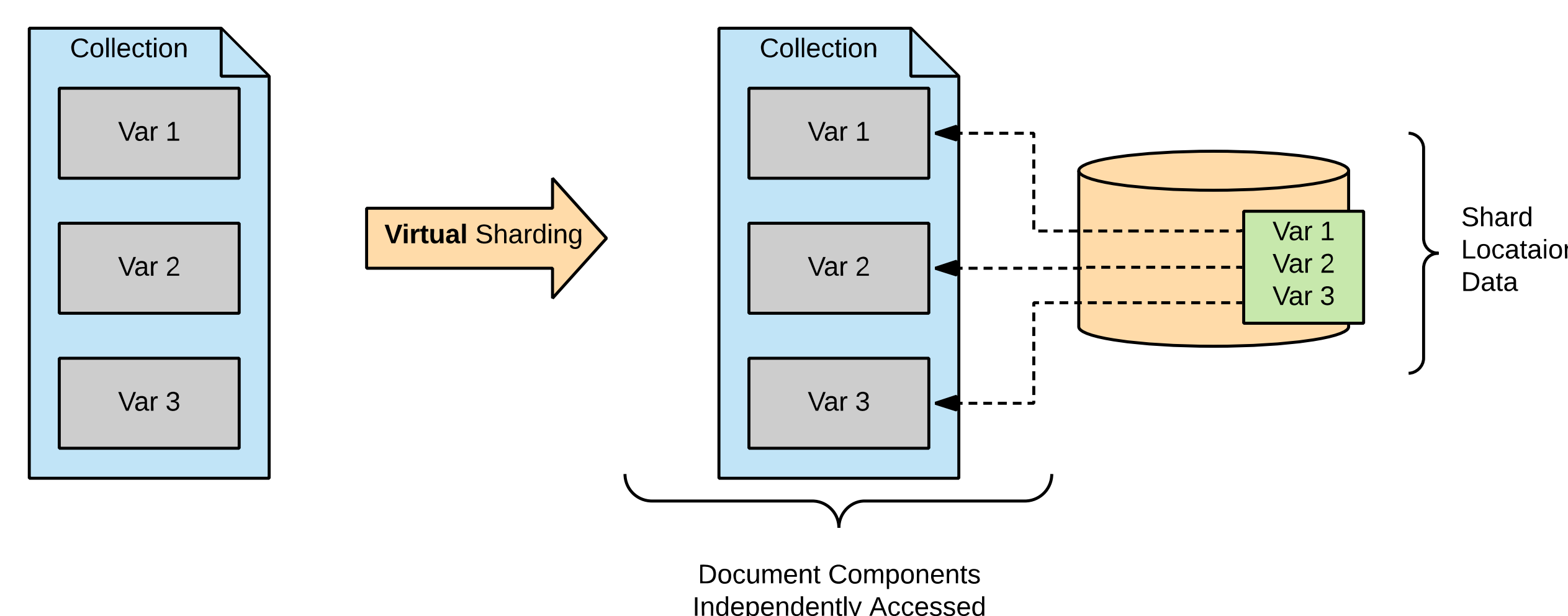
- Sharding deconstructs data into smaller pieces, often to achieve faster I/O performance.
- However, sharding can be used for other reasons.
- Virtual Sharding provides many of the benefits of Sharding, without altering the original structure of the data.
- Virtual Sharding provides a way to read portions of a file/object directly from a Web Object Store (e.g., S3, GCS).
- Virtual Sharding also opens the door to high-performance aggregations.
- Both Cloud Optimized GeoTIFF and Zarr use Virtual Sharding.

Sharding changes the data organization



Sharding is breaking up something that was once a single entity into several smaller containers. Extra information must be stored so software can find the deconstructed data.

Virtual Sharding requires no change to the document

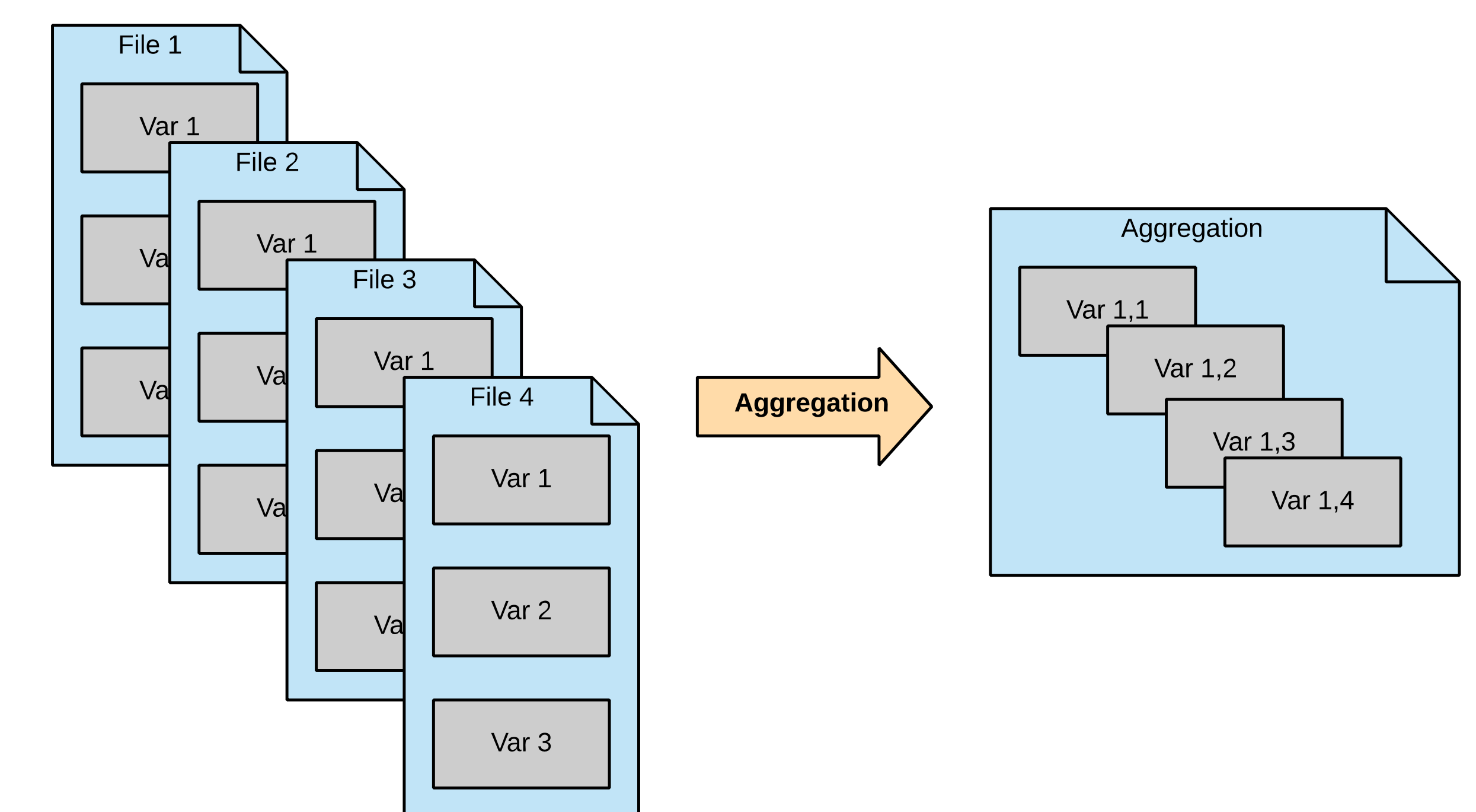


Virtual Sharding: Parts of a document are accessed independently, but the document itself retains its original structure. Extra information enables some software to treat the document as if it was **deconstructed** while other software accesses the document without modification.

Aggregation Using Virtual Sharding

- Previously, Hyrax had to open each document and build extensive metadata to form aggregations.
- Using Sharding, the server eliminates this time consuming step.
- Using Virtual Sharding, the servers can get these performance benefits without reformatting the data

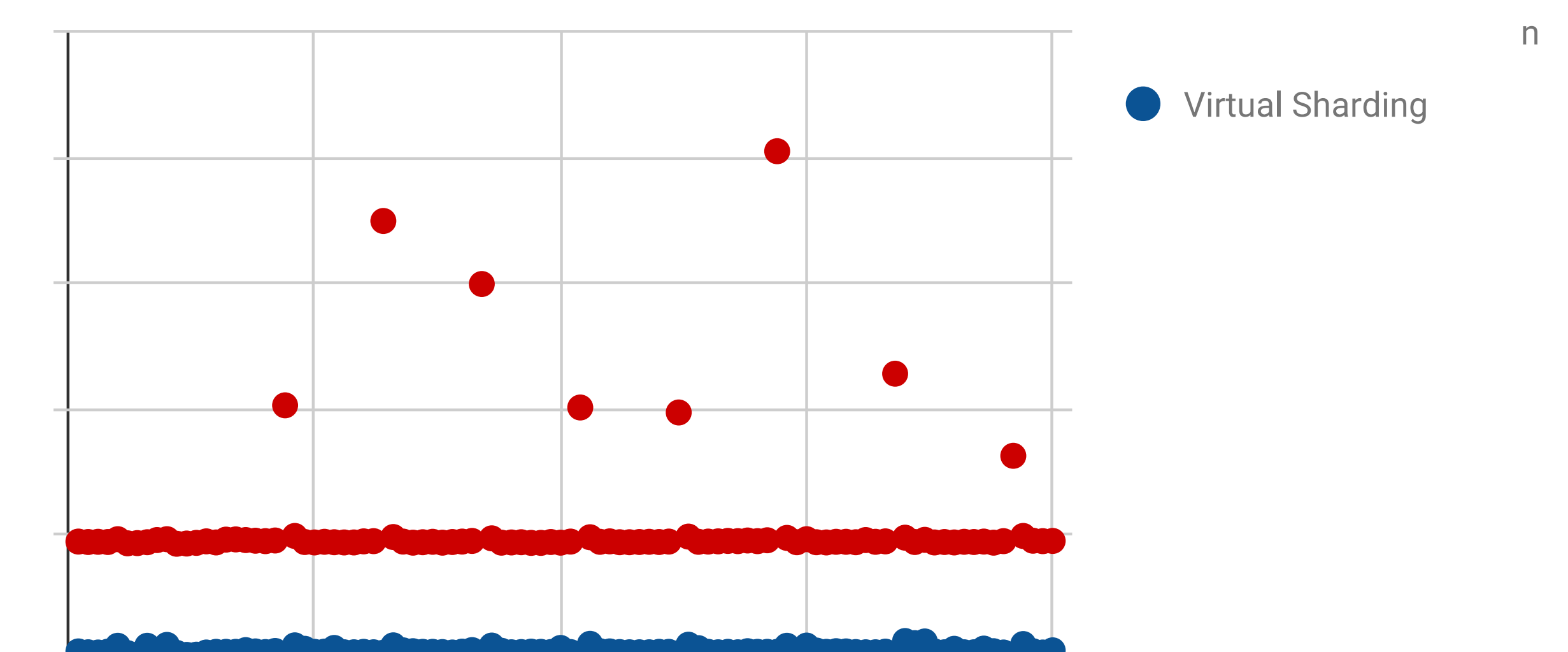
Aggregation Combines Data from Different Files



This is a limited view of aggregation – building a cube from a set of 2D slices – but it is particularly useful for time-series satellite data. Hyrax and other OPeNDAP servers can perform other kinds of aggregation.

Aggregation: Performance Improvement with Virtual Sharding

Time to access all the data from 365 files, 100 repetitions, Data on S3



Summary

- Data do not have to be sharded to be subset in-place on S3.
- Virtual Sharding reduces aggregation response times by an order of magnitude.
- This technique can be used with local data, too.

This work was supported by NASA/GSFC under Raytheon Co. contract number NNG15HZ39C



Contact US

Use the opendap-tech@opendap.org mail list or, send a message to support@opendap.org.
{jgallagher, ndp, kneumiller}@opendap.org