

A satellite image of Earth showing the Western Hemisphere, with the Americas on the left and the Atlantic and Pacific Oceans on the right. The image is overlaid with a semi-transparent blue rectangle containing the title text.

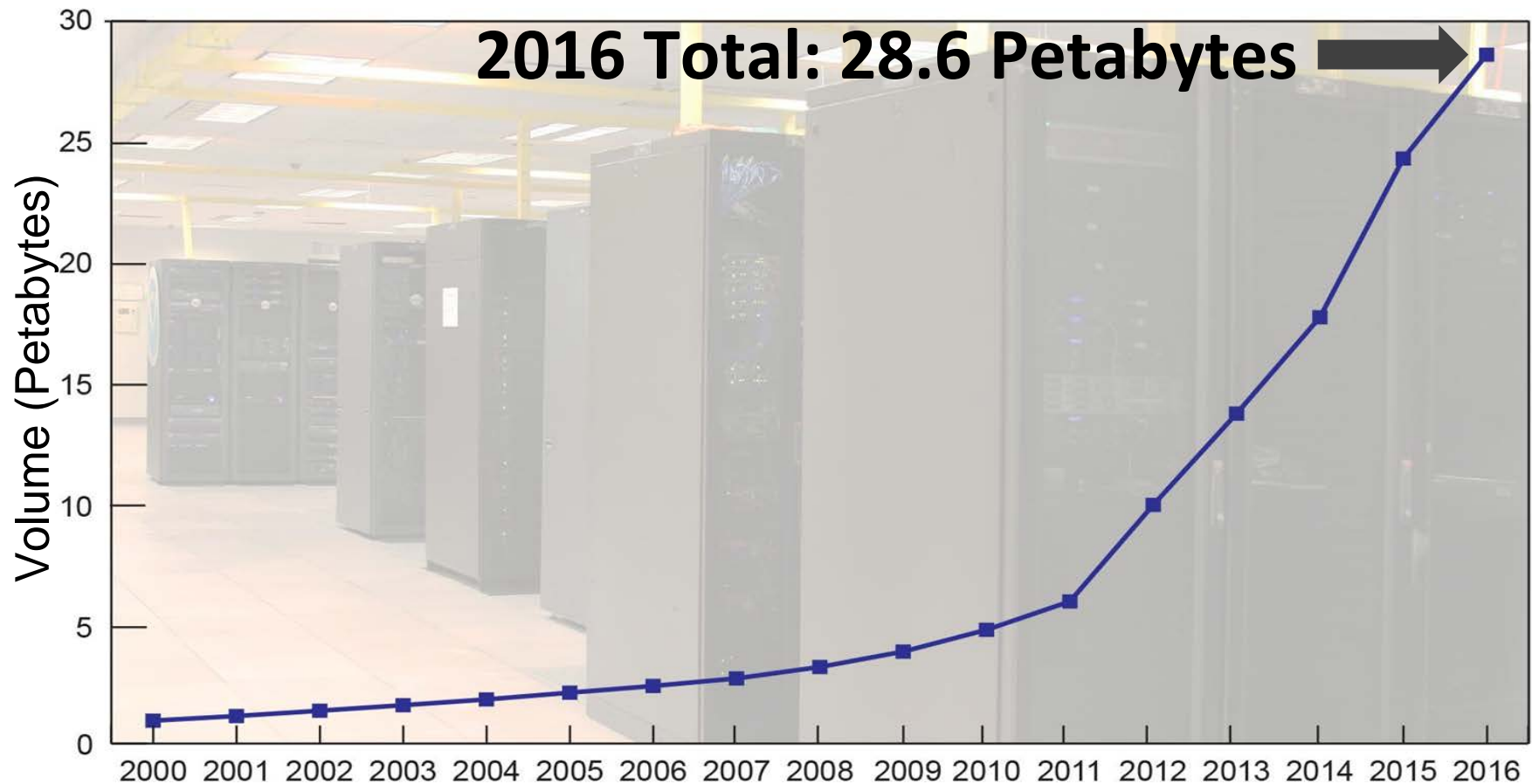
NOAA's Big Data Project: Status Update

Shane Glass
Project Manager, NOAA Big Data Project

IOOS Advisory Committee Fall Meeting
October 27, 2017

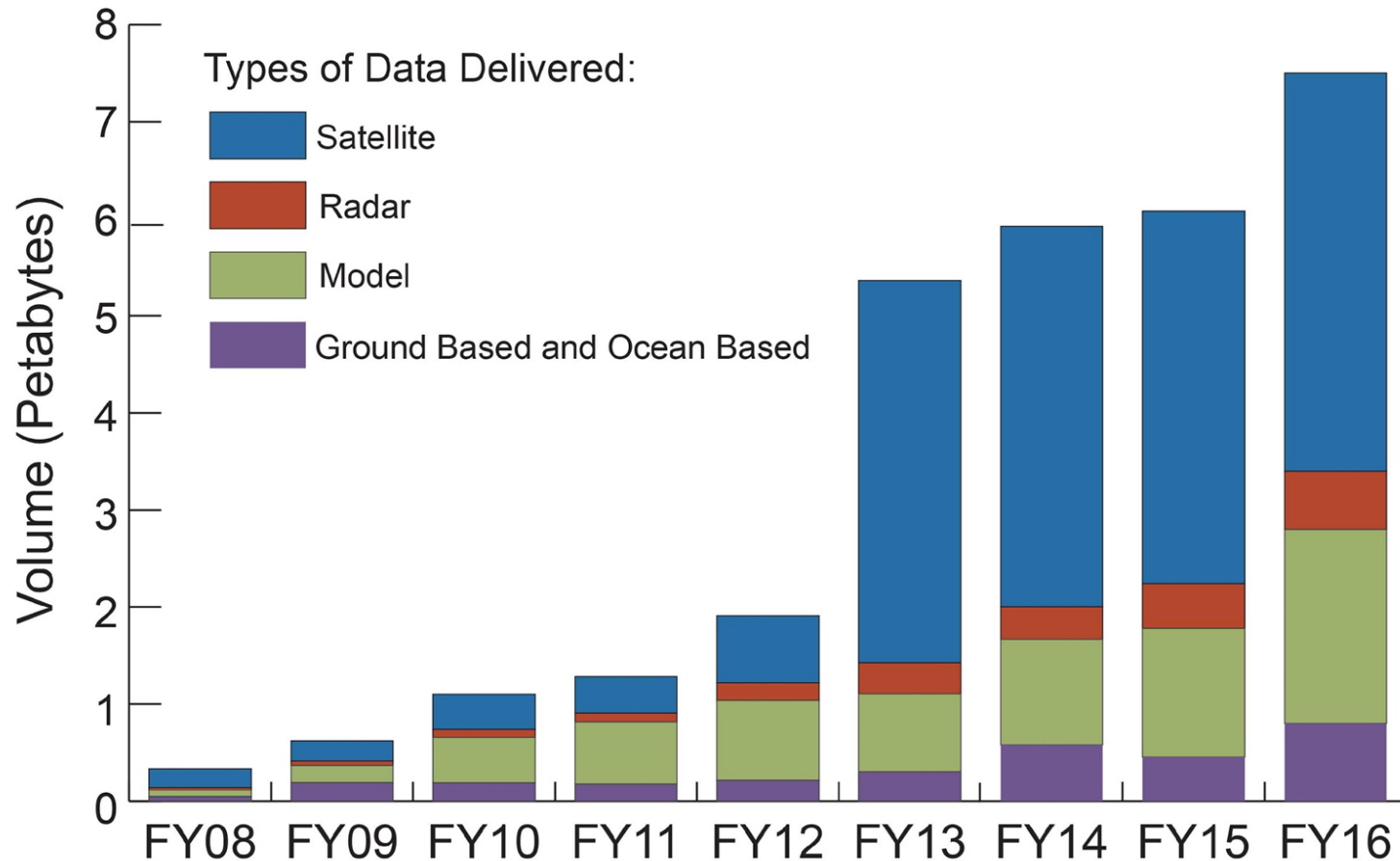
Why Open Data Partnerships?

Archive Storage Volume Increasing Over Time



Why Open Data Partnerships?

Data Dissemination Volume Increasing Over Time



Why Open Data Partnerships?

- NOAA's data are increasingly popular and valuable.
 - **Budgets** for capacity and security: **Static**
 - **Costs** for data access support: **Increasing**
- Create **new economic opportunities** for the Nation
 - **Democratize** data access

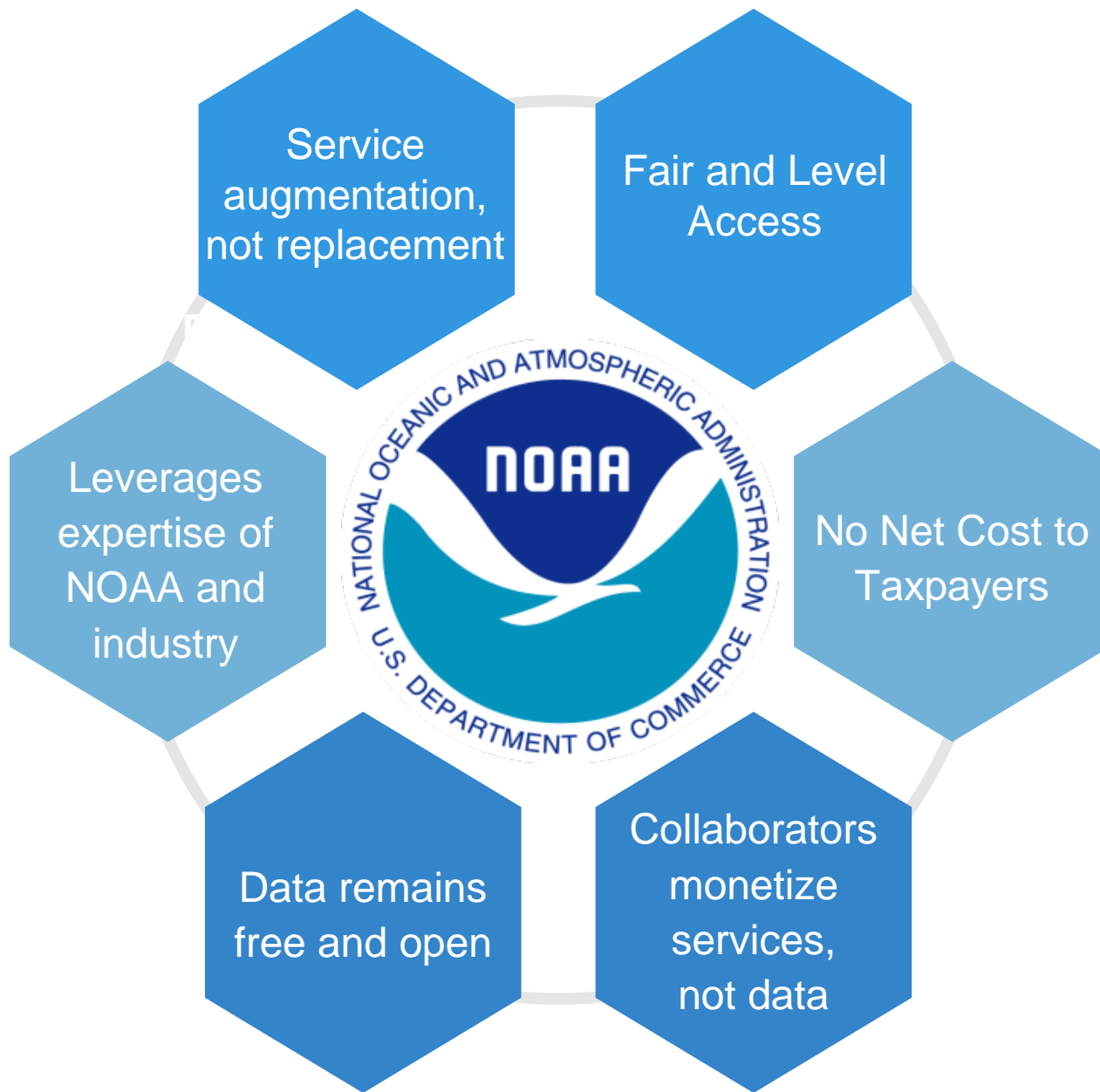
Big Data Project Objective:

Define enterprise business and technical models for data-centric Public-Private Partnerships



Google Cloud Platform





Leverage value of NOAA's data to increase their utilization ⁵

BDP Data Access Strategy

Leveraging Industrial Partners

Augment



Add
Capabilities

Amplify

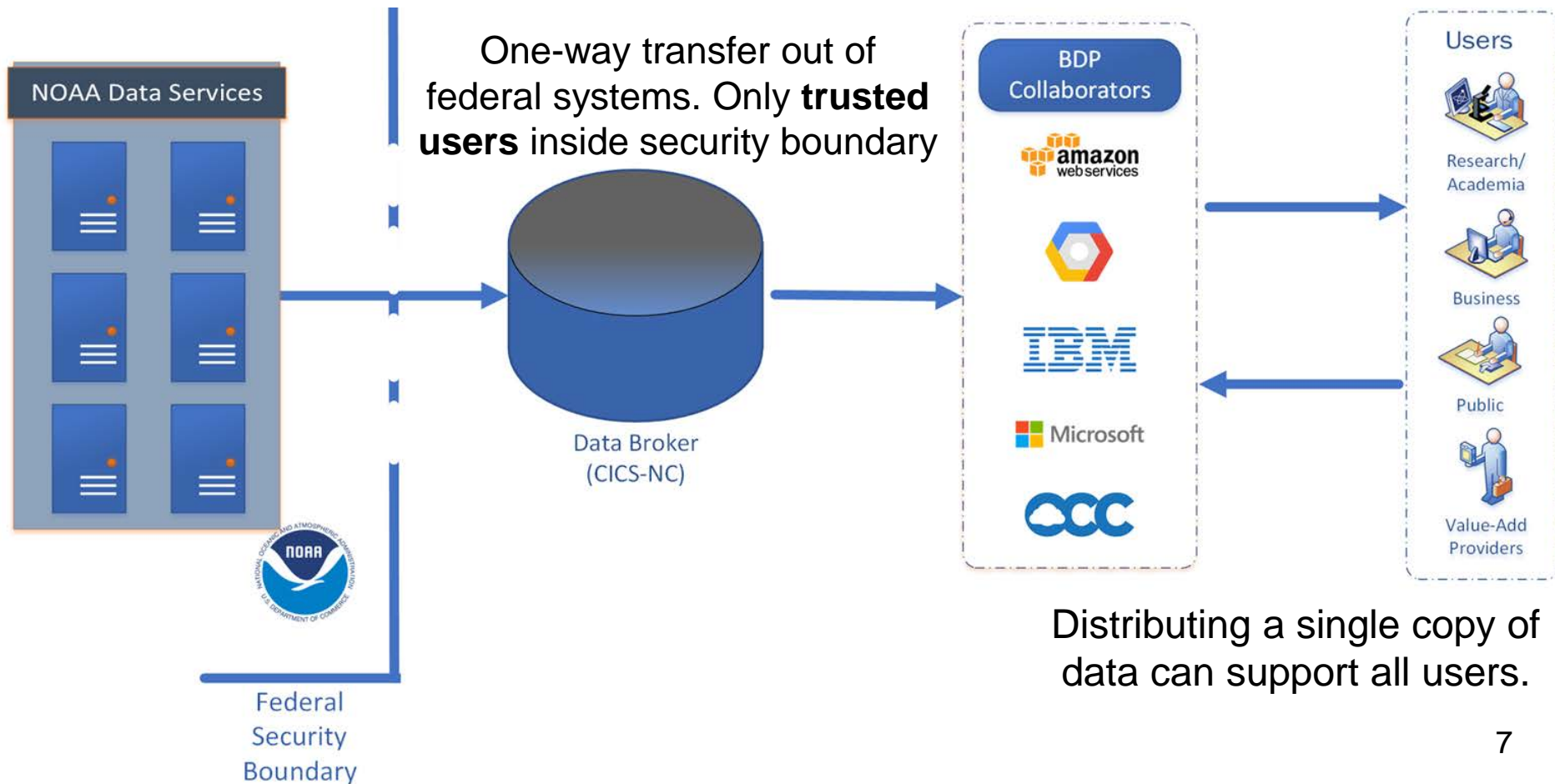


Add
Capacity

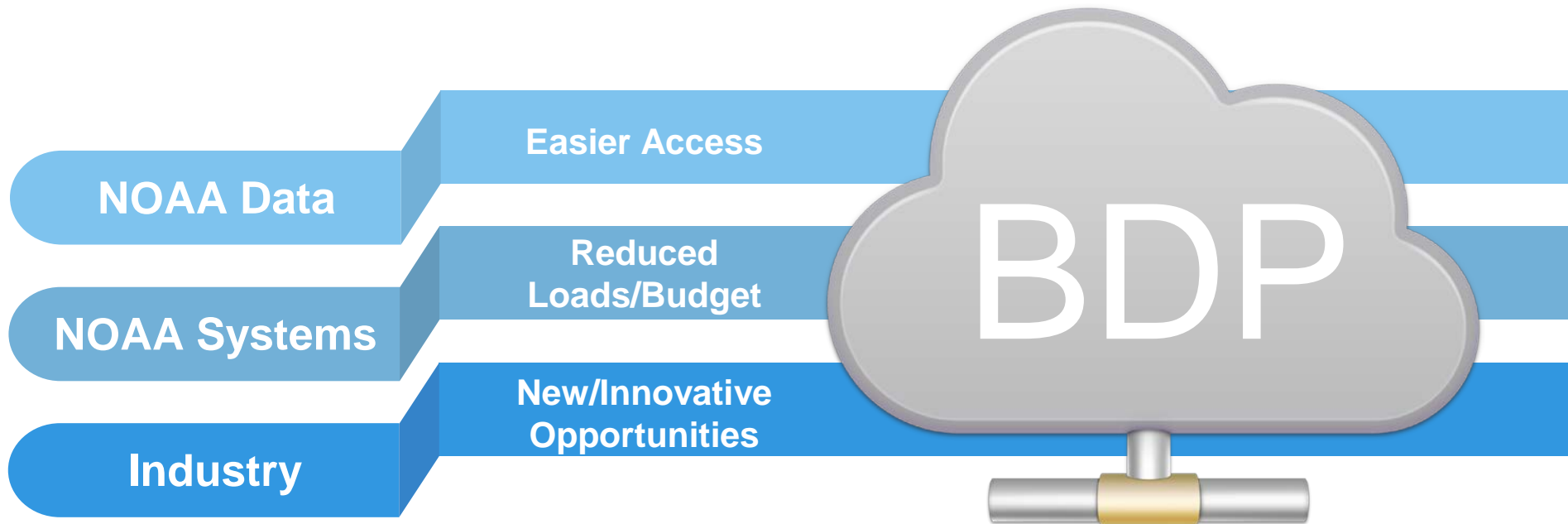
Bring Processing to the Data

BDP Distribution Scheme

A potential paradigm shift for providing data access



Tangible BDP Benefits



Increased Data Utilization

AWS and NEXRAD L2

Data Usage

Increased 2.3X



Decreased 60%



NCEI Server Load

Increased Data Utilization

GCP and Historical Observation Data

Climate Data Online

Climate Data Online (CDO) provides free access to NCDC's archive of global historical weather and climate data in addition to station history information. These data include quality controlled daily, monthly, seasonal, and yearly measurements of temperature, precipitation, wind, and degree days as well as radar data and 30-year Climate Normals. Customers can also order most of these data as [certified hard copies](#) for legal use.



Browse Datasets

Browse documentation, samples, and links



Certify Orders

Get orders certified for legal use (requires payment)



Check Status

Check the status of an order that has been placed



Find Help

Find answers to questions about data and ordering

Increased Data Utilization

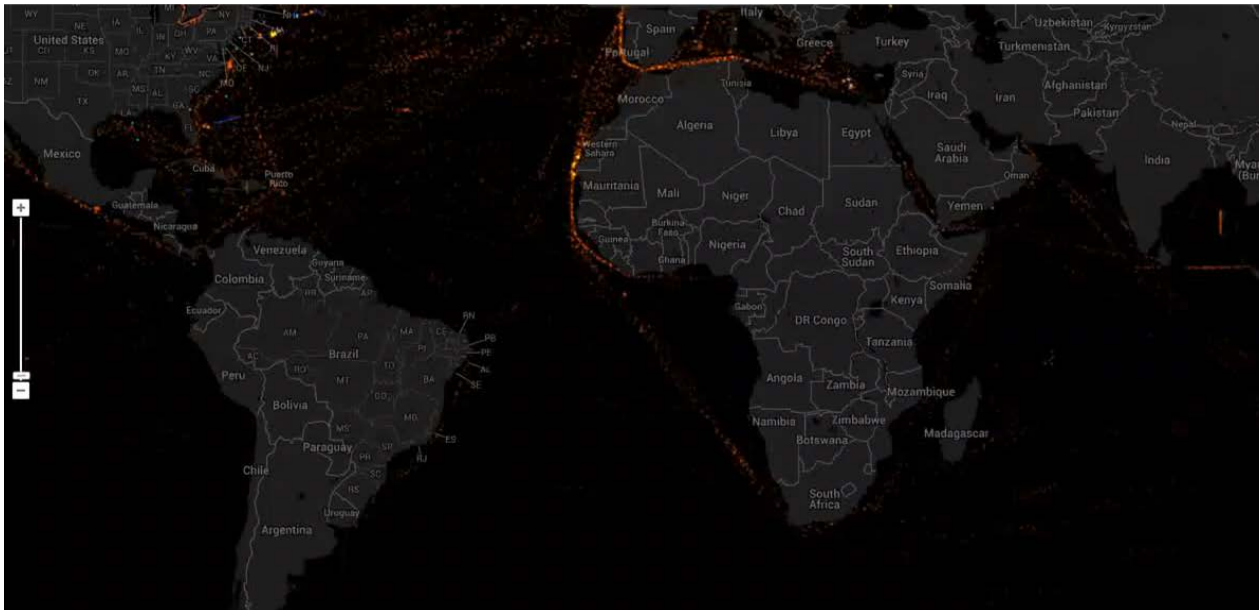
GCP and Historical Observation Data

- Three datasets uploaded in 2017:
 - Global Summary of the Day (GSOD)
 - Global Historical Climatology Network (GHCN)
- Early metrics are extremely positive
 - 800,000 data requests between Jan and Apr 2017
 - 1.2 **PBs** of data delivered
 - Completely **organic utilization**
- More datasets available through Google Earth Engine

Increased Data Utilization

GCP and Historical Observation Data

- International Comprehensive Ocean-Atmosphere Dataset (ICOADS)



Global marine data
from ships and buoys...

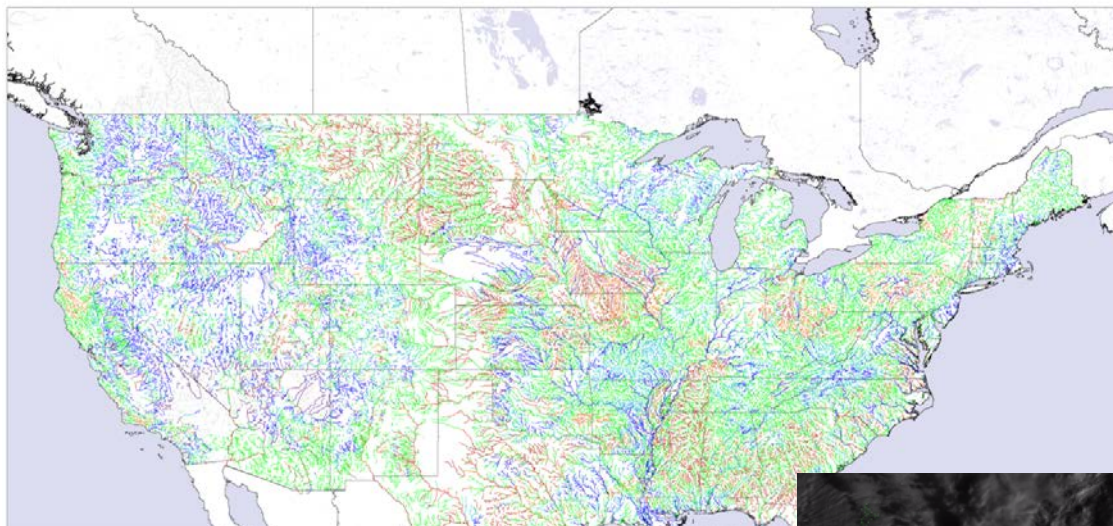
...as far back as 1662

Ongoing and Upcoming Efforts

National Water Model Streamflow Anomaly Guidance

Analysis valid for 2017-06-01 19:00:00 UTC

Model initialized at 2017-06-01 16:00:00 UTC



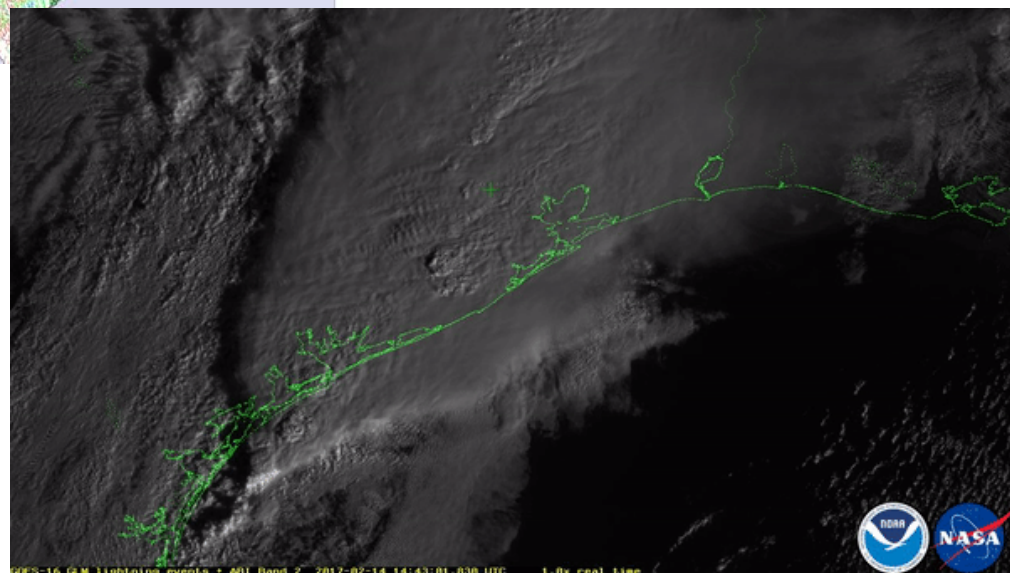
National Water Center: <http://water.noaa.gov/tools/nwm-image-viewer>

National Water Model:

- 23-year reanalysis
- Real-time forecast
- Data transfer underway

Provisional GOES-16:

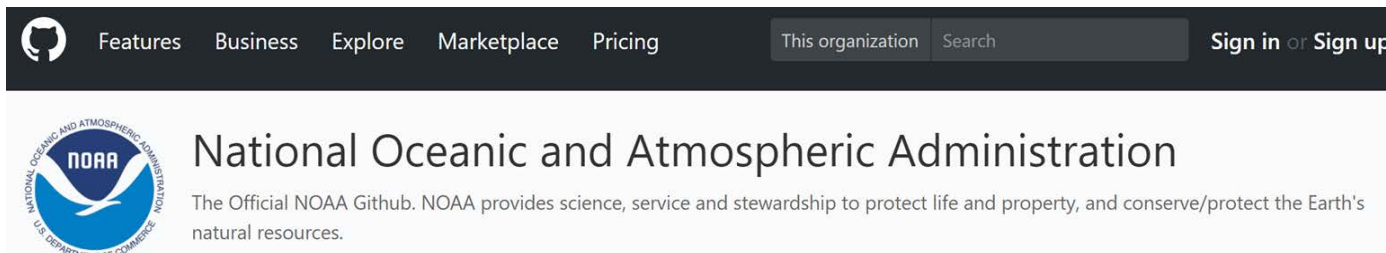
- Now: L1b ABI Products
- Soon: L2 products (GLM)



NOAA NESDIS: <https://www.nesdis.noaa.gov/content/flashy-first-images-arrive-noaa%E2%80%99s-goes-16-lightning-mapper>

A New Approach to Enable Ocean Data

- Target a problem instead of moving individual datasets
 - Deliver multiple datasets and code to the platform to facilitate use
- Compute in “sandboxes” alongside NOAA (and others) data
- NOAA has approved the use of GitHub (NOAAGov)
 - Guidance from NOAA CIO released in September



The image shows a screenshot of the NOAA GitHub repository page. At the top, there is a dark navigation bar with the GitHub logo on the left and links for 'Features', 'Business', 'Explore', 'Marketplace', and 'Pricing'. To the right of these links are 'This organization' and 'Search' buttons, and 'Sign in or Sign up' text. Below the navigation bar, the NOAA logo is displayed on the left, featuring a blue circle with a white bird and the text 'NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION' and 'U.S. DEPARTMENT OF COMMERCE'. To the right of the logo, the text reads 'National Oceanic and Atmospheric Administration' in a large font, followed by 'The Official NOAA Github. NOAA provides science, service and stewardship to protect life and property, and conserve/protect the Earth's natural resources.' in a smaller font.

Way Ahead for the BDP

- **Seek extension of CRADAs** by one year (April 2019)
- **Finalize Lessons Learned Document**
 - Captures information gathered in first 2.5 years
- **Define requirements** for any future agreement
 - Actively engaging with NOAA offices to gather requirements
 - Examine level of service and commitment requirements
 - Define the “Data Broker” role more clearly
- **Determine the optimal partnership models options** for the time period beyond the CRADAs

Discussion

shane.glass@noaa.gov

#NOAABigData

<http://www.noaa.gov/big-data-project>