

Making Data Available on the Cloud for Decision Support Applications through NOAA's Big Data Project

NOAA Big Data Project
Ed Kearns, Chief Data Officer
Tony LaVoi, Geospatial Information Officer

National Geospatial Advisory Committee April 4, 2018

NOAA's Geospatial Programs are Diverse

Geospatial technologies provide the framework to collect, store, analyze, and disseminate 'NOAA's Environmental Intelligence'

Geodetic Control

Nautical Charts

Shoreline

Coastal Topography/Bathymetry

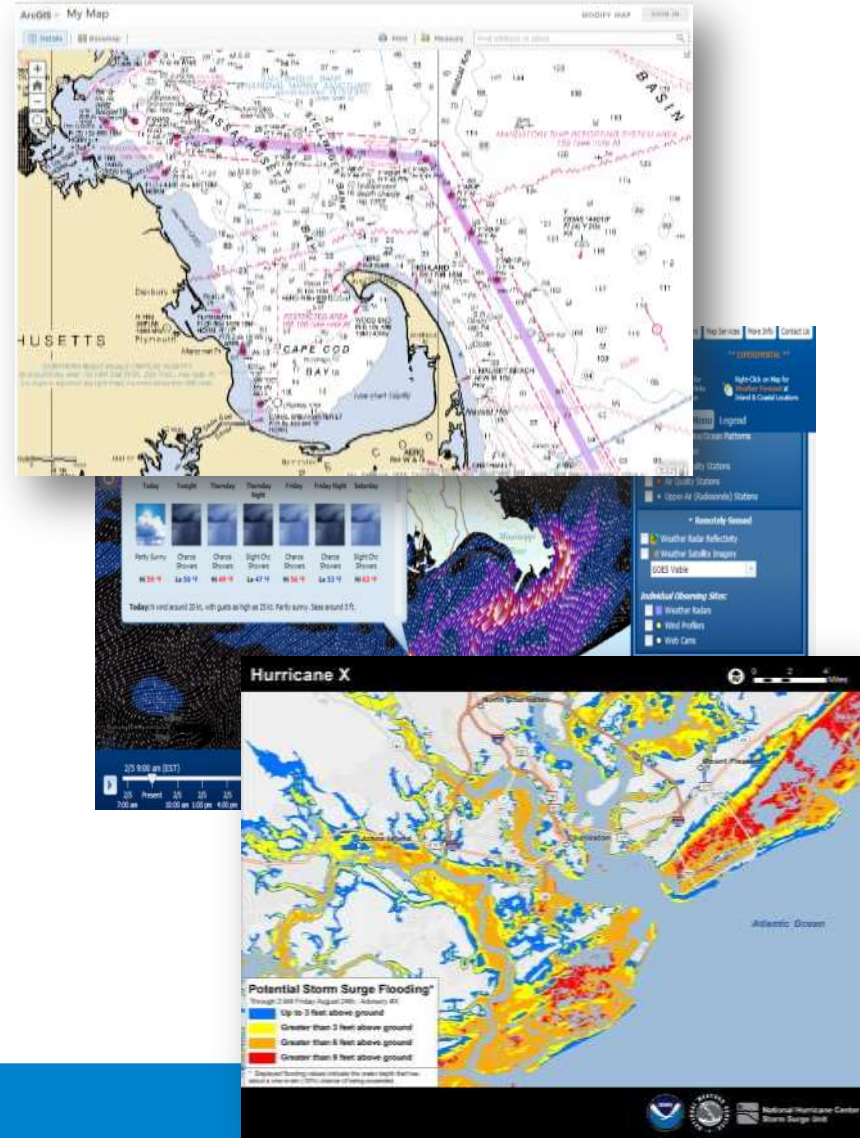
Marine and Coastal Protected Areas

Fisheries Catch and Habitat

Weather and Climate

Satellite and Aircraft Imagery

Authoritative
GeoData



NOAA has "Big Data"

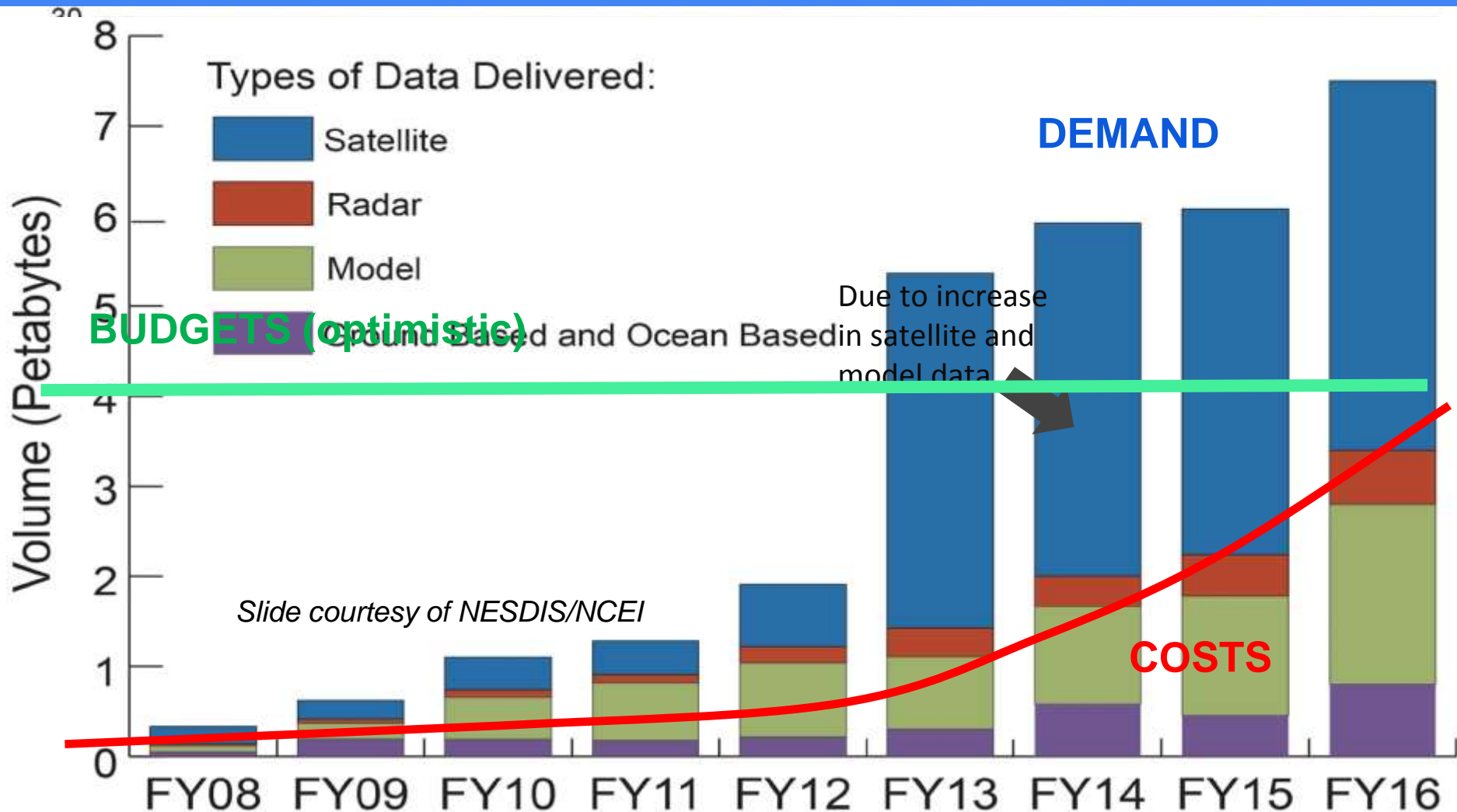
- Volume: 30+ PB
- Velocity: real-time to historical
- Variety:
 - Satellites
 - Weather radars
 - Buoy networks
 - Tide gauges
 - Ships
 - Aircraft
 - Human observers
 - Animal telemetry
 - Numerical models

NOAA data are unique, valuable, and irreplaceable, but most users want answers and not just data



Why? Proportionally Increasing Costs with Data Volume & Demand

NOAA/NCEI's Environmental Data Archive



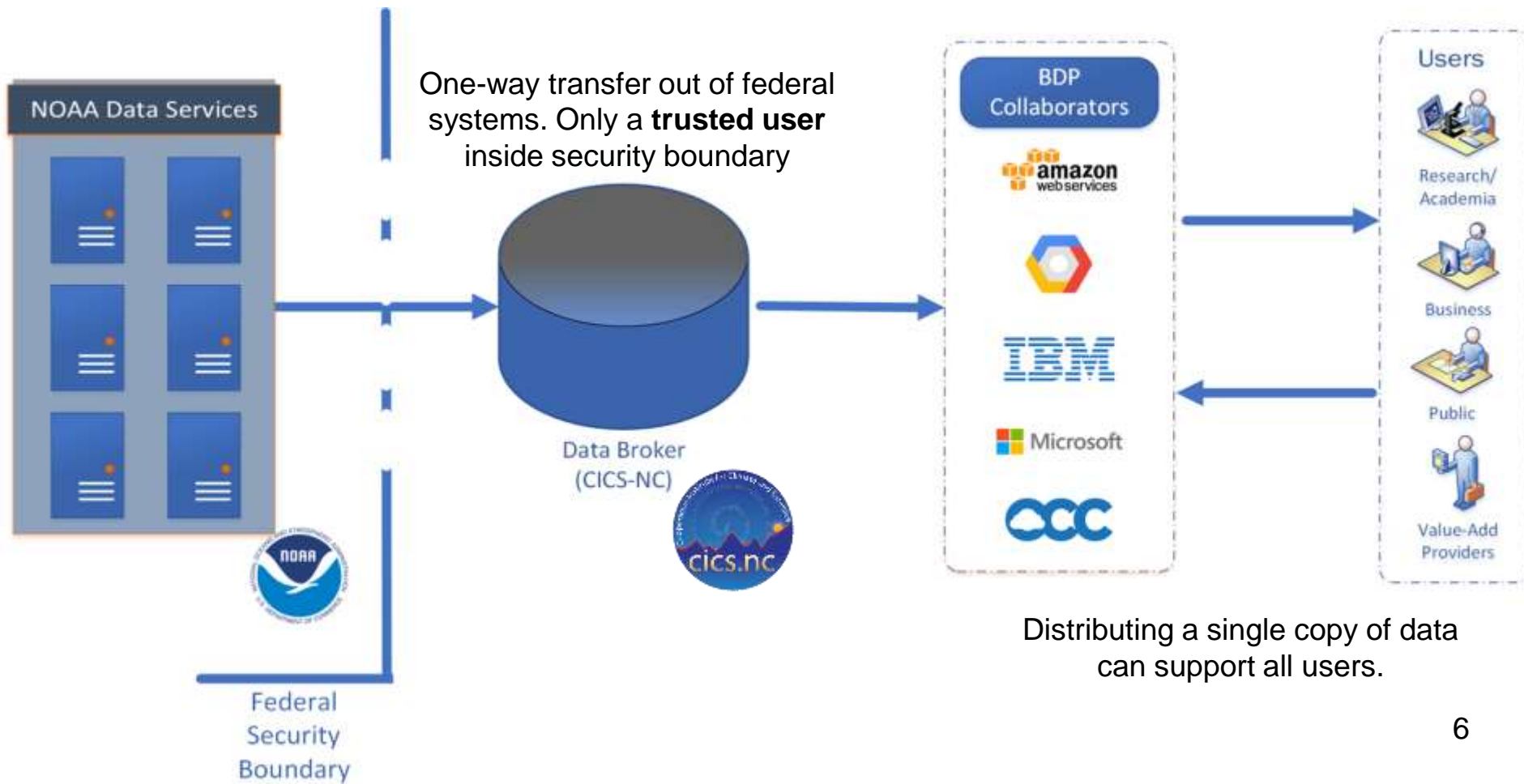
Big Data Project Basics



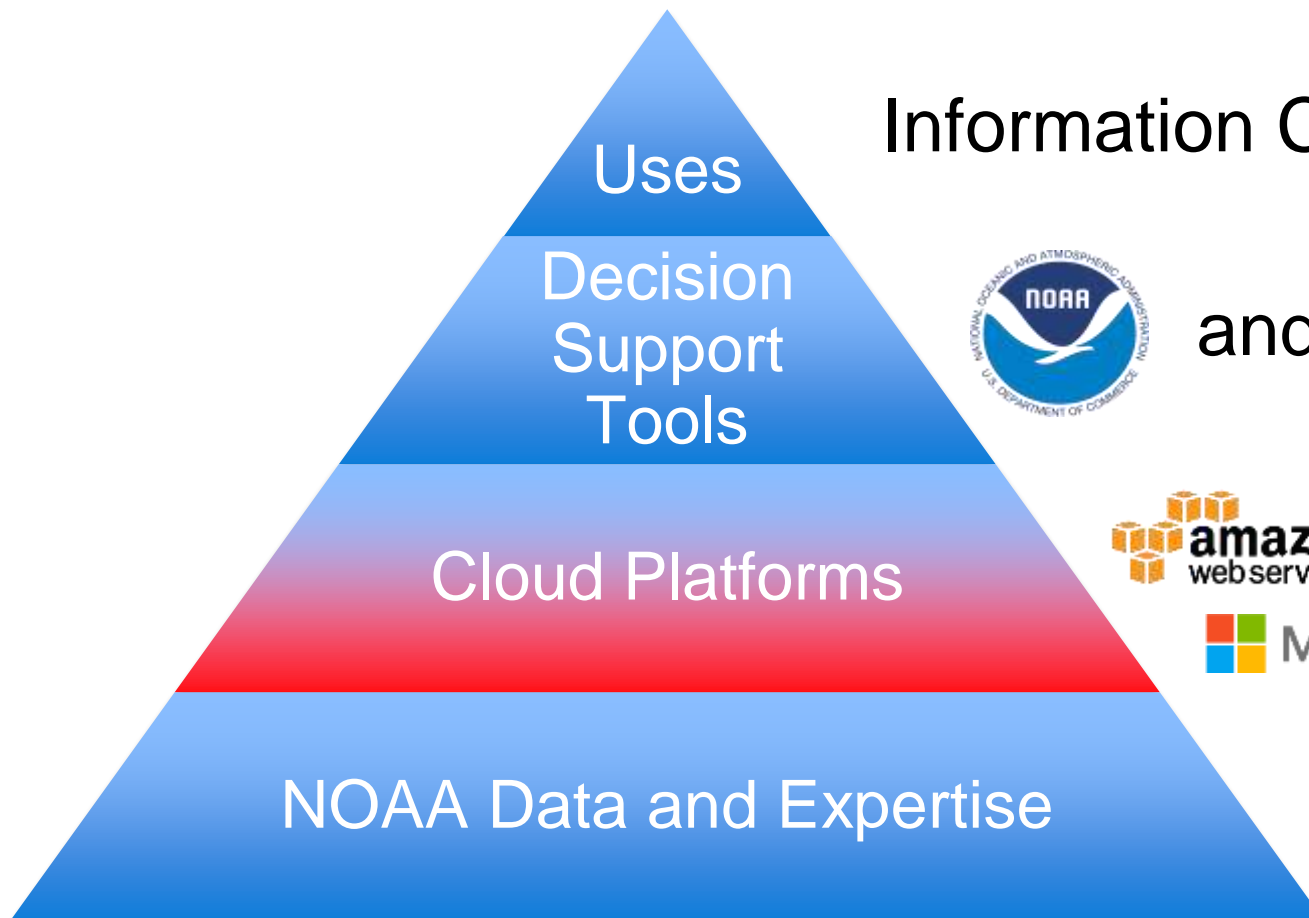
- Experiment with Industry via Cooperative Research and Development Agreements
 - 5 separate, identical agreements – extending to Apr 2019
- Industry provides access to NOAA's open data to all users
 - Data remain open, are not to be sold
 - Collaborators monetize services based on data
- Uses open data to combine 3 powerful resources:
 1. NOAA's **science and subject matter expertise**
 2. Industry's **data storage and cloud** expertise
 3. Industry's **user-focused tools and processing**

NOAA Big Data Project

How to best provide cloud access to NOAA data?



Industry Builds What is Needed to Support its Users



Information Consumers



and 3rd Parties



Summary

- The Big Data Project is working with its CRADA Collaborators to publish NOAA data to their platforms.
- The Collaborators are providing services to enable data use, including decision support applications.
- Through available stats are limited, the BDP has seen increases in data usage from 2 to 100 times.
- NOAA has seen a decrease load on its servers and networks to support data access services

Big Data Project

Collaborators' Offerings

35+ NOAA datasets are now available



- **AWS**

- <https://aws.amazon.com/noaa-big-data/>



- **Google Cloud Platform**

- <https://cloud.google.com/bigquery/public-data/>
- <https://explorer.earthengine.google.com/#index>



- **IBM**

- <https://noaa-crada.mybluemix.net/>



- **Microsoft**

- No public services to date



- **Open Commons Consortium**

- <http://edc.occ-data.org/>

Questions?

ed.kearns@noaa.gov
tony.lavoi@noaa.gov

#NOAABigData

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