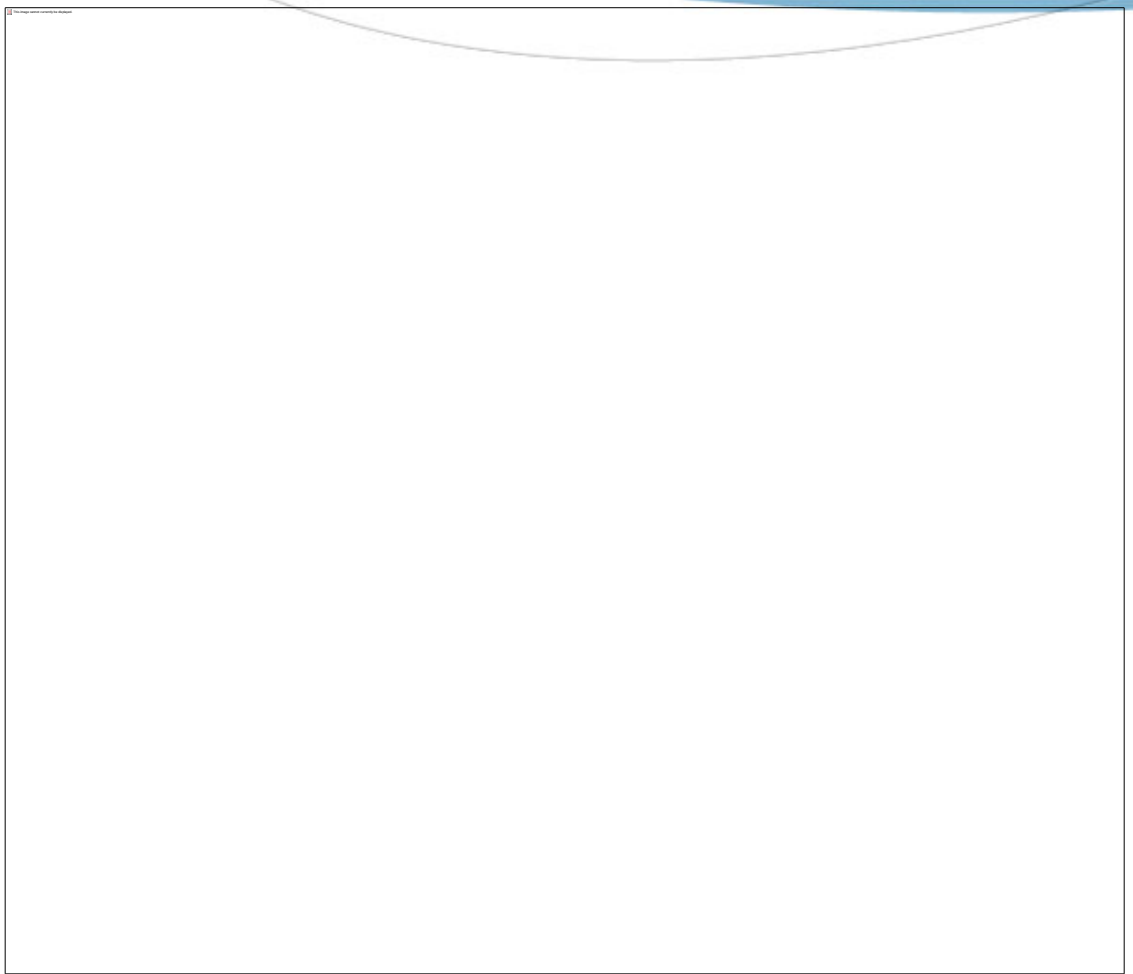
The NOAA Big Data Project Overview

Amy Gaskins, OCIO
March 2016

What is the Big Data Project?

- An innovative approach to **publishing NOAA's vast data resources** and positioning them near cost-efficient high performance computing, analytic, and storage services **provided by the private sector**
- High-quality environmental data at low cost
- Increased public access to taxpayer-funded data
- Opportunities for new data products and services

Why Start Now?

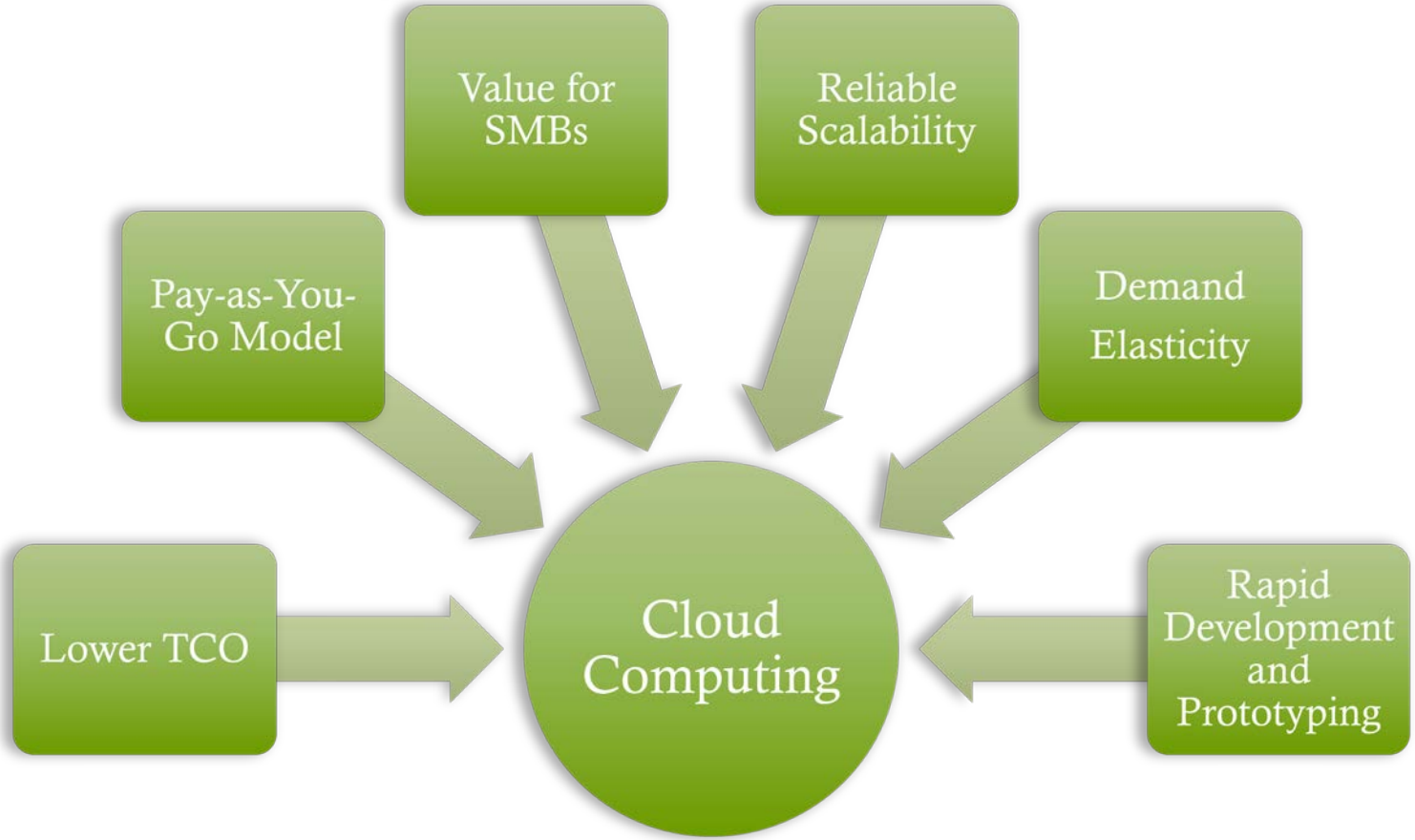


- NOAA's projected data growth is exponential
- By 2020, we'll have 160 petabytes of archived data
- The rate of data capture will only continue to increase

Why a CRADA?

- Existing NOAA infrastructure and funding cannot support growing amount of data or level of external demand
- CRADA is a low-risk mechanism for creating and testing a smaller version of the full market ecosystem
- Provides an iterative approach to dissemination without disrupting NOAA operations
- Allows for lessons learned and real-time modifications to the dissemination process

Why Cloud Services?



The BDP Dissemination Model

Traditional Data Request

Big Data Project

Important CRADA Rules of the Road

- Valid for three years with annual renewal option
- Collaborators and/or NOAA may choose to terminate with 30 days' notice
- Collaborators have non-exclusive access to NOAA data
  - All NOAA data is up for discussion (excluding ITAR-restricted and national security sensitive data)
- Collaborators must provide users with equal access to NOAA data on equal terms

The partnership model…

- Allows access to the entire historical archive of large datasets
- Uses a market to provide choices to industry without NOAA interference
- Increases the reach of NOAA data
- Frees up NOAA personnel from repeating data extracts

Example Case Studies

Consumer Packaged Goods & Retail
Oil & Gas
Insurance
Transportation & Shipping

Consumer Packaged Goods & Retail

- Scientists at a consumer goods company want to predict future changes in humidity to develop lotions that better absorb moisture
- A grocer wants to understand whether next year's crop can meet the rising demand for locally sourced produce

Oil & Gas

- An oil company is designing a rig and needs to make sure it can withstand future hurricane conditions
- Rig workers benefit from better sea condition forecasts that help ensure their safety

Insurance

- Underwriters are assessing hurricane damage risks in order to determine whether a company should return their business to Florida
- Auto insurance underwriters can adjust the policy pricing of drivers who don't drive during bad weather conditions

Transportation & Shipping

- A marine shipping company is looking for a better method of finding ideal paths when rerouting its ships during storms
- Airlines want more accurate winter weather predictions to prevent travel delays and control monetary losses (>$100M)

Multiple Industries, Same Interest

Wants to…

…develop new soil additives to increase crop production

…make smart bets on wheat commodities

…understand how weather will affect next year's crop

…accurately forecast this year's cereal production

…predict cereal manufacturers' quarterly earnings

Current Status

- Released NEXRAD Level II data
  - Amazon, Microsoft, OCC, and Google
- Planning release of next data set(s)
  - Currently GOES/GOES-R and MRMS, but looking for others
- Still in early stages, but moving faster
  - Setting up operational and technical working groups with representation from line offices
  - Increasing industry outreach beyond private weather enterprise
  - Engaging with other federal agencies and policymakers

Questions?

Contact Information

Email: Amy.Gaskins@noaa.gov

Twitter: @AmyVGaskins

LinkedIn: www.linkedin.com/in/amygaskins