



NOAA

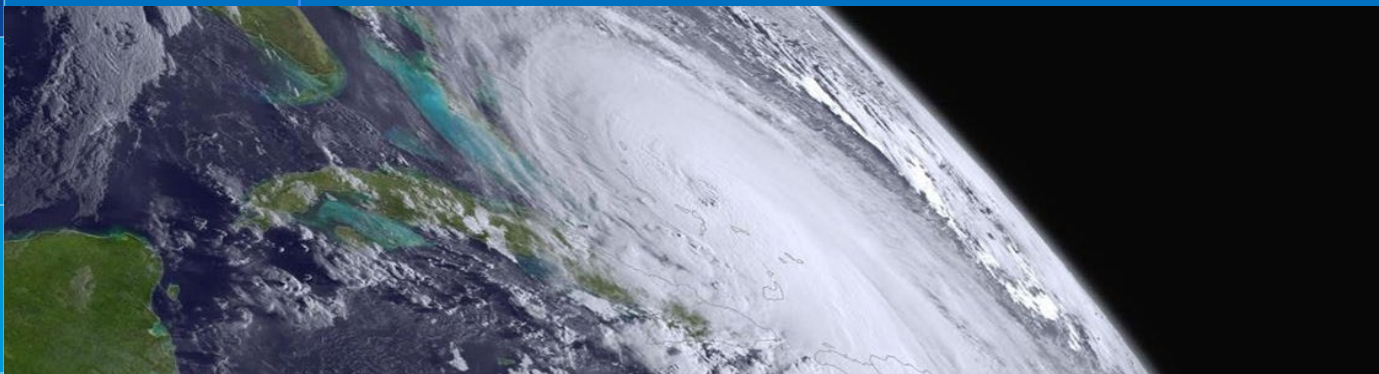
**National
Environmental
Satellite, Data, and
Information Service
(NESDIS)**

The Environmental Leadership Series: NOAA/NESDIS Transformation with Innovative Technologies

Presented by:

Irene Parker

Assistant Chief Information Officer - Satellites





Abstract



NOAA's mission is complex and ever changing due to the availability of new commercial and partner data sources that can enhance the quality of scientific research and numerical weather prediction. This seminar will outline efforts underway to leverage innovative technologies such as commercial cloud and artificial intelligence (AI) to transform how NOAA/NESDIS does business, and can provide more timely and cost effective products and services to the public and its' partners.





Agenda



- NOAA Cloud Vision & Overview of NOAA End-to-End Product and Data Management – *Section 1*
- NESDIS Common Cloud Framework & Services – *Section 2*
- Role of Data Management – *Section 3*
- Artificial Intelligence & Machine Learning – *Section 4*
- Closing Remarks - *Section 5*





Section 1:

NOAA Cloud Vision & Overview of NOAA End-to-End Product and Data Management





NOAA's Cloud Vision

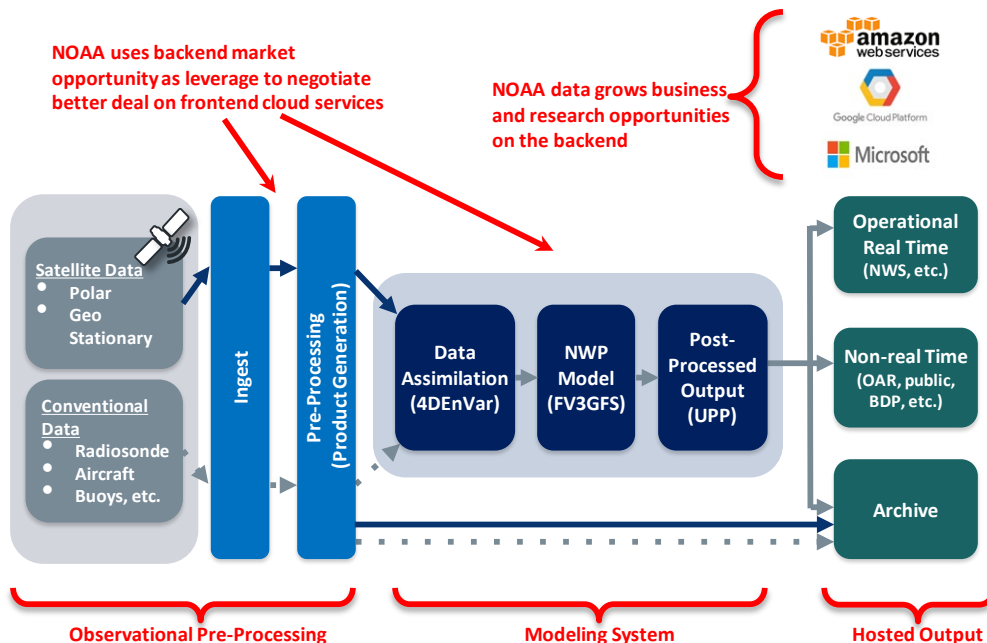


Identify and **migrate** suitable IT services to a commercial cloud computing environment that supports:

- Reducing costs
- Increasing IT provisioning speed
- Unlimited seamless scalability
- Maintaining high level of security

Key Tenet of Vision: Ensure that NOAA Numerical Weather Prediction enterprise is no longer constrained by hardware resources, which enables us to produce the preeminent global model

NOAA's Business Proposition Strategy of Moving to the Cloud



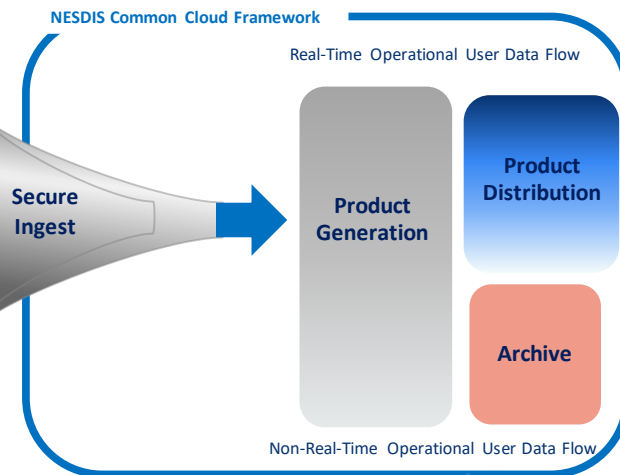


NESDIS Future Vision for Common Products & Services

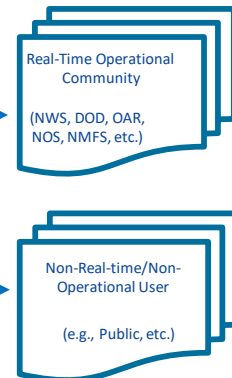
DATA SOURCES



NESDIS Common Cloud Framework



USERS





Motivation for a NESDIS Common Cloud

The NESDIS Ground Enterprise (NGE) is comprised of a disparate set of systems across all offices

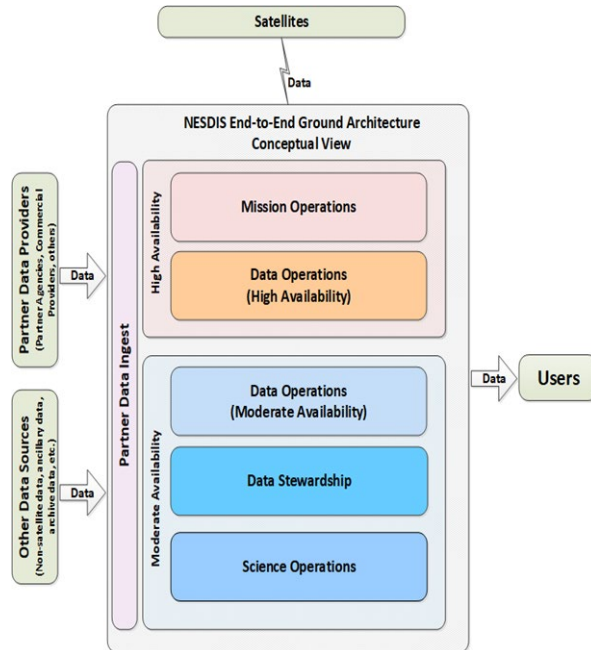
- Separate approaches support operations, science, and archive
- Duplicative functions and data from system-specific standards

The NGE is unsustainable - we will overrun our IT sustainment budgets by 2024 with the current approach

Therefore, **NESDIS must transform its business operations approach** to accommodate growing observing system data volumes

- Need a secure and highly available environment to support operational demands
- The environment must scale to meet new volume needs and be flexible enough to leverage all types of data
- Mission Operations functions are not part of the scope

Why not a service-based enterprise commercial cloud solution based on industry best practices?





Key Takeaways



- NOAA needs to get all of its environmental, satellite, and other data into a common commercial cloud repository
- The common cloud infrastructure to manage the common repository will be vertically and horizontally scalable to meet current and future organizational requirements – *(section 2.0 NESDIS Common Cloud Framework NCCF)*
- Essential to enable robust discoverability and dissemination is to structure and manage the data properly *(section 4.0 Role of Data Management)*
- This will enable the meaningful application of innovative tools & techniques such as Artificial Intelligence (AI) and Machine Learning (ML) to further enhance data discoverability; climate change analytics; and weather forecasting *(section 5.0 Artificial Intelligence & Machine Learning)*





Section 2:

NESDIS Common Cloud Framework (NCCF)





NESDIS Common Cloud Framework (NCCF)



The NCCF is a suite of services that allows NESDIS to provide end-to-end ground capabilities

- Cloud agnostic common enterprise architecture that enables NESDIS business functions
- Provides secure infrastructure, tools, and services to meet NESDIS user needs
- The solution is:
 - **Secure:** FISMA compliant FedRAMP Moderate cloud services
 - **Fault-Tolerant:** Redundant and highly available services lead to robust applications
 - **Scalable:** Capacity to accommodate all current and future workloads
 - **Data Agnostic:** Enable any data type and workflow within the framework
 - **Decoupled:** Interchangeable and independent services
 - **Cloud Agnostic:** Workloads and services can run in any Cloud Service Provider (CSP)
 - **Resources On-Demand:** Rapid provisioning of services based on business need
 - **Agile:** Supports agile processes with DevOps



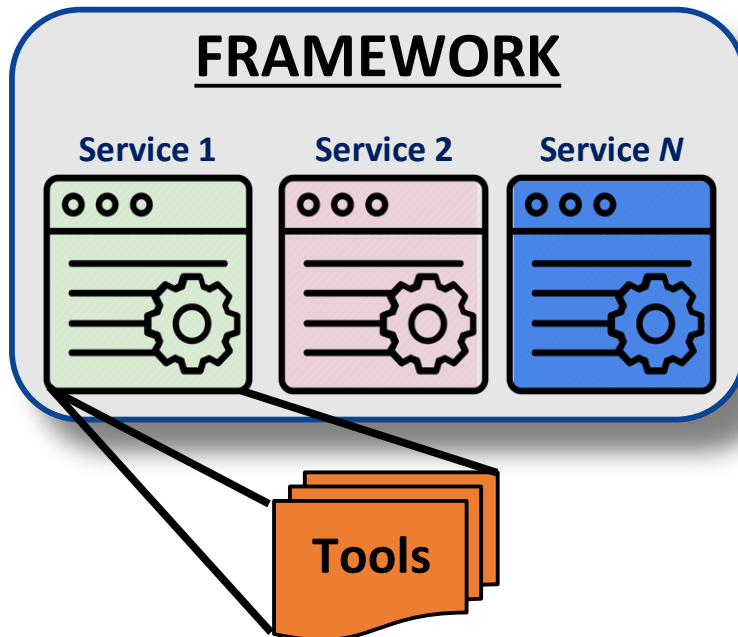
We have demonstrated that the NCCF works over a *representative set of algorithms* and workflows to meet organizational and mission requirements



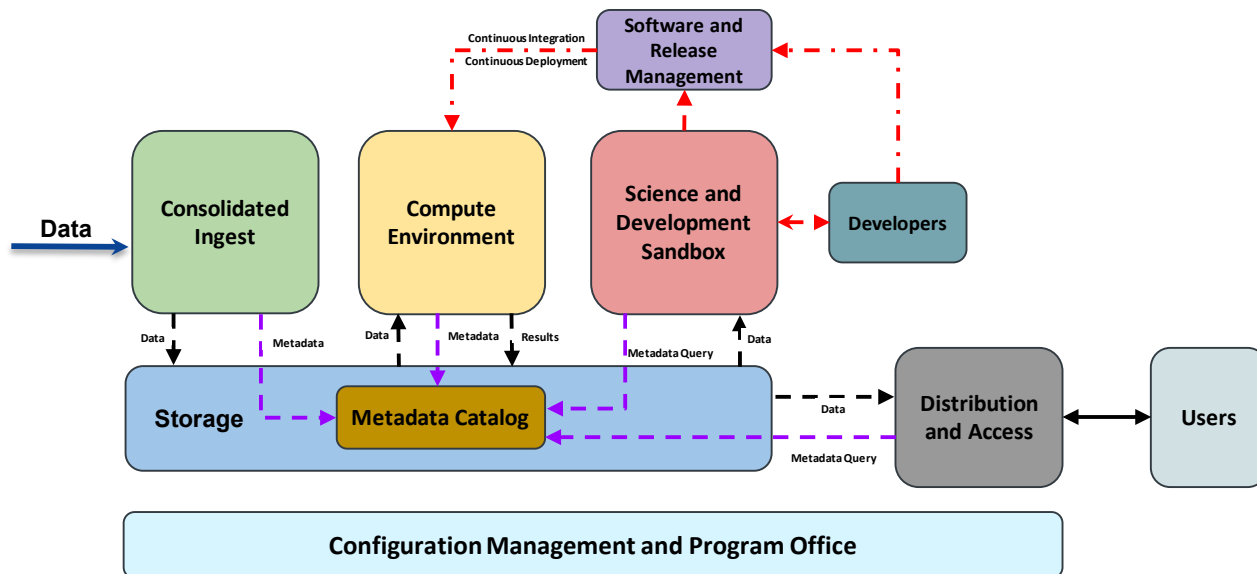


NESDIS Cloud Lexicon

- **Framework:** enables a set of services to work together to deliver the mission value
- **Service:** how to meet the core NESDIS IT functions
- **Tool:** cloud software application(s) used to implement the service



NCCF Services Overview





NCCF Services



<u>NCCF Service</u>	<u>Description</u>
Consolidated Ingest	Ingest service that securely ingests the data, serving as the entry point to the NCCF
Storage	Foundational service that centrally stores, manages, and archives all the data in the NCCF
Metadata Catalog	Elastic centrally managed metadata repository to quickly search the data
Compute Environment	Fault tolerant High Performance Computing (HPC)-based scalable and flexible compute functionality, that includes support for the Product Generation (PG) function with integrated orchestration and processing services
Science and Development Sandbox	Common development environment for developers and scientists to develop and test scientific algorithms, visualize data, and conduct basic research
Software and Release Management	Set of tools that support common configuration management practices and allow Continuous Integration/Continuous Deployment pipelines
Distribution and Access	Data agnostic and flexible service that supports access and dissemination of data to NESDIS partners and consumers





NCCF - A Working Solution



- NESDIS has an approved/documented security approach and ATO process to enable the NCCF to go to operations
- We have demonstrated that the common NESDIS approach works
 - Implemented a **Minimally Viable Product (MVP)** in 7 months
 - Represents a successful demonstration of consolidated ingest, storage, metadata catalog, compute, and distribution and access services in near real-time
- This is a **framework of services** to support the NESDIS mission
- Capability of a particular service in the NCCF expands and changes based on NESDIS mission needs
- Secure, scalable, fault tolerant, data source agnostic solution capable of processing near real-time data flows (working code!)
- **Services are built to evolve** and we are enhancing them with an eye toward operationalization





NCCF Independent Technical Evaluation



- Objective to validate that the NCCF met stated success criteria in both design and demonstrated functionality
- Review panel consisted of experts from USAF, NASA, USGS, OAR & NOAA OCIO
- Panel issued 17 Request for Action (RFAs) [8 Green, 9 Yellow, 0 Red]

Notable Strengths:

- NCCF could be foundation cloud architecture for other LOs
- Architecture strongly supports science computing requirements with integrated HPC stack based services
- Demonstrates enterprise distribution capabilities
- Flexible, scalable storage architecture

Notable Weaknesses:

- Insufficient detail on how NESDIS will address workforce transformation issues
- Cost control measures & tools are far along but need to mature to support operations
- Data boomerang or hair pinning effect with respect to support TIC requirements – requires further analysis of cost and latency implications to store and retrieve data





Section 3:

The Role of Data Management





Data Management



- For NOAA & public to maximize the value of NOAA data, the data must:
 - Be in cloud optimized formats;
 - Be analysis ready data; and
 - Have optimized metadata with appropriate tagging
- These attributes will facilitate more meaningful discoverability and value for the scientific and novice user population
- This will enable new innovative tools & techniques such as AI and ML to be applied to realize enhancements of our on-going research & development and numerical weather forecasting



Data Management



	Real Time Data	Monthly	Yearly	Retrospective Data
NOAA Products	Weather Prediction	Seasonal Forecasts	Trends, Maps and Indices	Reanalysis models
Winds	<ul style="list-style-type: none"> NOAA Weather Monitoring Commercial Weather Use Disaster and Risk Mgmt 	<ul style="list-style-type: none"> Insurance industry to assess disaster event impact Agriculture Industry NOAA Seasonal Forecast 	<ul style="list-style-type: none"> Insurance industry to assess disaster event impact NOAA Forecast Improvements 	<ul style="list-style-type: none"> Renewable Energy Industry Agriculture Commodities Prediction
Precipitation	<ul style="list-style-type: none"> NOAA Weather Monitoring Commercial Weather Use Disaster and Risk Mgmt 	<ul style="list-style-type: none"> Insurance industry disaster event impact Agriculture Industry NOAA Seasonal Forecast Municipal Water Resource Mgmt 	<ul style="list-style-type: none"> Urban and Development Planning Agriculture Industry Municipal Water Resource Mgmt 	<ul style="list-style-type: none"> Insurance Industry Urban and Development Planning Improving Forecast for Disaster
Aerosols	<ul style="list-style-type: none"> Municipal Air Quality Aviation Industry 	<ul style="list-style-type: none"> Municipal Air quality Insurance industry to assess disaster event impact Agriculture Industry Aviation Industry 	<ul style="list-style-type: none"> Municipal Air quality Forest and Land Mgmt Agriculture Industry Aviation Industry 	<ul style="list-style-type: none"> Municipal Air quality Industrial production Aviation Industry Forest and Land Mgmt
Temperature	<ul style="list-style-type: none"> NOAA Weather Monitoring Commercial Weather Use Disaster and Risk Mgmt Aviation Industry 	<ul style="list-style-type: none"> Weather Risk Management (Trade Markets) NOAA Seasonal Forecast 	<ul style="list-style-type: none"> Agriculture Commodities Prediction 	<ul style="list-style-type: none"> Agriculture Commodities Prediction Renewable Energy Industry



Section 4:

Artificial Intelligence & Machine Learning



What is AI & ML



Artificial Intelligence (AI) is wide-ranging branch of computer science concerned with building smart machines capable of performing tasks that typically require human intelligence.



Machine learning (ML) is an application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. **Machine learning** focuses on the development of computer programs that can access data and use it to learn for themselves.



How is NESDIS Exploring AI & ML



- In September 2020 NOAA and Google entered into an Other Transactional Authority (OTA) agreement that will leverage NESDIS science and IT/cloud subject matter experts with Google's AI/ML world leading practitioners to explore if these tools can enhance our weather forecasting and research capabilities
- From a technical perspective the OTA's goal is to develop an AI based system to do multi-sources data fusion and assimilation to generate a rich analysis of multiple parameters, with user driven features. The analysis should be useful for nowcasting and numerical weather prediction.

What is an Other Transactional Authority (OTA):

With the passage of P.L. 115-423, the National Drought Information System Reauthorization Act of 2018, Congress authorized NOAA to set up Other Transaction Authority (OTA) agreements to partner with commercial, academic, non-profit, and government organizations in order to enhance the effectiveness of its data and satellite systems.

An OTA is a contractual arrangement between the Government and a commercial entity in which they jointly contribute to a work product or effort AND both contribute resources and funding



Goals: Google OTA Agreement



One: Explore if a relationship like this can yield benefits for NOAA (this is the first executed OTA at NOAA)

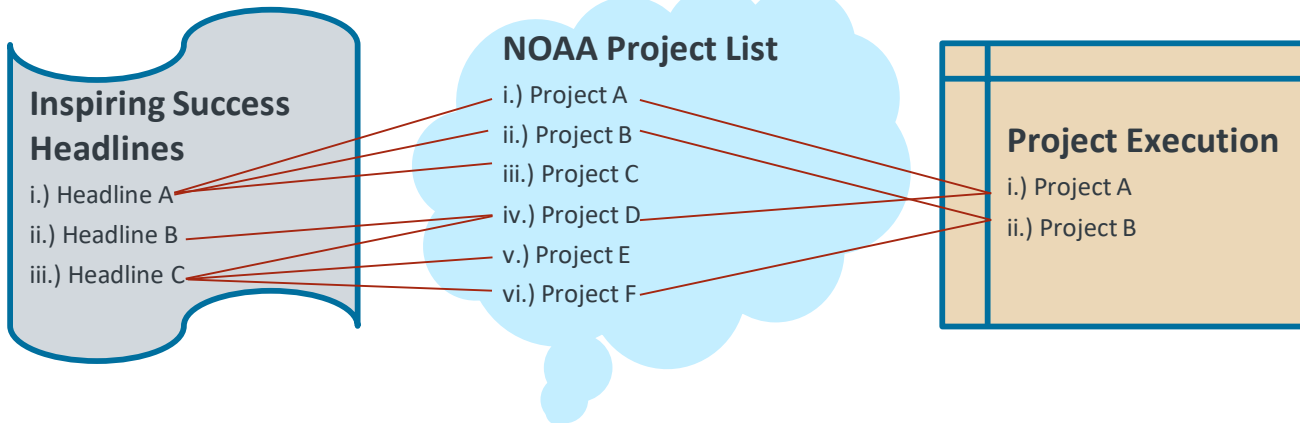
Two: Determine if we can enhance our weather forecasting and research capabilities with AI/ML tools & techniques. Some of the notable benefits that we expect from the agreement are:

- Gain direct exposure and training from Google AI/ML SMEs that may enhance our research capabilities
- Jointly develop usable code – NOAA will have an unlimited use of that source code
- Gain insights into how to better address our challenges with Big Data in ‘nowcasting’ and ‘numerical weather prediction’





Google OTA Project Framework



Inspiring Success Headlines:

- Combining AI and Physics to Improve Forecast Skills (including global model, hurricanes, storms, etc.)
- Using AI to increase the rate of data assimilation in NOAA models from 3% (current) to 30%
- Using AI to increase spatial resolution of forecast models (for example from county level to block level)
- Using AI to generate added-value info to society (help with drought, fire, flood, coast human activity, etc.)
- Using AI to help ingest crowd-source and citizens data to improve forecast skills





Section 5:

Closing Remarks





Closing Remarks



- a) NOAA's mission is complex, dynamic, and has reliance on stove-piped systems and antiquated methodologies
- b) The availability of new partner and commercial data sources provide opportunities to greatly enhance numerical weather prediction and climate analytics
- c) New innovative technologies & techniques allow for NOAA to transform how we do business and provide more timely and cost effective services to the public and our partners
- d) Getting NOAA's data into the cloud is not enough, we need to structure and manage the data properly to enable robust discoverability & dissemination
- e) Future efforts will focus to transform the data into cloud optimized formats (*where appropriate*) to enable AI and ML tools





Questions / References



Questions?

References:

Google OTA:

<https://www.nesdis.noaa.gov/nesdis-google-artificial-intelligence-prototyping-initiative>

NOAA Cloud Strategy:

[https://nrc.noaa.gov/Portals/0/Final Cloud Strategy.pdf?ver=2020-07-02-122459-813](https://nrc.noaa.gov/Portals/0/Final%20Cloud%20Strategy.pdf?ver=2020-07-02-122459-813)

NOAA Artificial Intelligence Strategy:

<https://nrc.noaa.gov/LinkClick.aspx?fileticket=0l2p2-Gu3rA%3D&tabid=91&portalid=0>



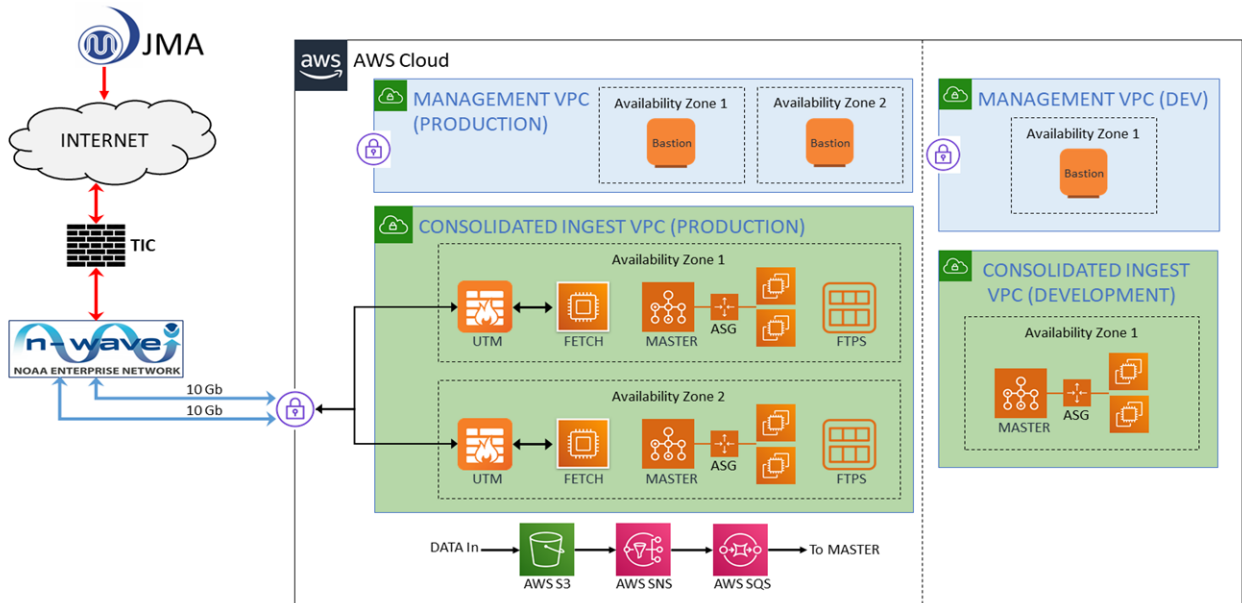
Back-Up Slides



Consolidated Ingest Service



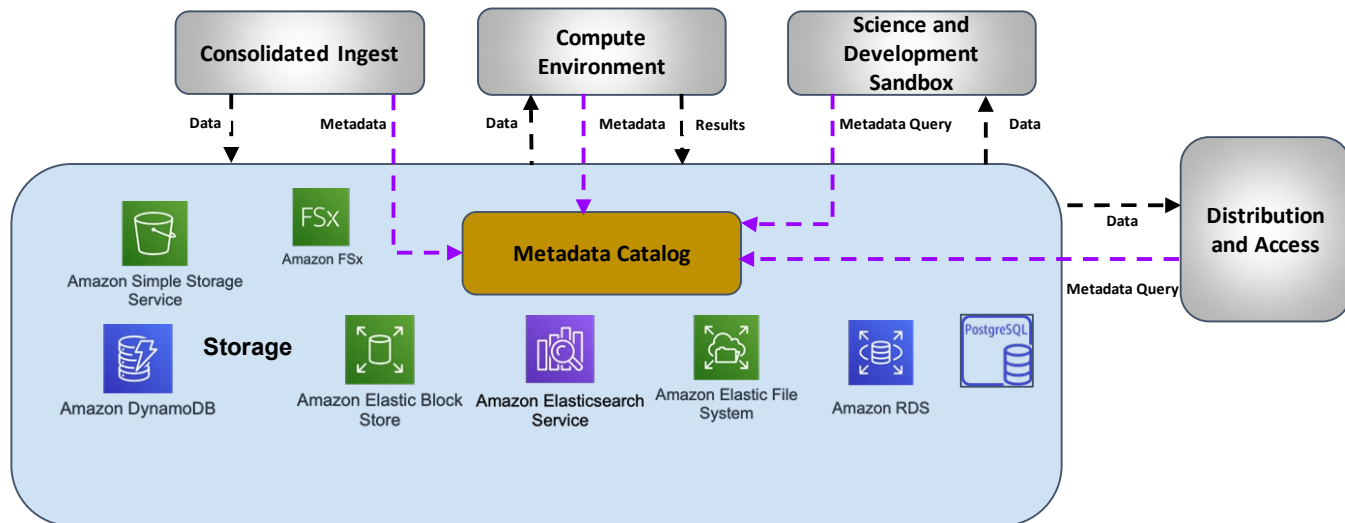
- Single point of entry for all NESDIS data types and protocols
- Customized security screening based on source and business agreements
- Highly available, fault tolerant, and scalable gateway to the NCCF



Enterprise Storage Service



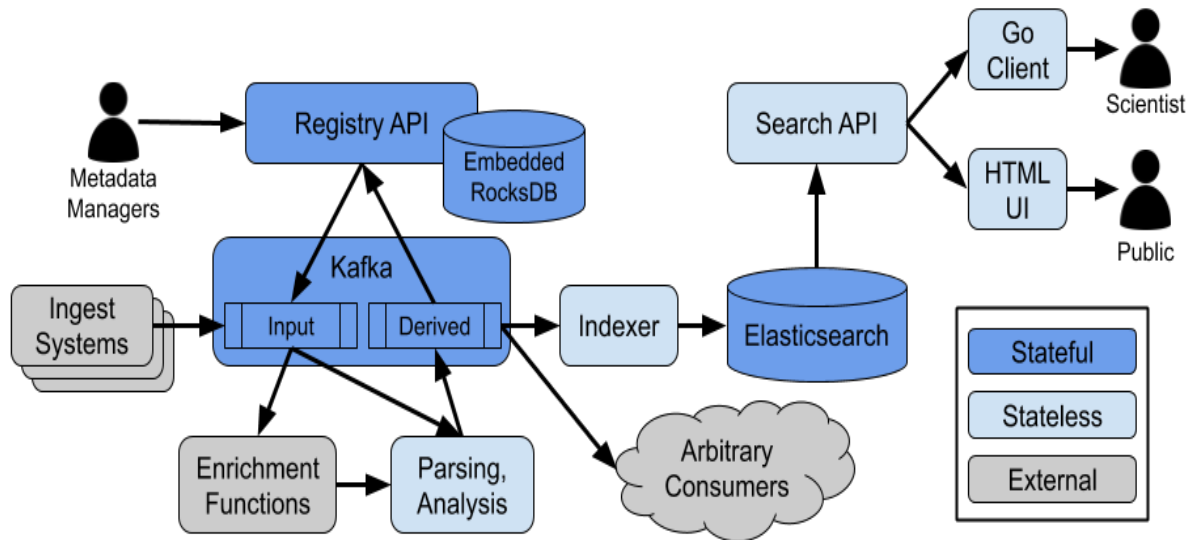
- Scalable and centralized data repository capable of storing all types of NESDIS data
- Supports the full data lifecycle - from ingest through archive
- Leverages all types of cloud vendor provided storage and underpins the entire NCCF



Metadata Catalog Service

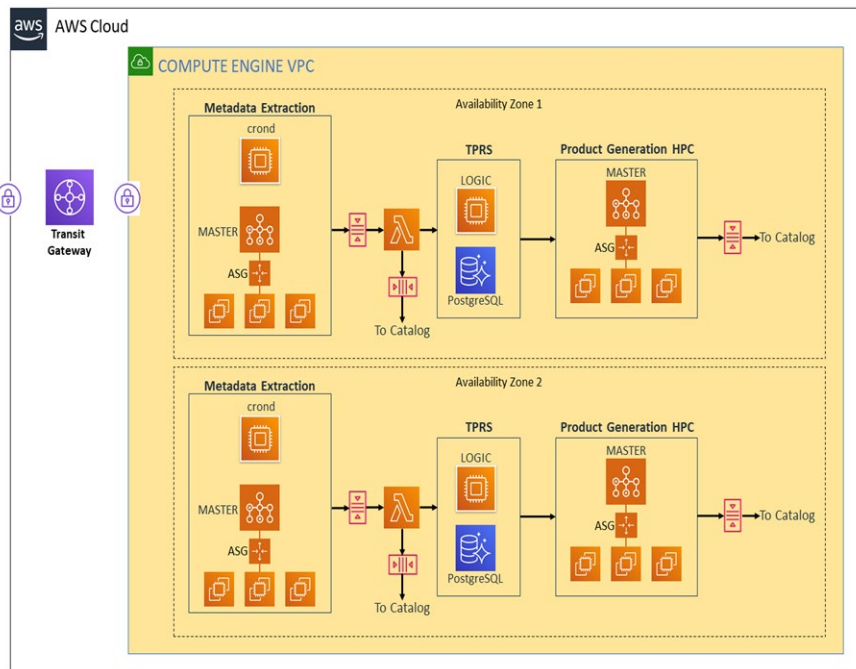


- Search and discovery approach to support all data within the NCCF
- Enables the efficient identification of datasets and associated metadata
- Designed to support data stewardship requirements





Compute Environment Service



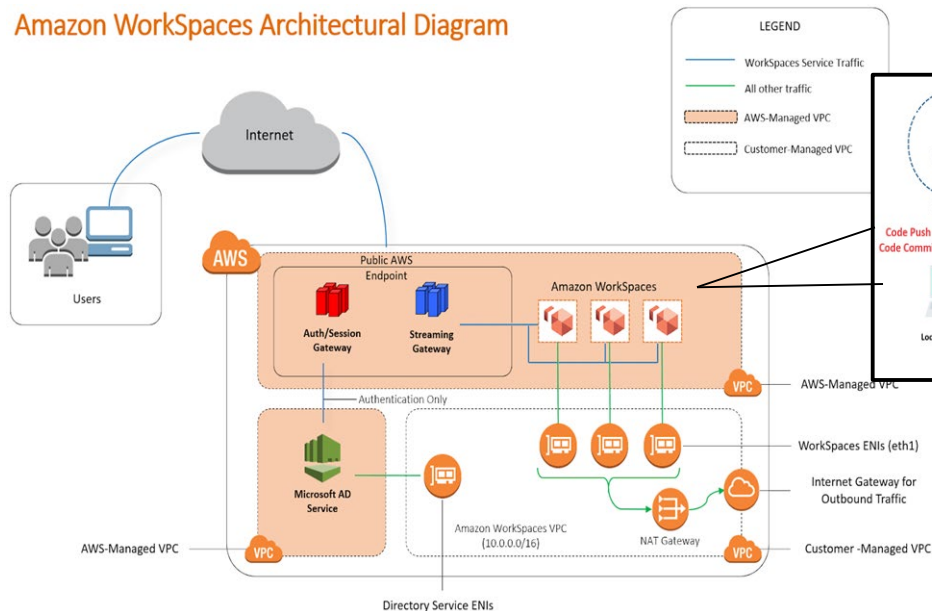
- Scalable, fault tolerant, and data agnostic service
- Supports all NESDIS compute needs
- Supports scientific data processing and development
 - **17 science algorithms** integrated into NCCF from STAR and NCEI
 - Scalable infrastructure to support current and future product generation needs



Science & Development Sandbox Service

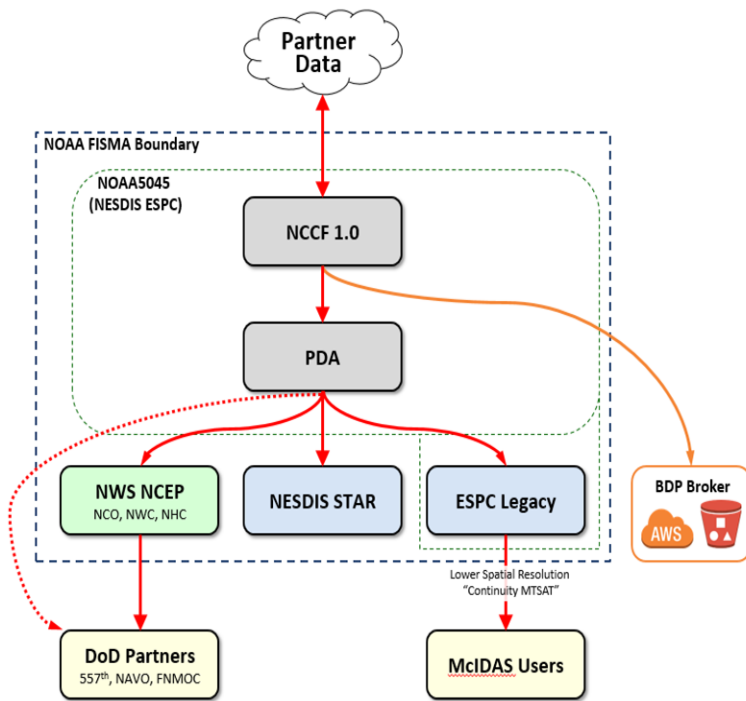
- Secure and scalable scientific development environment
- Provides streamlined integration between developers and scientists
- Supports data visualization tools

Amazon WorkSpaces Architectural Diagram





Distribution & Access Service



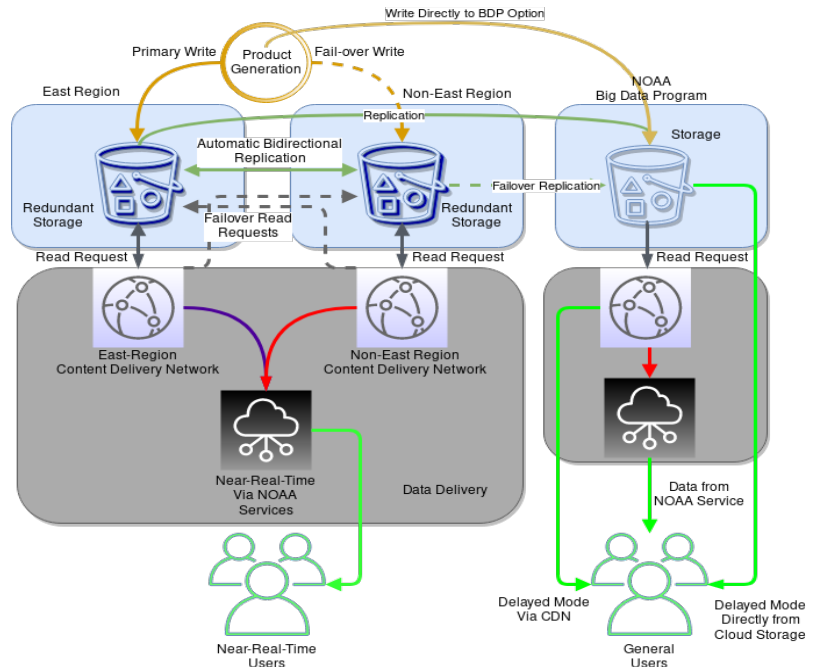
- Provides all required NESDIS data delivery pathways:
 - Low-latency push/pulls
 - Cloud-to-cloud transfers
 - Public access capabilities
- We will continue to leverage the PDA system until the NCCF distribution service is complete

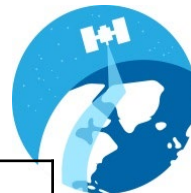




COOP Ready Dissemination Architecture

- Demonstrated and tested data redundancy and high availability (HA) using Cross-Region Replication (CRR)
- Demonstrated dissemination with **99.9% uptime** using Content Delivery Network (CDN)
- This is a Fault Tolerant and HA Dissemination Architecture suitable for disaster recovery (DR)





Acronyms

Acronym	Description
AI	Artificial Intelligence
AWS	Amazon Web Services
API	Application Programming Interface
BDP	Big Data Projects
CSP	Cloud Service Provider
CDN	Content Delivery Network
CRR	Cross Region Replication
DR	Disaster Recovery
HA	High Availability
HPC	High Performance Computing
ML	Machine Learning
MVP	Minimally Viable Product
NCCF	NESDIS Common Cloud Framework
NCEI	National Centers for Environmental Information
NGE	NESDIS Ground Enterprise
NWS	National Weather Service
OAR	Ocean and Atmospheric Research

Acronym	Description
OAR	Ocean and Atmospheric Research
OTA	Other Transactional Authority
PDA	Product Distribution & Access
PG	Product Generation
RFA	Request for Action
SME	Subject Matter Expert
STAR	Center for Satellite Applications & Research
UI	User Interface
VPC	Virtual Private Cloud