NOAA OCEANIC AND ATMOSPHERIC RESEARCH (OAR)

CLOUD COMPUTING STRATEGY

Released: February 2021

OAR Cloud Computing Strategy

1. Executive Summary

The National Oceanic and Atmospheric Administration's (NOAA) Office of Oceanic and Atmospheric Research (OAR) conducts an integrated program of research, development, and services to improve the understanding of the Earth system and support the delivery of world-class science. Cloud computing has the potential to radically advance OAR's capabilities and transform collaboration with partners. The OAR Cloud Computing Strategy vision is to implement cloud-enabled, mission-centric capabilities whenever possible. This strategy will leverage key benefits of the cloud such as dynamically scalable computing resources, consistent easy-to-use tools and techniques which speed up innovation, improve security, and provide cost benefits due to economy of scale. This strategy envisions a secure, scalable, cost effective, and data-source agnostic path to forward the four goals and objectives outlined in the OAR Strategy 2020-2026[1].

Additionally, research-specific strategic goals were developed to facilitate an overarching OneNOAA vision of cloud-enabled innovation, and to guide the initial evolution of cloud capabilities for OAR. While cloud may be a viable, optimal solution for many OAR technologies, there is a continued need for on-premise HPC research infrastructure to emulate the NOAA operational HPC environment for testing, evaluation, and transition. The path forward for OAR entails a hybrid on-premises/cloud approach involving a cycle of experimentation, evaluation, procurement, and adoption toward cloud where applicable, combined with workforce development and training.

Through demonstration and assessment of capabilities across key thematic program areas, OAR expects the adoption and utilization of cloud capabilities to increase agility and accelerate the transition of research and development (R&D) outputs (i.e., Research-to-Applications or R2X) into NOAA's operations, applications, commercialization, and other uses for societal benefits in a cost effective manner.

Through this vision OAR will empower its research endeavours with streamlined access to cloud capabilities, today and in the future.

2. Introduction

OAR's scientific research is foundational to the core NOAA mission and priorities of reducing impacts of extreme weather and water events, exploring and understanding the ocean, maximizing the sustainable economic contributions of the American blue economy, understanding climate change, and advancing the rate and breadth of community innovation in numerical modeling, data assimilation and analysis, and machine learning. Scientific computing is a critical tool in achieving NOAA goals, and is currently performed mostly on agency-furnished assets. OAR leverages new technologies and advanced computing to more efficiently and effectively accomplish NOAA's strategic goals. Cloud computing brings various potential benefits that warrant further exploration under this strategy such as agility, economies of scale, and data management efficiencies. The agility offered by the cloud is expected to provide added flexibility in fulfilling OAR's mission while also sustaining an important aspect of how OAR engages with the external (non-NOAA) community across many applications.

OAR will expand its engagement with the external community to maintain awareness of new cloud technology and explore innovative ways to acquire and use it. OAR will work across NOAA to develop a requirements-based approach to cloud computing, accelerate its adoption where applicable, and invest in infrastructure to advance Earth-system modeling and enable next-generation research. The need for on-premise HPC research infrastructure is required to emulate the NOAA operational HPC environment and will continue to be required in the foreseeable future.

Migration to the cloud is strongly encouraged by federal policy intended to accelerate the pace at which it will realize the value of cloud computing by requiring agencies to evaluate safe, secure cloud computing options before making any new information technology (IT) investments [2,3]. This document presents a strategy consistent with federal policy and by which OAR will leverage cloud computing technology to transform OAR applications and models as applicable. The specific approaches by which targeted applications and capabilities will evaluate and migrate on merit from on-premises systems to cloud-based services to meet OAR needs will be outlined subsequently in a separate implementation plan.

3. What is cloud computing for OAR?

> Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction[4].

The cloud architecture presents an environment that challenges conventional software development and hosting paradigms[4]. With a wide range of turn-key solutions presented by cloud service providers (CSPs), development targeted for cloud deployment allows researchers to focus efforts on the application and science objectives, and focus less on on-premise computing infrastructure and its costs. The opportunities presented by cloud hosting will be a catalyst for innovation as legacy applications destined for cloud hosting are reimagined and novel approaches to address new scientific computing needs are developed.

Migration to a cloud-based environment can provide key benefits such as efficiency, agility, innovation, surge capacity, and potential cost reduction. OAR's vision is to implement cloud-enabled capabilities whenever possible to further its goals and objectives[1]: 1) Explore the Marine Environment; 2) Detect Changes in the Ocean and Atmosphere; 3) Make Forecasts Better; and 4) Drive Innovative Science. The guiding principle for how OAR will leverage cloud computing technology is to demonstrate benefits and risks across its R & D portfolio. Six themes which encompass OARs breadth of activities will provide paths to realistically and objectively evaluate the potential for cloud computing.

While the adoption of cloud computing offers multiple potential benefits to OAR, it also presents critical challenges that must be considered when evaluating and deciding to use cloud computing. Noteworthy challenges include the use of vendor-specific software or proprietary customization that can result in vendor "lockin", costs associated with accessing and retrieving data ("egress/exfiltration"), potential transfer to other cloud vendors, and Federal Information Security Management Act (FISMA[5]) impacts on access. Project managers and developers considering using the cloud need to know how storage, networking, security, monitoring, funding, cloud porting, and provisioning will work. These must be in place in order to plan and budget for any cloud application, regardless of whether such cloud is private or public. OAR must do a thorough analysis of these challenges. Government and industry best practices, as well

as collaboration with other federal agencies, should be leveraged to develop efficient and effective plans for cloud adoption [2,6].

Successful cloud migration relies upon an enhanced understanding of the specific needs, challenges, and requirements informed by the spectrum of R & D activities at OAR. OAR will implement cloud computing such that all capabilities will be in compliance with federal security, data and acquisition regulations [7,8,9,10].

The path forward on cloud migration relies upon assessing potential benefits and risks to specific OAR themes derived from a use case analysis. Common to these themes is continued or increased engagement with community development projects with the intention to eventually transition projects within these themes from R2X and X2R based on demonstrated benefits, risks, and costs.

- **A Cloud-hosted Development Environment** will enable users to collaboratively work on model and application development. This common development environment includes hosting a JupyterHub[11] platform and using remote workstation capabilities with access to the necessary development and/or application tools, including code repository, compile, test, and continuous integration/deployment capabilities.
- **High Performance Computing (HPC)** which involves highly parallel processing, is critical to OAR for development and testing of Numerical Weather Prediction (NWP) models, data assimilation, genome assembly and analysis, highly computational applications, surge computing demands such as rapid reanalysis and reforecast production, as well as processing data quickly and efficiently.
- **Machine Learning (ML)/Artificial Intelligence (AI) Cloud Capabilities** provide an opportunity for OAR to efficiently access high-memory hardware and leverage ML/AI software to modernize and improve NOAA products and services, as well as share lessons learned that could be valuable for many groups inside and outside OAR.
- **Data Analysis** is a key function to OAR research and development. Cloud capabilities allow the co-location of processing with NOAA's environmental modeling and data assets. As the volumes and data and model analysis results increase this co-location reduces the need to migrate large volumes across networks and adds efficiencies towards extracting value from the high cost NOAA assets.

- **Data Management** includes hosting NOAA environmental data in the cloud for availability to shared resources and other tools. Cloud solutions will be explored to manage data workflows from acquisition (in-situ or derived) for use by research applications, and archival.
- **Migrating Applications and Websites to Cloud** will be necessary as part of the transition to cloud computing. While migration will be somewhat specific to individual applications, knowledge gained and lessons learned about benefits of the cloud versus on-premises capabilities, along with best practices for adapting applications for optimal operations on the cloud, can be shared across OAR.

4. Strategic Goals

Scientific research, application and product development, and the successful transfer of this research and/or products to NOAA Operations or the private sector will continue to be the foundation of the OAR mission. OAR will align with and uphold this principle while implementing cloud functionality. To help position OAR for adoption and use of cloud computing resources in alignment with the NOAA Cloud Strategy[12], the following goals have been identified. OAR will collaborate with NOAA Office of the Chief Information Officer (OCIO), Line Offices and other Enterprise cloud initiatives (e.g., NOAA Big Data Program), to maximize the benefits of cloud adoption for internal use, to ensure cohesive cloud approaches across NOAA, and to accelerate the research to operations (R2O) process.

a. OAR will evaluate the benefits and adequacy of and pathways (e.g., rehosting, replatforming, refactoring) for using cloud to foster both R2O and operations to research (O2R), using pilot projects, trade studies, cost-benefit analysis, and enterprise risk assessment.

b. OAR will explore and pursue the use of cloud-native capabilities to enhance enterprise compute and to reduce costs in data storage and analysis capabilities.

c. OAR will help define a high performance computing procurement and implementation strategy that allows NOAA to evaluate costs and benefits for classes of applications on cloud versus on-premise hosting.

d. OAR will provide effective governance by defining roles, responsibilities, and procedures both internal to the organization and within the cloud.

e. OAR will foster a cloud-ready workforce by identifying training and other educational resources to ensure the workforce has the skills and competencies needed to increase cloud adoption.

f. OAR will reduce barriers towards the adoption of cloud-based services by establishing self-service capabilities and templates for easy access and provisioning of cloud development environments. The goal is a cloud development environment with low barriers of access for developers, scientists, and experimenters to utilize capabilities presented by this dynamic and rapidly evolving environment.

g. OAR will define requirements for the cloud architecture and foster implementations that ensure confidentiality, integrity, and availability of data, applications, and cloud services.

h. OAR will leverage best development and architecture practices to optimize cloud resources and maximize utility.

5. Path Forward with Cloud Computing

Success of cloud initiatives comprises more than just efficient use of cloud technology and the right deployment approach. Changes must happen in the following stages as the ability of OAR to leverage cloud matures:

1. **Experimentation -** The emerging OAR activities in the cloud will increase understanding directly through demonstration projects, training, and leadership support. OAR will support architectural experiments, pilot studies, and training that span the breadth of its R&D activities. OAR will collaborate closely with NOAA Line Offices using lessons learned and recommendations on cloud implementation based on their experiences to accelerate experimentation.
2. **Migration -** OAR will mature its cloud-enabled governance, technical, and operational business practices to effectively and efficiently migrate targeted elements to the cloud as appropriate. OAR anticipates a hybrid approach combining on-premise and cloud computing that begins to explore and realize the benefits of the cloud environment ahead of a stepped migration, and then subsequently through downstream cost and service optimization based on cloud pathway strategy analysis. Thus, support for on-premise and cloud computing will likely continue in varying proportions.
3. **Optimization -** OAR will use continuous improvement processes to optimize the R & D capabilities in the cloud. New cloud services can then be implemented to augment or replace existing services to add value. In this way, OAR can accurately quantify and manage use of cloud computing resources and subsequently develop the most cost-efficient utilization methodology[6].

4. **Workforce Development and Training -** OAR will invest in staff hiring, training and other educational opportunities to create a cloud-ready workforce.
5. **Knowledge Sharing -** OAR will use the NOAA Research and Development Database[13] to capture all R&D projects associated with the strategy as the primary tool to provide information across OAR and NOAA on projects designed to implement the Strategy.

Migrating to the cloud is an iterative process of trial and assessment that will evolve as OAR develops new skills, processes, tools, and capabilities. OAR will therefore continuously seek to refine and mature the cloud computing approach. Establishing a solid foundation is key to a successful migration, as initial migrations help build experience and depth of knowledge that will accelerate later migration efforts. The migration process should balance the business and technical efforts needed to complete a cloud migration.

Successful migration to the cloud also requires clear and continuous communications with key stakeholders. OAR will establish a robust cloud communications plan to keep key stakeholders engaged in the cloud migration efforts. Active participation and commitment of all OAR Laboratories and Programs is critical to ensure a consistent and optimized cloud migration.

Success of the Strategy will depend greatly on the resources allocated to workforce development and training as well as those purchased solely for experimentation and learning. Success also depends on maintaining and modernizing HPC and other on-premise systems so that existing R2X and X2R capabilities are not diminished as the strategy is executed.

6. Conclusion

OAR's vision is to empower its research endeavours with streamlined access to the capabilities offered by cloud service providers, both today and in the future and in line with the strategic goals formulated in this document. In the near term, the path forward for OAR entails a hybrid on-premises/cloud approach to emulate the NOAA operational HPC environment.

OAR’s cloud utilization relies upon assessing demonstrated benefits to specific themes such as cloud high performance computing, machine learning/artificial intelligence, software development, data analysis, data management and migration of applications

and websites to the cloud. OAR will seek to leverage lessons learned and guidelines developed by other NOAA units as well as collaborate as appropriate. OAR expects cloud adoption and utilization to increase agility and accelerate the transition of R&D outputs into NOAA's operations, applications, commercialization, and other uses for societal benefits (i.e., R2X) in a cost effective manner.

Cloud computing potentially brings a fundamental change to how OAR conducts its business. Therefore, for each key area of business impacted by cloud adoption, OAR will assess the migration readiness and adequacy, and create a migration plan to address gaps in staff skills and organizational processes. OAR acknowledges that migrating to the cloud is an iterative process that evolves as OAR develops new skills, processes, tools, and capabilities. OAR will continuously seek to refine and mature the cloud computing approach and maintain open communications with all OAR stakeholders.

7. References

1. OAR Strategy 2020-2026
2. Federal Cloud Computing Strategy (Cloud Smart)
3. White House Executive Order on Maintaining American Leadership in Artificial Intelligence (EO 13859)
4. National Institute of Standards and Technology (NIST), Special Publication 800-145 The NIST Definition of Cloud Computing
5. Federal Information Security Modernization Act (FISMA) of 2014
6. NESDIS Cloud Computing Strategy NESDIS-PLN-1120.1 Version 1.0
7. Federal Data Center Optimization Initiative (DCOI)
8. Foundation for Evidence-Based Policymaking Act (H.R.4174)
9. Federal Information Technology Acquisition Reform Act (FITARA)
10. Federal Risk and Authorization Management Program (FedRAMP)
11. https://jupyter.org/hub (JupyterHub)
12. NOAA Cloud Strategy Maximizing the Value of NOAA's Cloud Services
13. NOAA Research and Development Database

8. List of Acronyms

AI	Artificial Intelligence
CSP	Cloud Service Provider
FISMA	Federal Information Security Management Act
HPC	High Performance Computing
IT	Information Technology
ML	Machine Learning
NOAA	National Oceanic and Atmospheric Administration
NWP	Numerical Weather Prediction
O2R	Operations to Research

OAR	Office of Oceanic and Atmospheric Research
OCIO	Office of the Chief Information Officer
R&D	Research and Development
R2O	Research to Operations
R2X	Research transferred to undetermined end-user
X2R	R&D efforts

9. Acknowledgements

OAR Cloud Writing Team Members:
Krishna Kumar, Ph.D (Weather Program Office/Cherokee Nation)
Daniel Meléndez, Ph.D. (Weather Program Office and National Weather Service)
Jebb Stewart (Global Systems Laboratory)
Melissa Petty, Ph.D (Global Systems Laboratory)
Kirk Holub (Global Systems Laboratory)
Eugene Burger (Pacific Marine Environmental Laboratory)
Tom Hamill, Ph.D (Physical Sciences Laboratory)
Sundararaman Gopalakrishnan, Ph.D (Atlantic Oceanographic and Meteorological Laboratory)
V. Balaji, Ph.D (Geophysical Fluid Dynamics Laboratory)
Leah Dubots (Weather Program Office)

Members of OAR Cloud Tiger Team

OAR Cloud Tiger Team Sponsors:
Chair: Jennifer Mahoney (Global Systems Laboratory)
Vice Chair: DaNa L. Carlis, Ph.D (Global Systems Laboratory)
Kathryn Mozer, (OAR Office of Policy Planning and Evaluation)

(END OF DOCUMENT)