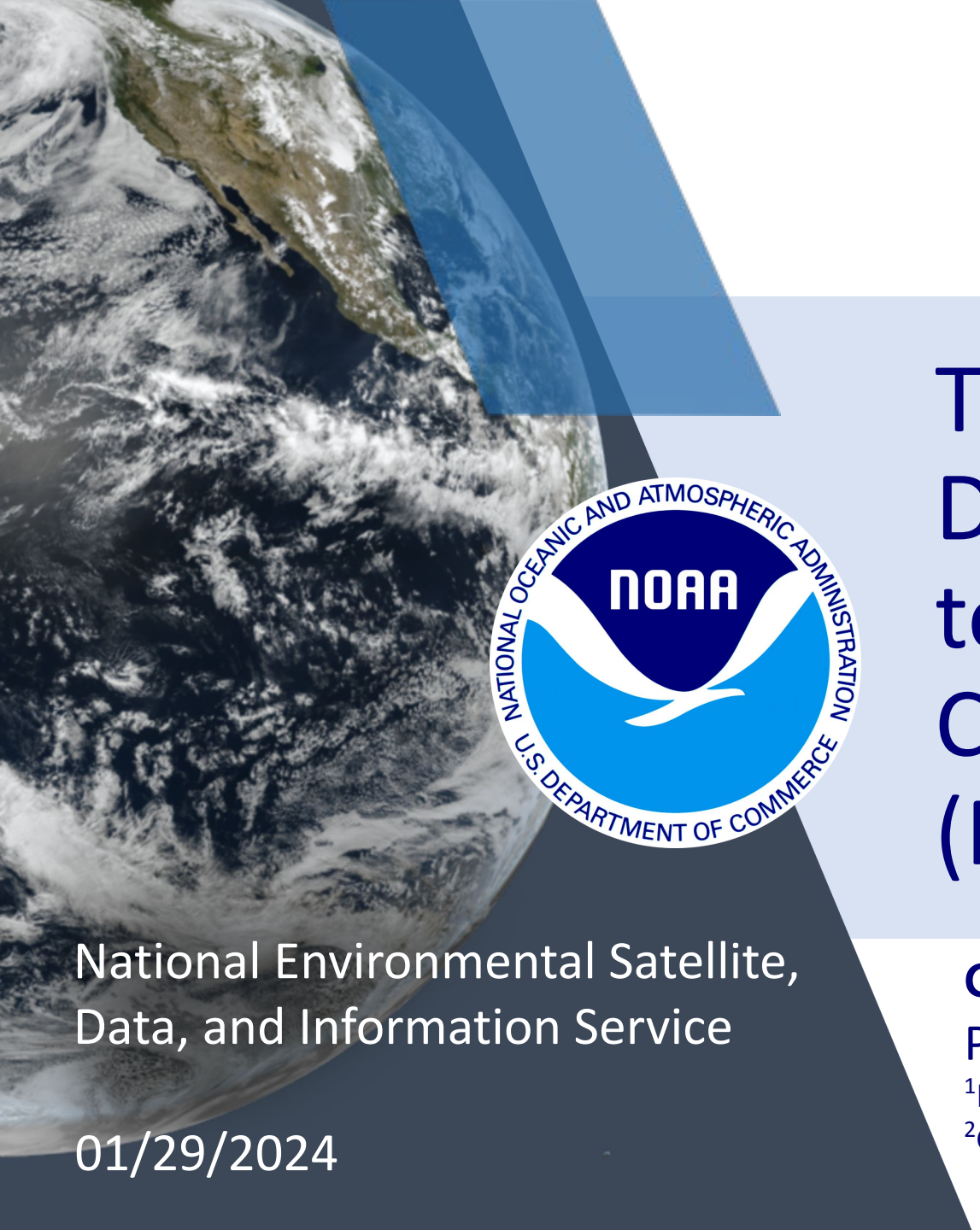


Transitioning Science Team Development From On-Prem to the Cloud via the NESDIS Common Cloud Framework (NCCF)

Gian Villamil-Otero[1], Joseph Hellmers[2], Melissa Zweng[1], Pura Perez[1], Walter Wolf[1]

[1]NOAA/NESDIS/OCS/PMD

[2]GAMA-1 Technologies

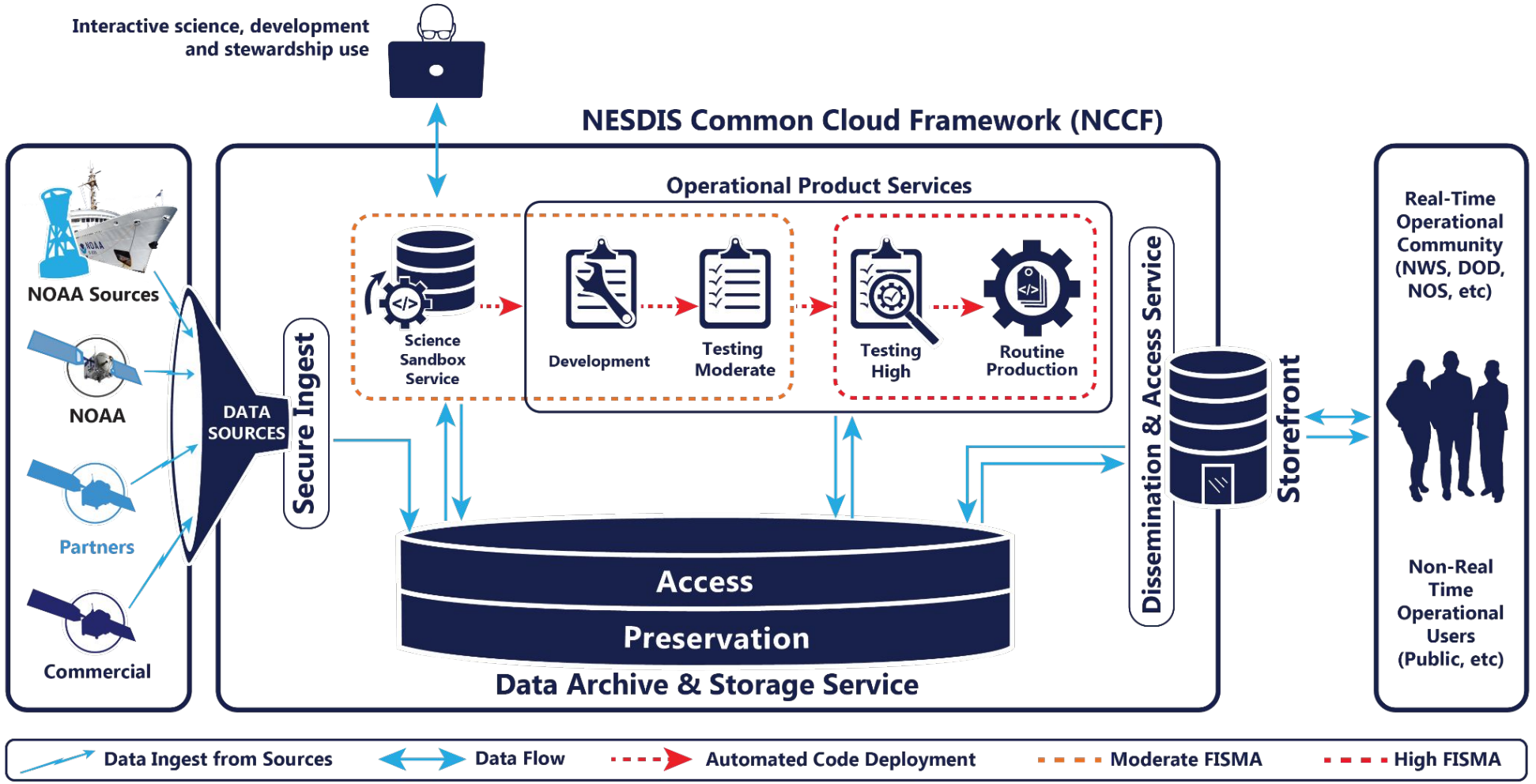
National Environmental Satellite, Data, and Information Service

01/29/2024

Background

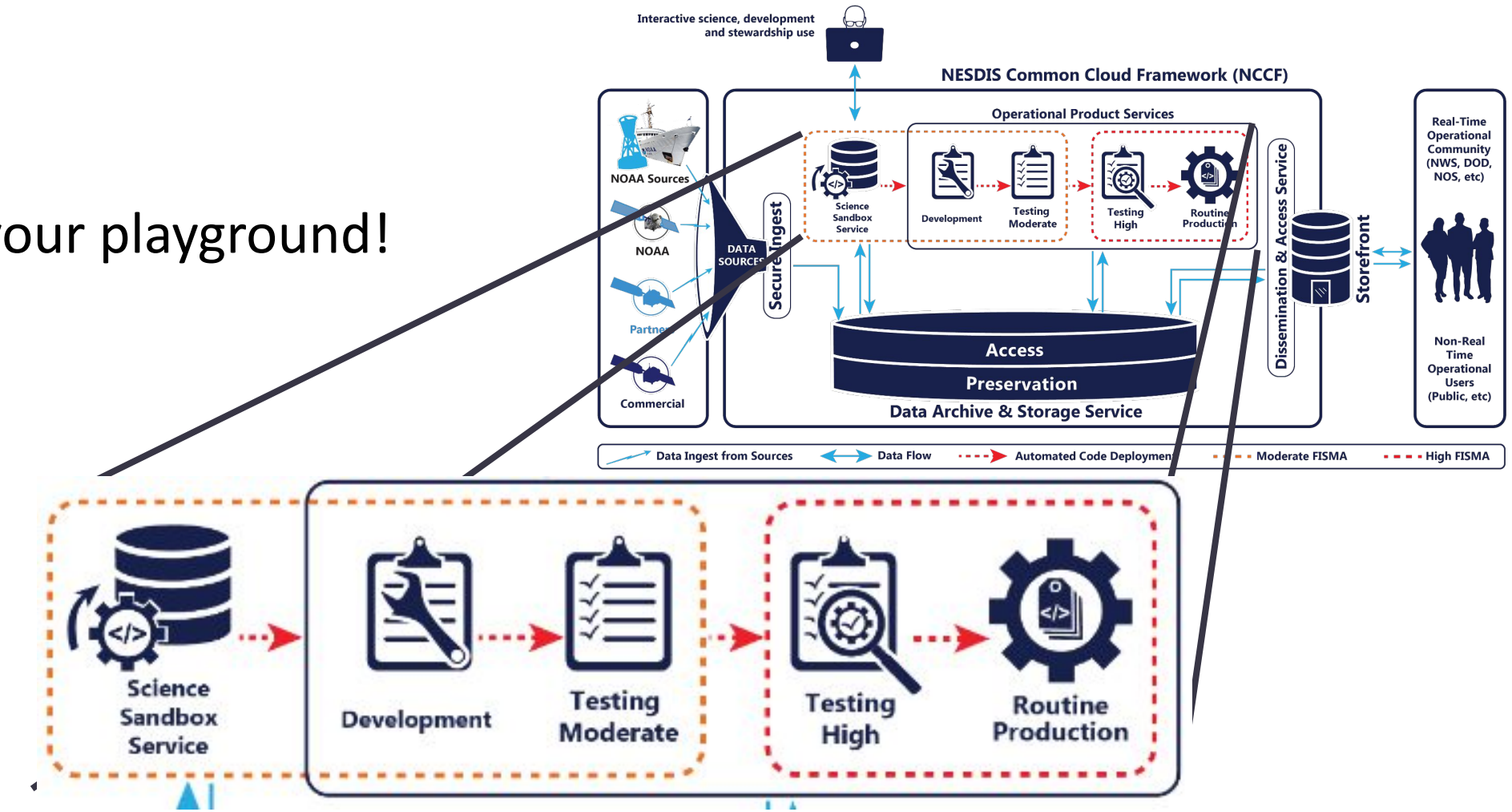
- NESDIS is implementing the the NESDIS Common Cloud Framework (NCCF) to support NESDIS data and science operations.
- Office of Common Services (OCS) works to maintain and advance agile, scalable, and efficient solutions to make sure that satellite data gets to users.
- The Product Implementation Branch (PIB) manages the actual implementation of science algorithms into the cloud.
- PIB provides application software engineering and lifecycle process reviews as well as coordinates user readiness for transition to operations.

What is the NCCF?



The Science Sandbox

The Sandbox is your playground!

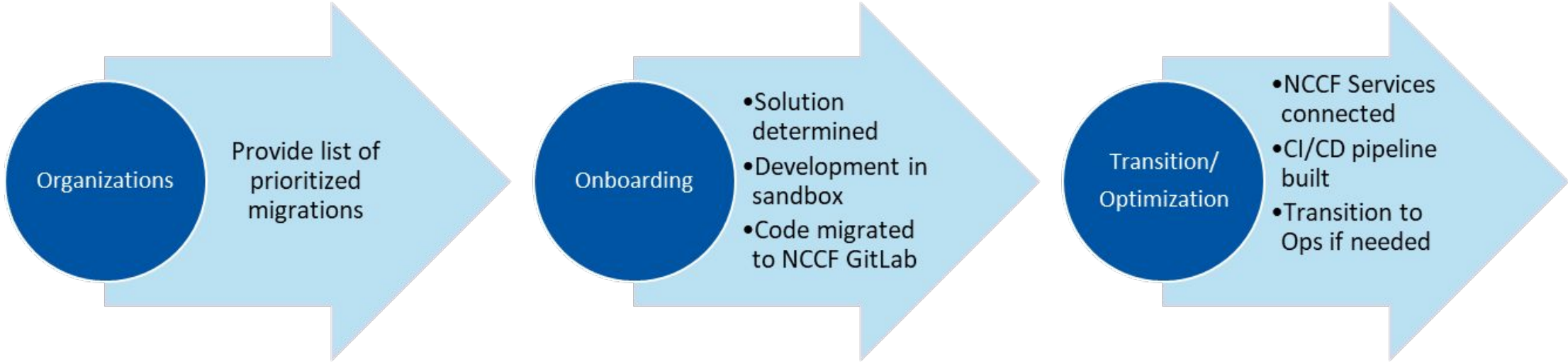


NCCF Services

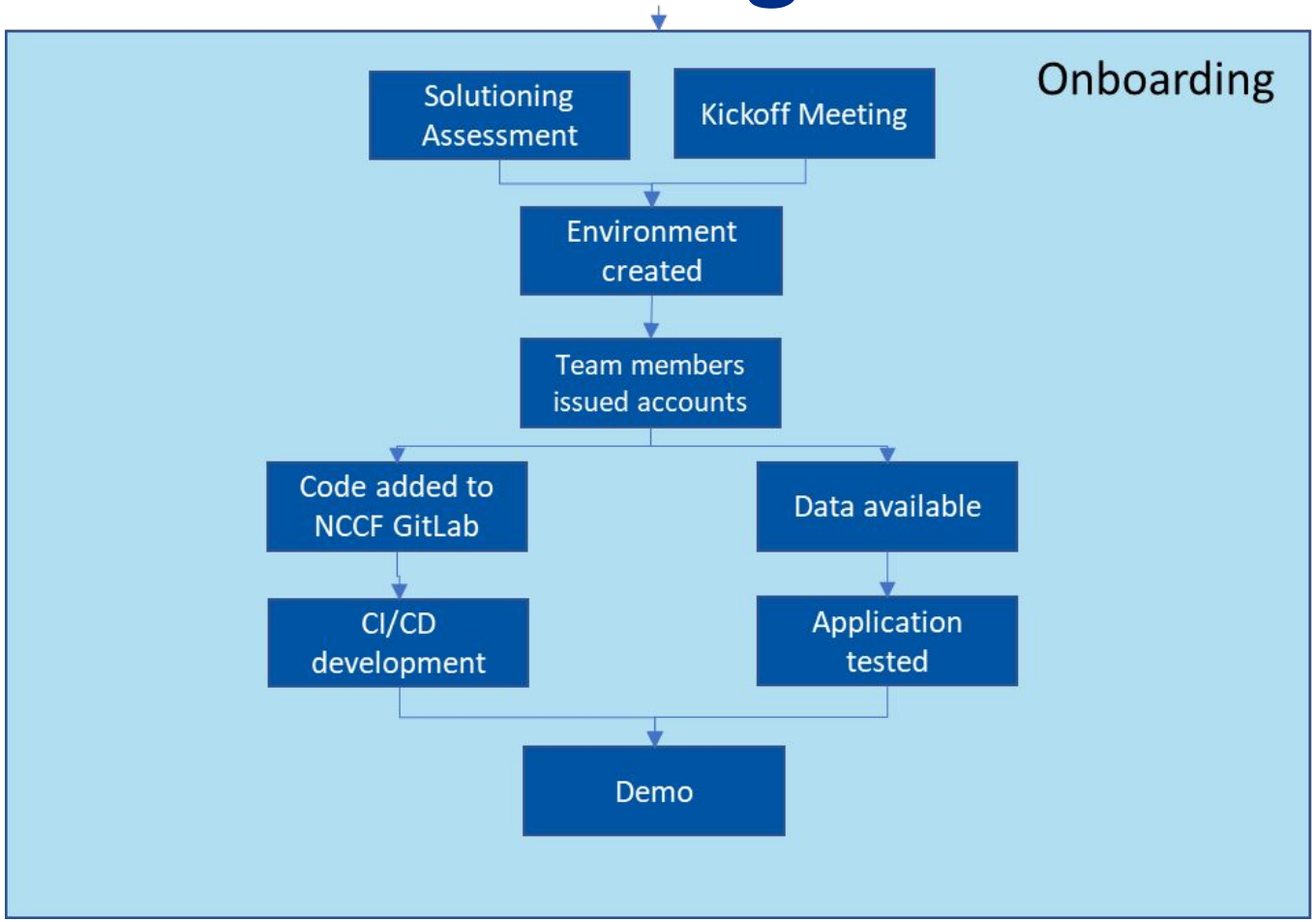
NCCF Service	Description
Consolidated Ingest	Ingest service that securely ingests the data, serving as the entry point to the NCCF
Storage	Foundational service that centrally stores, manages, and archives all the data in the NCCF
Metadata Catalog	Elastic centrally managed metadata repository to quickly 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

Transitioning of Science Teams

- Migration of science teams has been separated into 2 phases once teams are identified.
- Onboarding phase sets the team for working in the science sandbox.
- Transition/Optimization phase integrates the application with other NCCF services and optimizes it for speed and cost savings. This phase transitions the application to operations if applicable.

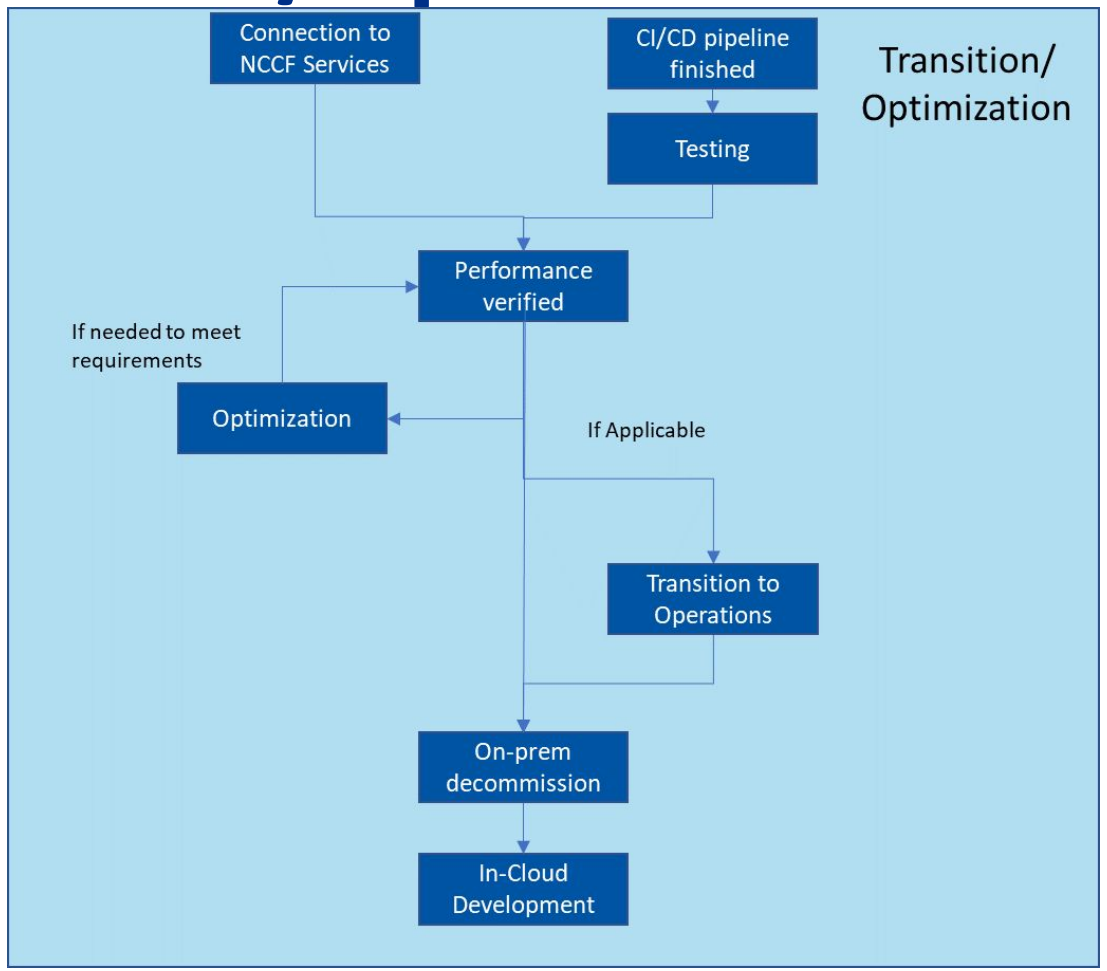


Onboarding Phase



Goal: Get the team in a position to do work in the Science Sandbox

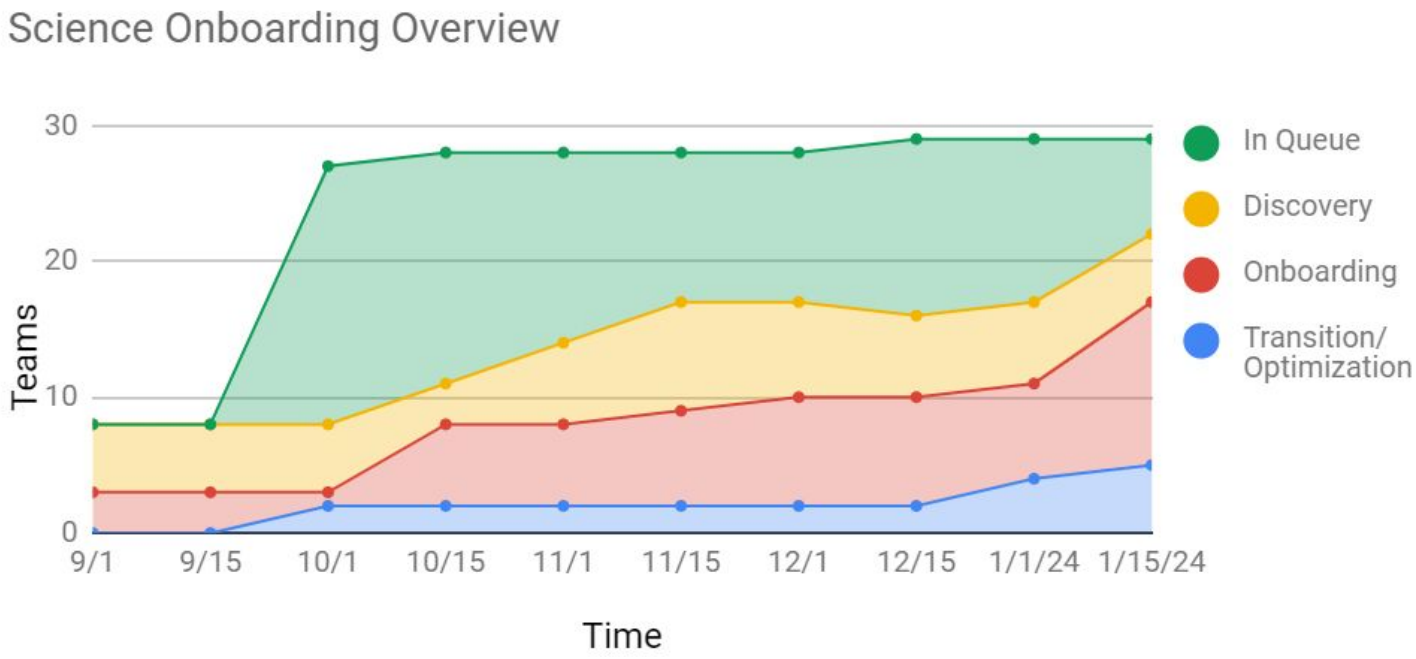
Transition/Optimization Phase



Goal: Get the team fully established in the cloud and transitioned to operations

Current Progress

- In the first six months of the pilot, we have onboarded 6 teams.
- Discovery has accelerated and we continue getting more teams in queue to begin the process.
- Uptick in velocity is clear and average onboarding time is around 3 months.
- Working on strategies to improve velocity by shifting work to the left and beginning some of the process earlier such as the CI/CD pipeline development.



Where are we going?

- Continuing our work on onboarding and transition/optimization of applications. We have over 200 applications other program offices are looking to migrate to the NCCF.

- Continue improving discovery process to identify ahead of time NCCF services that are needed and connect teams early on.

- Teach the process! Teams can see what we do and why we do it and can start working on those before getting to us.

Conclusions

- The NCCF provides a unique architecture for science teams and their applications to perform development in an environment where all data they need is centralized in one location.
- The science sandbox provides a location for teams to develop their applications and get technical support to optimize their application for cloud computing and/or for operational use.
- Separating the transitioning procedure into two distinct phases enables faster transition to development in the sandbox, and focus on optimization and operationalization is performed as a second step post-transition.
- Performing a solutioning assessment early in the process has allowed us to identify dependencies early.
- Opportunities for improving transition velocity are starting to be identified.