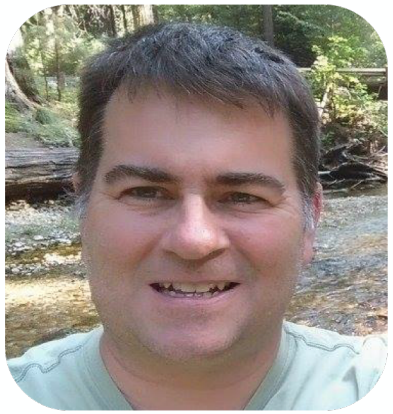
How to Use Kubernetes to Build a Data Lake for AI Workloads

Pete MacKinnon & Uday Boppana

Red Hat

Speaker Introduction



Pete MacKinnon

Red Hat AI Center of Excellence
Principal Software Engineer

pmackinn@redhat.com

Uday Boppana

Principal Product Manager
AI/ML and big data solutions

uboppana@redhat.com

AI/ML execution challenges

Long turnaround times

Centralized request processing and limited understanding of data scientist needs

Readily usable data lacking

Siloed data sets and different authentication mechanisms

Lack of collaboration across teams

Unable to implement quickly due to slow, manual and siloed operations.

Unavailability of infrastructure & software

No rapid availability of infrastructure and software tools slows data scientists and developers

AI/ML solution design points

Long turnaround times

Centralized request processing and limited understanding of data scientist needs

Readily usable data lacking

Siloed data sets and different authentication mechanisms

Lack of collaboration across teams

Unable to implement quickly due to slow, manual and siloed operations.

Unavailability of infrastructure & software

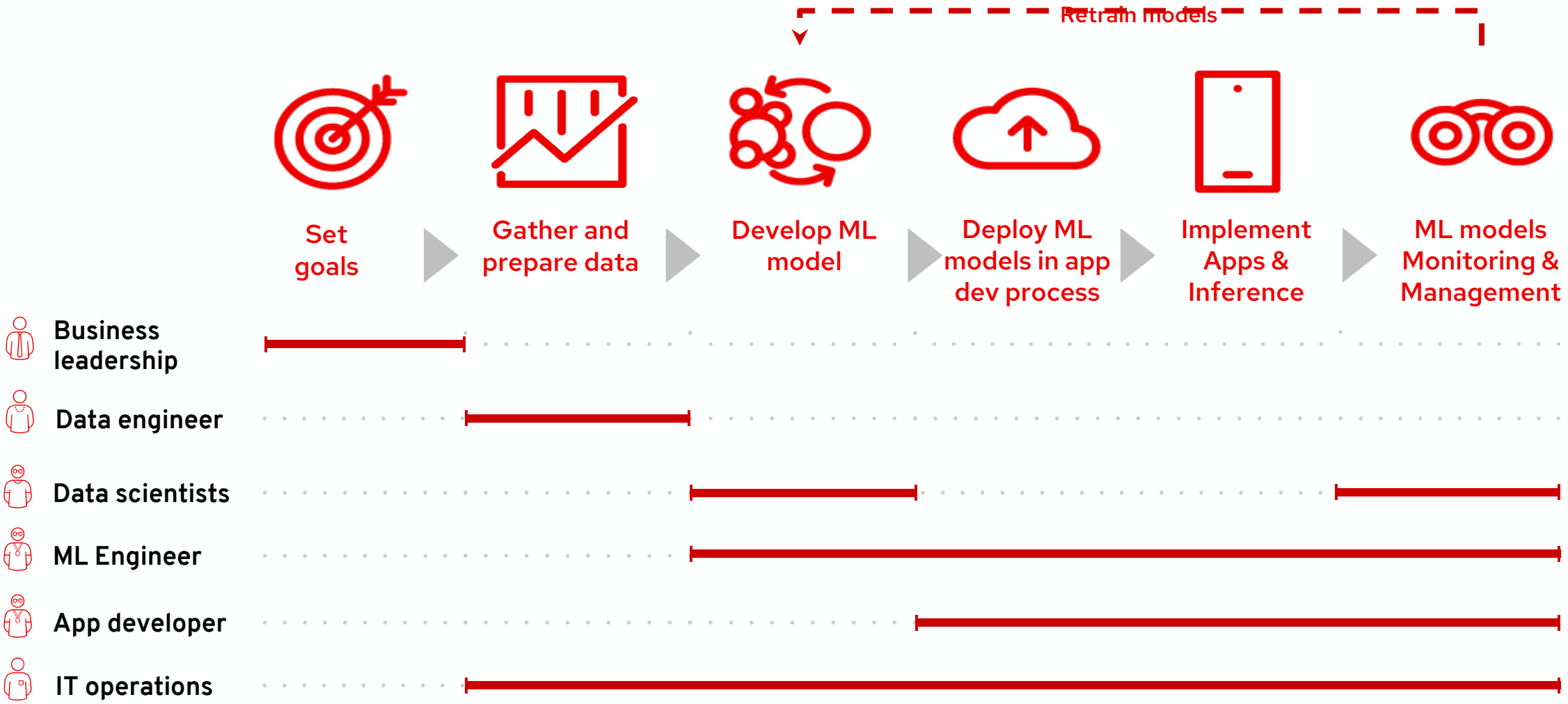
No rapid availability of infrastructure and software tools slows data scientists and developers

Self service capabilities

Centralized and shared infrastructure

Packaged tools and flexibility

AI/ML lifecycle and key personas



Kubernetes and Containers for AI/ML

Agility

Respond quickly with automated compute resource management.

Portability

Develop and deploy ML models consistently across data center, edge, and public clouds.

Flexibility

Provision AI/ML environments as and when you need them.

Scalability

Autoscaling and high availability of the AI/ML solution stack.

STORAGE FOR AI/ML DEPLOYMENTS

Gather and prepare data

Develop ML model

Deploy ML models in app dev process

Implement Apps and inference

ML models monitoring & management

ML Software, data tools & services

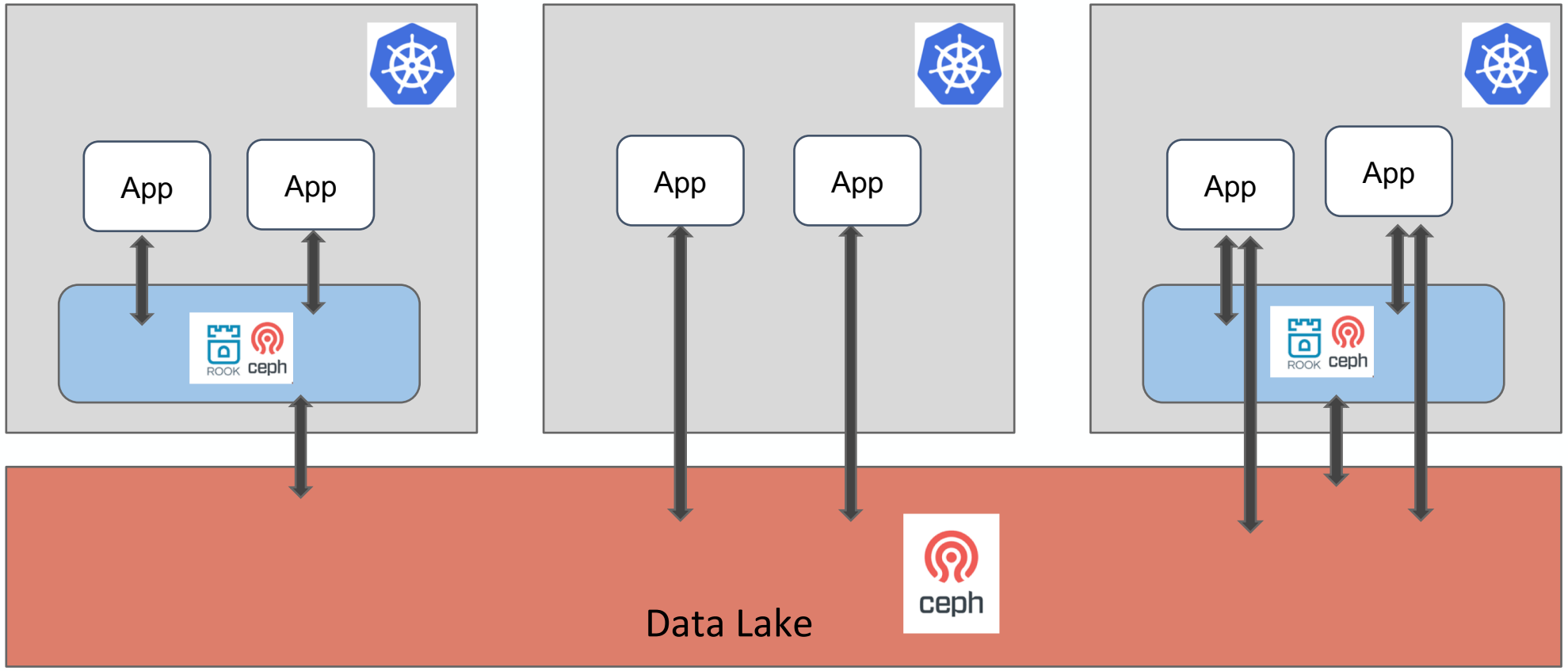
Hybrid, Multi Cloud Platform with self service capabilities

Data lake for storage & data management

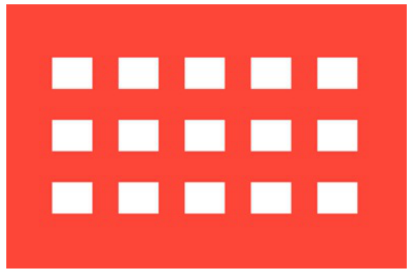
ceph

ROOK

Goal - Storage and data infrastructure



CEPH UNIFIED STORAGE + ROOK OPERATOR



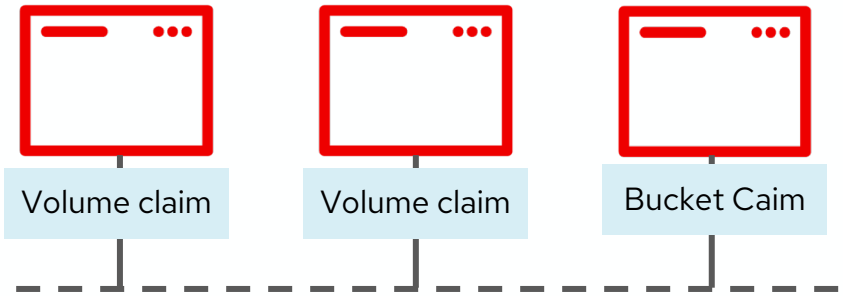
OBJECT STORAGE

BLOCK STORAGE

FILE SYSTEM

ROOK

Rook + Ceph



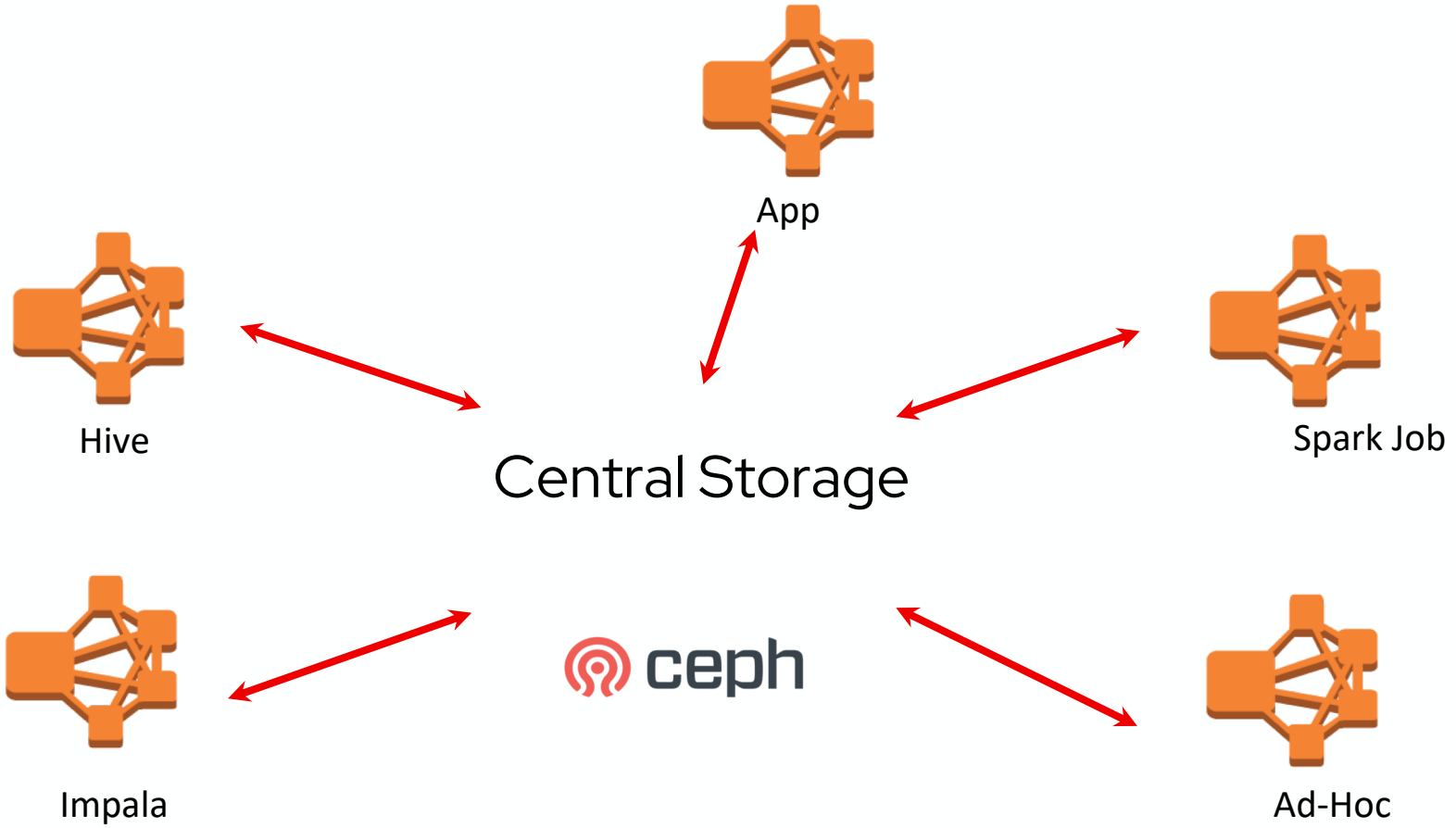
ROOK

Agent | Agent | Operator | Agent | Agent | Agent

Mon | MGR | MDS | Mon | RGW | RGW | Mon

OSDs | OSDs | OSDs | OSDs | OSDs

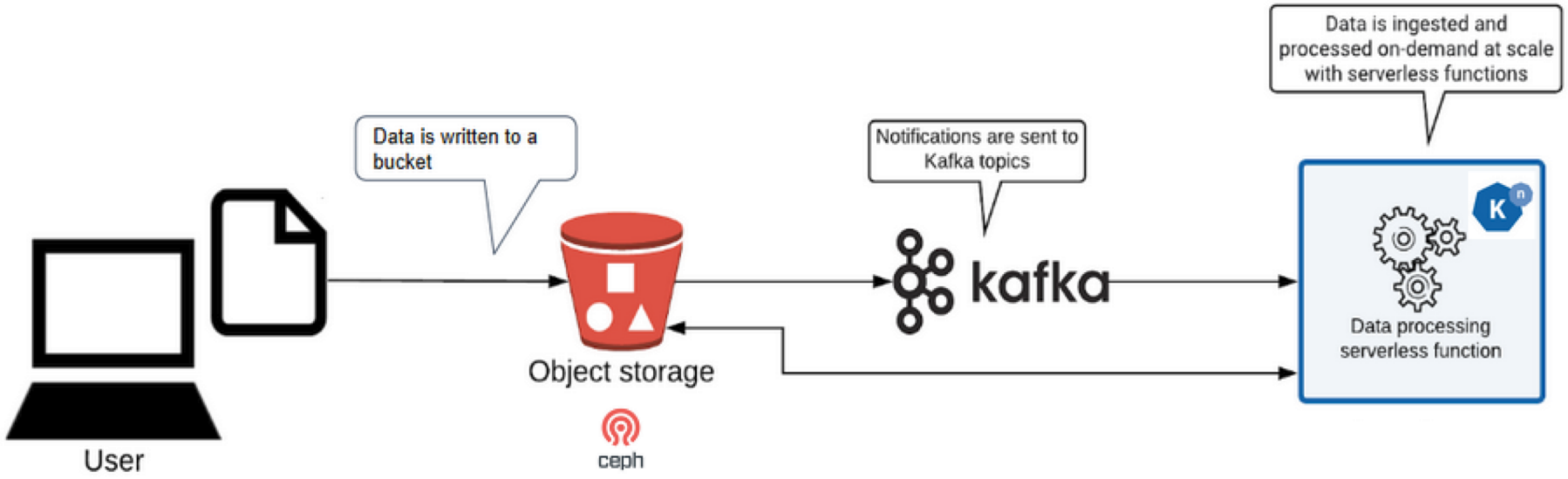
One Data Lake



Automatic ingest data processing and mobility

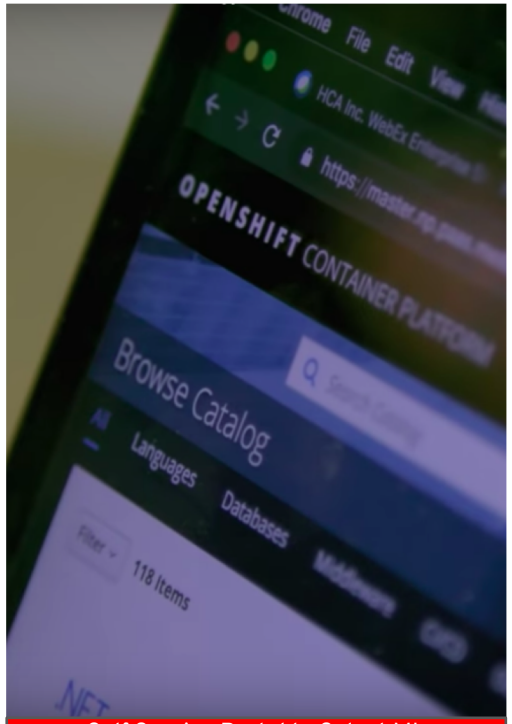
Automatic and real-time processing of ingest data

- Ceph bucket notification
- Kafka
- Knative

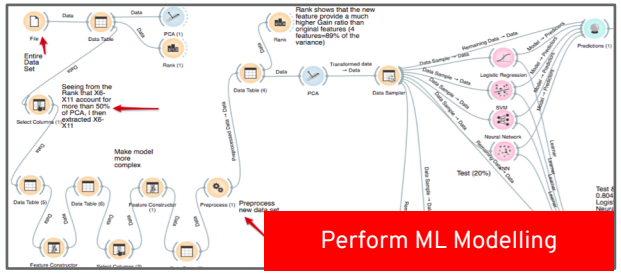


BUILDING A PLATFORM FOR DATA SCIENCE

As a Data Scientist, I want a "self-service, cloud-native" experience for my Machine Learning projects, where I can access a rich set of modelling frameworks, data, and computational resources, share and collaborate with colleagues, and deliver my work into production with speed, agility and repeatability to drive business value!



Self Service Portal to Select ML Frameworks, and Data Sources



Perform ML Modelling

Inferencing w/ Hardware Acceleration

Model Deployment in App Dev

Open source community projects in the AI/ML space

NVIDIA NGC
GPU optimized and curated

ceph

ROOK

Kubeflow

ML toolkit for Kubernetes

Tensorflow

jupyter

Spark

PYTORCH

Others

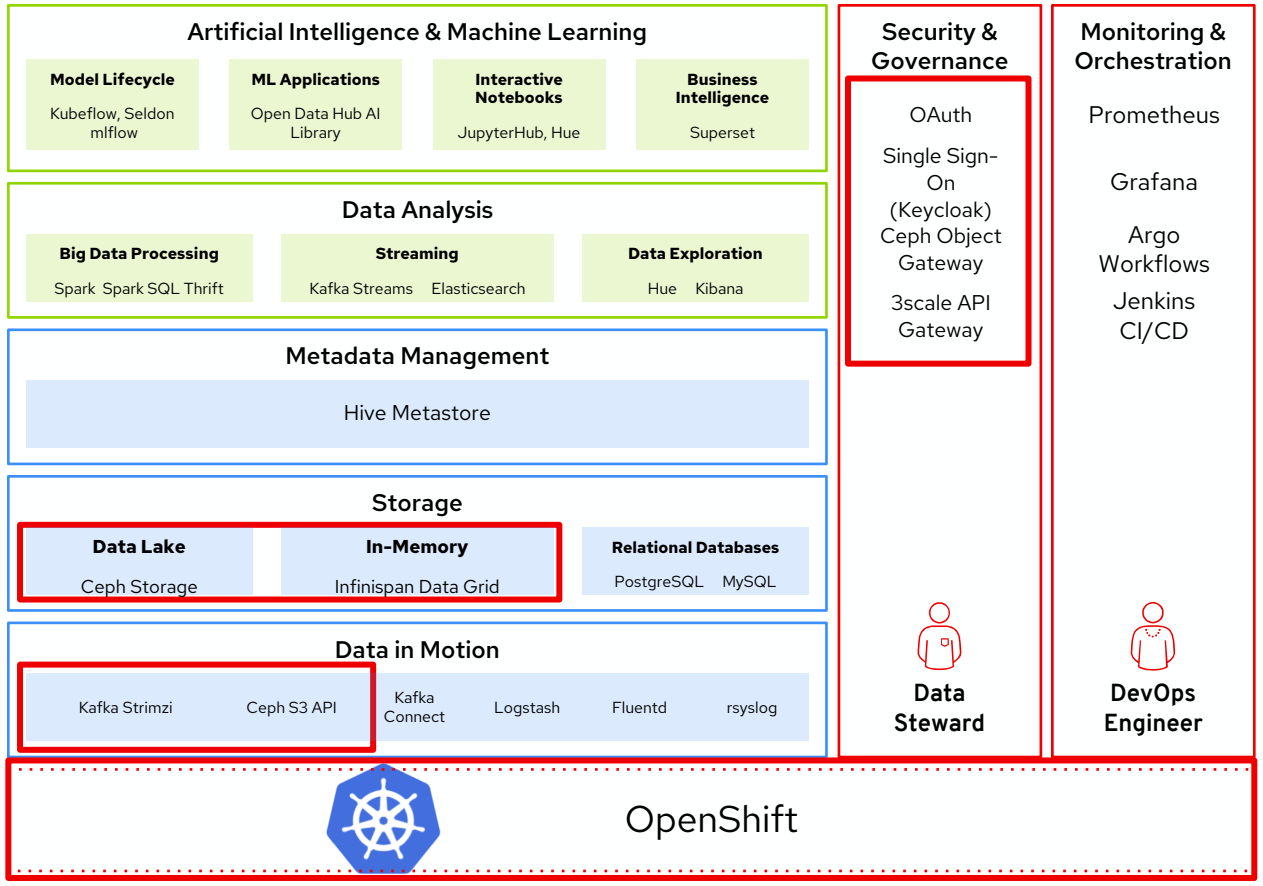
OPEN DATA HUB
AI Platform powered by Open Source

ML-as-a-service platform based on OpenShift, Ceph storage, Kafka, JupyterHub and Spark

The registry for Kubernetes Operators
OperatorHub.io

Home for k8s community to share operators for various apps/tools

Open Data Hub Solution Architecture

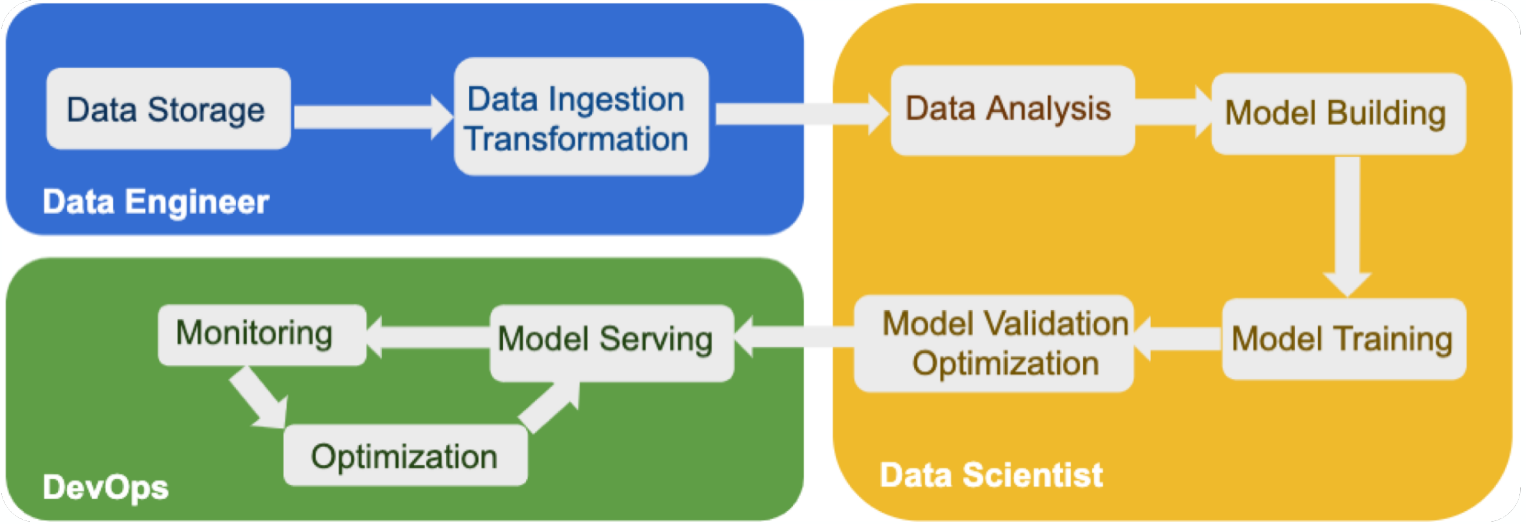


Open source community project

- Demonstrate the value of Red Hat portfolio and open source tooling to accelerate the AI/ML lifecycle.
- Automated deployment of open source AI/ML tooling.

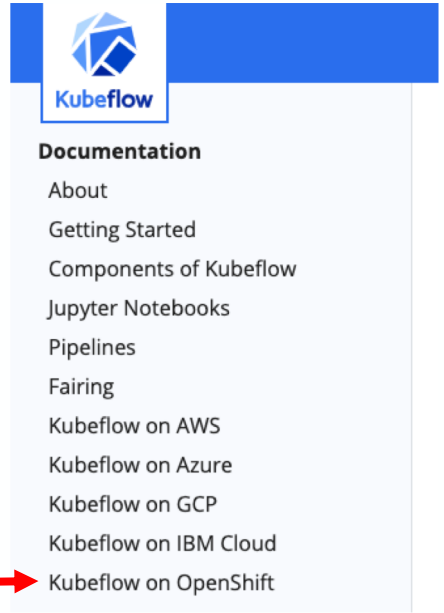
The Open Data Hub Community Project

- Collaborate on a Data & AI platform for the Hybrid Cloud - opendatahub.io/
- Open Source community meta-project to integrate open source projects into an end-to-end AI/ML platform
- Install today from the OperatorHub Community Operators

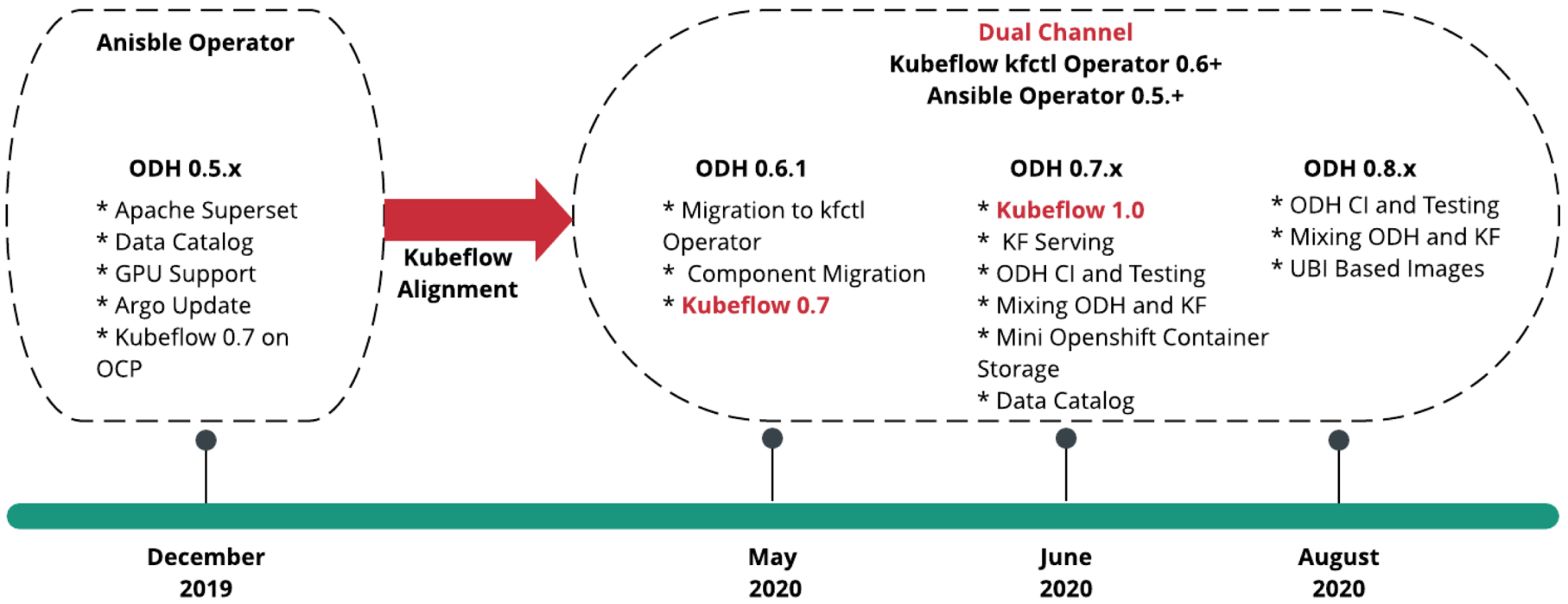


OPEN DATA HUB - Kubeflow

- Kubeflow - A new open source project dedicated to making deployments of machine learning (ML) workflows on Kubernetes simple, portable and scalable - kubeflow.org
- It is now integrated into Open Data Hub and runs on OpenShift
- Kubeflow brings multiple new AI/ML capabilities and features
  - Model training: TensorFlow, PyTorch, MPI, etc.
  - Model serving: Seldon and KFServing
  - Kubeflow Pipelines based on Argo



OPEN DATA HUB - roadmap status



OPEN DATA HUB - Contact and Engagement

ODH site: opendatahub.io

ODH-Kubeflow Github: https://github.com/opendatahub-io

ODH Community Mailing lists: announcements@lists.opendatahub.io, contributors@lists.opendatahub.io

ODH Community Meetings: https://gitlab.com/opendatahub/opendatahub-community


KubeCon
CloudNativeCon
Europe 2020
Virtual
HELM
KEEP CLOUD NATIVE
CONNECTED