

# Hadoop Migration Workshop

\$0 | 2 days (onsite/online)

## PARTICIPANTS:

- Hadoop/Spark Admins
- Data Engineers
- Infrastructure Engineers
- Data Analysts
- Data Scientists

## PRE-REQUISITES:

- Hadoop workload in PROD
- Prioritized use cases

## DISCOVERY

Understand EMR features and how to migrate Hadoop to AWS through a use case.

## KEY ACTIVITIES

**Amazon EMR Intro:** Expand the Customer knowledge of AWS EMR including best practices, architectural designs, migration patterns, and customer references.

**Demos:** Demonstrate steps involved to create an AWS EMR cluster, submit HIVE/Pig jobs on EMR, run SPARK-based ETL on EMR and integration with Sagemaker through pre-built demos.

**Use Case Analysis:** Decode current Hadoop workload by delving deeper into one customer-specific use case and its associated components.

- Understand pain points and share insights on how to address it in AWS
- Collaborate and jointly create a tailored high-level migration plan, approach and next steps

## OUTCOME:

Extend knowledge on Amazon EMR and how to migrate a Hadoop workload to AWS.

- Understand EMR features, migration options
- Hadoop to AWS migration approaches
- Future State AWS Architecture
- High-level Migration Plan
- Next Steps

“ 83% of data migration projects either fail outright or exceed their budgets and schedules – Gartner

“ Many Enterprises invest poorly on assessment and planning phase, setting themselves up for failure