CI/CD on Azure Databricks using Azure DevOps
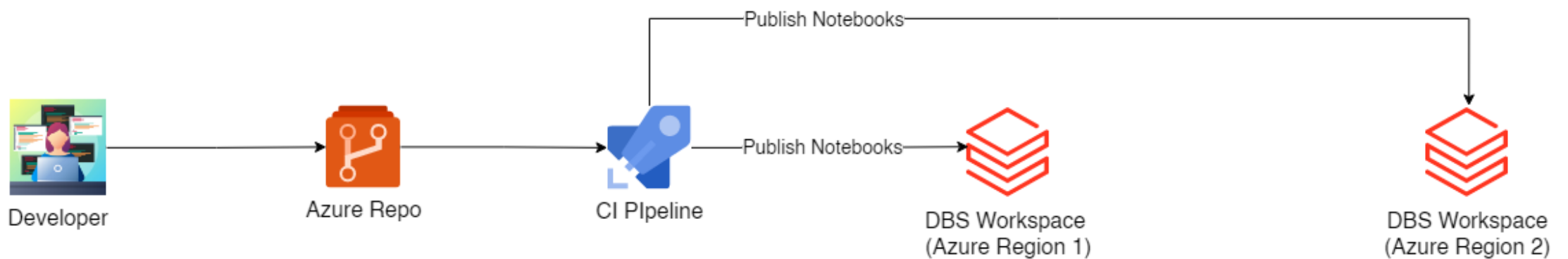
DISASTER RECOVERY
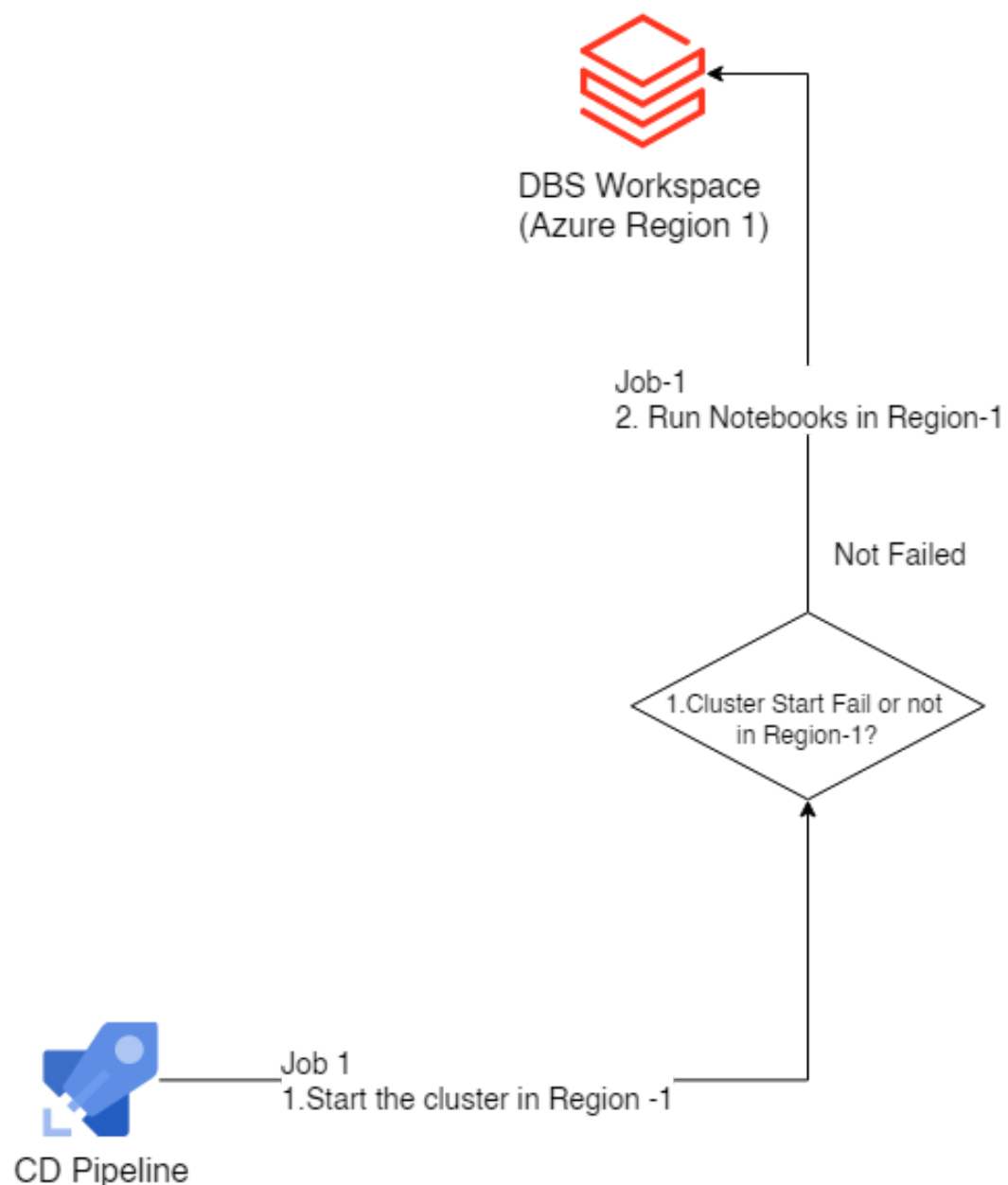
APPROACH

Architecture

CI Pipeline in Azure DevOps

- Pipeline will run to publish notebooks into both workspaces when developers push notebooks to Azure Repos.
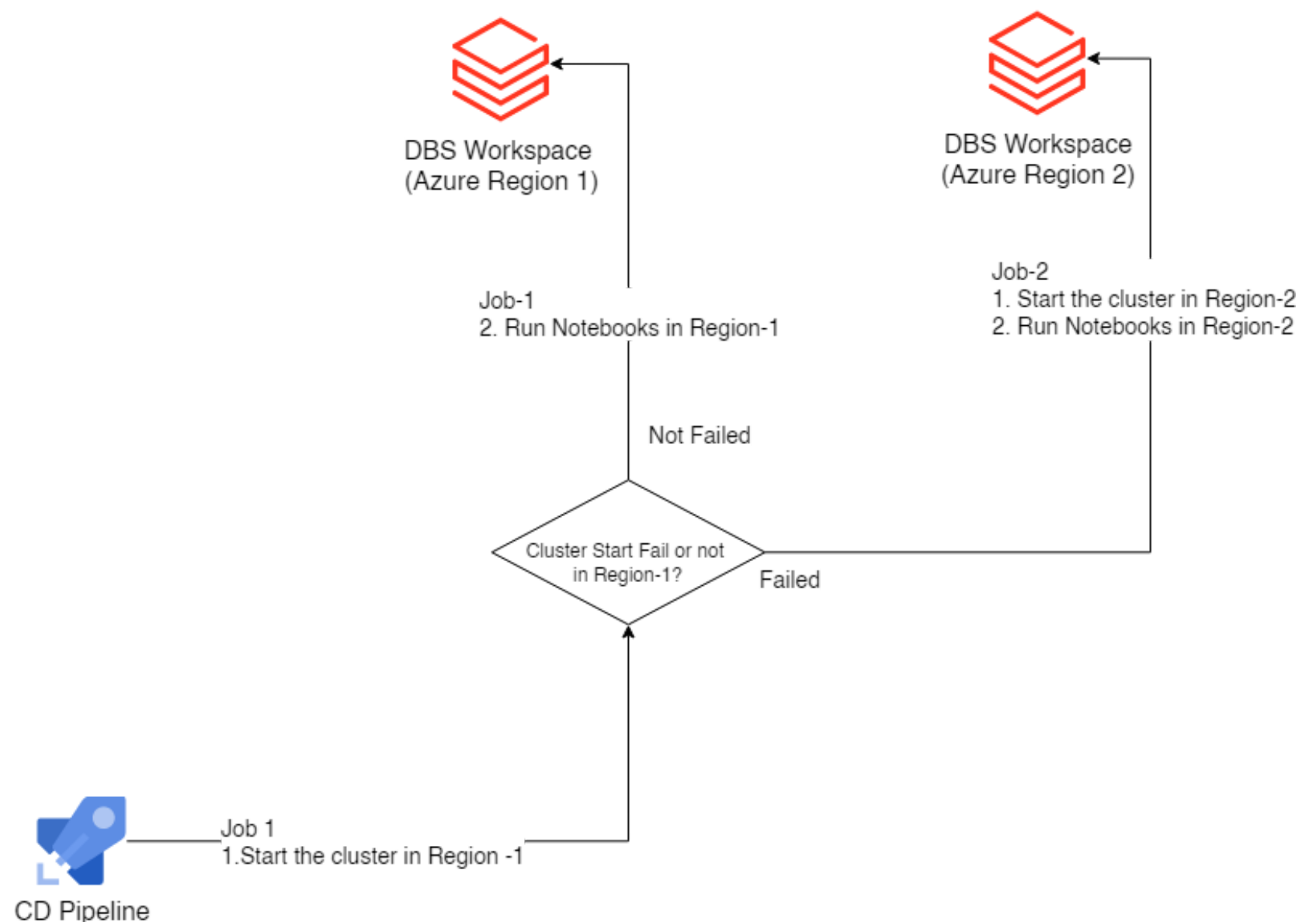
CD Pipeline(Job-1) in Azure DevOps

- Firstly, Job 1 in CD pipeline will run on a schedule to start cluster in primary region.
  - After cluster starting process is successfully finished in primary region, notebooks will be run on that cluster.

DBS Workspace
(Azure Region 1)

Job-1
2. Run Notebooks in Region-1

Not Failed

1.Cluster Start Fail or not
in Region-1?

Job 1
1.Start the cluster in Region -1

CD Pipeline

CD Pipeline(Job-2) in Azure DevOps

- If cluster starts to fail in Primary region,
  - Job-2 will be triggered to start cluster in secondary region and run notebooks on that cluster

DBS Workspace
(Azure Region 1)

DBS Workspace
(Azure Region 2)

Job-1
2. Run Notebooks in Region-1

Job-2
1. Start the cluster in Region-2
2. Run Notebooks in Region-2

Not Failed

Cluster Start Fail or not
in Region-1?   Failed

Job 1
1.Start the cluster in Region -1

CD Pipeline

THANK YOU

FOR READING