Kubernetes native workflow engines

How to make something happen in a predictable way

Whoami?

Customer Success Manager @ IBM

Experience as:

- Cloud Architect

- Cloud Engineer

- DevOps Engineer

- Linux Administrator

Studied software design

What is a workflow engine

- automates often manual steps in an automated way

- the steps to archive a goal a defines in some language or visual designer

- often inputs and outputs are used

- business critical processes can be delivered faster to increase

- automate repeating tasks

- is able to eliminate human error in processes which can be automated

What to do with a workflow engine?

- execute a CI/CD pipeline

- run machine learning jobs

- start containers only when needed (compare knative/lambda....)

- define complex workflows with multiple nodes and artifacts shared between steps

- automate data processing steps

- run many steps in parallel, ideally with autoscaling the servers running on

- trigger workflows on a specific time or event

DAG = Direct acyclic graph

- A DAG is a directed graph with no cycles (has a beginning and an end)

- a vertex is a node in the graph

- a edge connects two nodes

- a DAG allows steps for a computer to be executed in a specific order where one step can define which previous steps it depends on
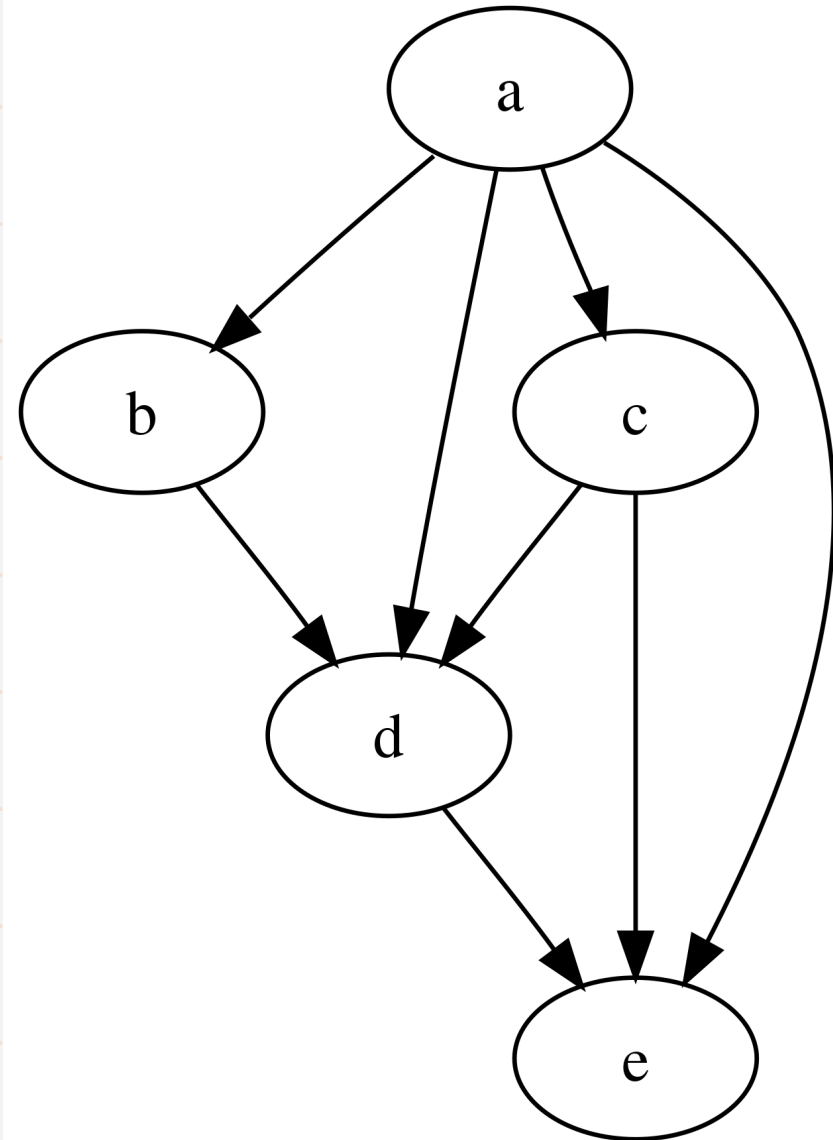
Argo Workflows

- Declarative and GitOps friendly using Kubernetes custom resources

- runs on any kubernetes distribution and on any cloud or on-prem solution

- used by big companies like Google, Datadog, NVIDIA, Github, SAP…

- can be controlled via UI, cli tool, REST API and Kubernetes CRs

- **CNCF** incubation level project

Argo workflows comparable products

- Tekton: kubernets native tool more focused on CI/CD

- Apache airflow: most used DAG tool

- kubeflow: kubernetes native tool focused on machine learning tasks

Demo time

How to trigger workflows? Argo events

- allows to trigger argo workflows and more things like AWS lambda serverless functions

- can be triggered by many sources like a slack channel, files uplodaded to s3 or git commits

- For everything else a generic http webhook can be used