class Profile(Struct):

project = Default(String, 'twttr-recos-ml-prod')

date = Required(String)

build\_target = Required(String)

job\_name = Required(String)

environment = Default(String, 'dev')

machine = Default(String, 'n2-standard-2')

frequency = Default(String, 'PT1H')

SimClustersIndexGenerationJob = Job(

name='{{profile.job\_name}}',

project='{{profile.project}}',

staging\_bucket='{{profile.project}}',

service\_account='{{profile.user\_name}}-shdw@twttr-dp-svc-accounts.iam.gserviceaccount.com',

region='us-central1',

worker\_config=WorkerConfig(

num\_workers=2,

worker\_machine\_type='{{profile.machine}}',

worker\_disk\_type=WorkerDiskType('HDD'),

),

extra\_args={

"date": '{{profile.date}}'

},

service\_identifier='twtr:svc:{{profile.user\_name}}:{{profile.job\_name}}:{{profile.environment}}:{{profile.cluster}}',

deployment\_config=BatchDeploymentConfig(

role='{{profile.user\_name}}',

environment='{{profile.environment}}',

build\_target='{{profile.build\_target}}',

gcp\_deployment\_credentials='/var/lib/tss/keys/{{profile.user\_name}}/cloud/gcp/dp/shadow.json',

statebird\_config=StatebirdConfig(

batch\_width='{{profile.frequency}}',

first\_time='{{profile.date}}',

),

workflow\_config=WorkflowConfig(

play=True,

),

timeout='PT4H'

)

)

jobs=[SimClustersIndexGenerationJob.bind(profile=Profile())]