class Profile(Struct):

project = Required(String)

date = Required(String)

environment = Default(String, 'prod')

job\_name = Default(String, 'mts-consumer-embeddings-tweets-ann-batch-job')

machine = Default(String, 'n2-highmem-4')

job = 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='prod',

build\_target='src/scala/com/twitter/simclusters\_v2/scio/bq\_generation/tweets\_ann:mts-consumer-embeddings-tweets-ann-batch-job',

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

statebird\_config=StatebirdConfig(

batch\_width='PT4H',

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

),

workflow\_config=WorkflowConfig(

play=True,

),

timeout='PT24H'

)

)

jobs=[job]