

Back Office DevOps test

The Back Office development team has implemented a task queue to process long running tasks. It's based on Celery and it uses Redis as a broker. To process the jobs there will be a Python based worker that can process a single job at a time.

Jobs can take from minutes to hours to finish.

The task queue load won't be consistent so we are interested in scaling up and down (down to zero) depending on the pending jobs in the queue.

Your task is to provide the code to orchestrate the infrastructure necessary to run these services in AWS:

- redis
- workers (number of instances will depend on a CloudWatch metric)

Instructions

- Use terraform to orchestrate the AWS
- Scale up to 4 workers if there's more than 4 pending jobs in the queue
- Stop the worker if there's no pending jobs to process
- Instructions should be provided on how to execute your code
- An AWS architecture diagram should be provided
- It's not necessary to configure services, only infrastructure
- Assume that you can find the number of pending jobs in an AWS CloudWatch metric

Points we will take into account

- Reusability (e.g. ability to start different environments from the same code)
- Availability
- Security
- Cost