scaleit

The purpose of this exercise is to write an auto-scaler that manipulates the number of replicas of a separate application based on CPU utilization metrics. As your auto-scaler changes the replica count, CPU utilization will be impacted accordingly.

Vimeo has provided such an application that has the following characteristics: * it starts on a user defined port (given by the `--port` flag) defaulting to 8123 * the JSON/REST API allows you to monitor the application's emulated CPU usage and change the number of replicas * its APIs occasionally return errors to mimic real life

The reported CPU usage will simulate a real application and rise and fall over time. At the same time, increasing the number of replicas will make the CPU usage go down and vice versa.

Your Challenge

Write an application that uses the API described below to automatically adjust the number of replicas to keep the average CPU usage reported to average 0.80 (80%).

Please write your application in one of the following languages: Python, Go, Ruby, Java, TypeScript/JavaScript, or C/C++

What We're Looking For

Your submission will be evaluated based on the following criteria: * clean, readable, testable, performant and well-documented code * demonstrated knowledge of software engineering best practices in the programming language of your choice * correctness and bug-free code

You should be thorough, but feel free to simply document where you would do things differently in a real production environment rather than an exercise.

We expect that this exercise will take roughly 60 to 90 minutes.

The API

The following describes the REST API implemented by the sample application.

Current Status

The `/app/status` API expects an HTTP `GET` request with an `Accept` header with a value of `application/json`. It, as you might expect, returns a JSON representation of the current app status that looks like this:

```
{
    "cpu": {
```

```
        "highPriority": 0.68
    },
    "replicas": 10
}
```

CPU utilization is returned as a float between 0 and 1. Replicas is an integer greater than or equal to 1.

Updating the Replica Count

The `/app/replicas` API expects an HTTP `PUT` request with a `Content-Type` header with a value of `application/json`. The body of the request should contain JSON in the following format.

```
{
    "replicas": 11
}
```

The replicas key should have an integer value greater than one. Invalid values will return HTTP 400 Bad Request.