

Python homework

Overview

This is a coding assignment for a backend developer position at Aiven.

The exercise should be relatively fast to complete. You can spend as much time as you want to. If all this is very routine stuff for you, this should not take more than a few hours. If there are many new things, a few evenings should already be enough time.

Homework evaluation is one of the criteria we use when selecting the candidates for the interview, so pay attention that your solution demonstrates your skills in developing production quality code.

If you run out of time, please return a partial solution, and describe in your reply how you would continue having more time.

Please use Python for the exercise, otherwise, you have the freedom to select suitable tools and libraries (with a few exceptions listed below), but make sure the work demonstrates well your own coding skills.

Be prepared to defend your solution in the possible interview later.

To return your homework, store the code and related documentation on GitHub for easy access. Please send following information via email:

- link to the GitHub repository
- if you ran out of time and you are returning a partial solution, description of what is missing and how you would continue

Your code will only be used for the evaluation.

Exercise

Your task is to implement a system that monitors website availability over the network, produces metrics about this and passes these events through an Aiven Kafka instance into an Aiven PostgreSQL database.

For this, you need a Kafka producer which periodically checks the target websites and sends the check results to a Kafka topic, and a Kafka consumer storing the data to an Aiven PostgreSQL database. For practical reasons, these components may run in the same machine (or container or whatever system you choose), but in production use similar components would run in different systems.

The website checker should perform the checks periodically and collect the HTTP response time, error code returned, as well as optionally checking the returned page contents for a regexp pattern that is expected to be found on the page.

For the database writer we expect to see a solution that records the check results into one or more database tables and could handle a reasonable amount of checks performed over a longer period of time.

Even though this is a small concept program, returned homework should include tests and proper packaging. If your tests require Kafka and PostgreSQL services, for simplicity your tests can assume those are already running, instead of integrating Aiven service creation and deleting.

Aiven is a Database as a Service vendor and the homework requires using our services. Please register to Aiven at <https://console.aiven.io/signup.html> at which point you'll automatically be given \$300 worth of credits to play around with. The credits should be enough for a few hours of use of our services. If you need more credits to complete your homework, please contact us.

The solution should NOT include using any of the following:

- Database ORM libraries - use a Python DB API compliant library and raw SQL queries instead
- Extensive container build recipes - rather focus your effort on the Python code, tests, documentation, etc.

Criteria for evaluation

- Code formatting and clarity. We value readable code written for other developers, not for a tutorial, or as one-off hack.
- We appreciate demonstrating your experience and knowledge, but also utilizing existing libraries. There is no need to re-invent the wheel.
- Practicality of testing. 100% test coverage may not be practical, and also having 100% coverage but no validation is not very useful.
- Automation. We like having things work automatically, instead of multi-step instructions to run misc commands to set up things. Similarly, CI is a relevant thing for automation.
- Attribution. If you take code from Google results, examples etc., add attributions. We all know new things are often written based on search results.
- “Open source ready” repository. It’s very often a good idea to pretend the homework assignment in Github is used by random people (in practice, if you want to, you can delete/hide the repository as soon as we have seen it).