Distributed Applications in the Edge-Cloud Continuum (2022W)

Homework 02 - BaaS services

Sashko Ristov

In Homework 01, you developed functions that load from and store data to a cloud storage. However, applications usually need to process that data in between. Therefore, in this homework, you need to extend the existing functions to process that data.

The aim of Homework 02 is to explore existing BaaS services, select two services of two providers, develop two or four serverless functions that use those BaaS services, and to explore the way BaaS services are invoked and communicate with data on storage.

1 Exploration task

Your task is to select two BaaS services (e.g., Object recognition and translation), both of which are supported from your preferred cloud providers.

Example 1 For instance, AWS offers AWS Rekognition and AWS Translate, while Google offers corresponding Google Vision AI and Google Translation AI.

Note: some cloud services may process data directly from the storage, without downloading data to the local file system of the function. Other services may store their output directly to cloud storage. Fig. 1 presents some examples of how a BaaS service may interact with the function and storage.

Function f1 - Call by reference in the same region. Both data and computing are placed in the same region (e.g., AWS NV). The function invokes the service with the location of data (by reference) and the service then accesses the storage in the same region through the reference.

Function f2 - Call with data by reference in another region. Data is stored on another region than the function. For instance, data is stored in AWS Oregon, while the function is deployed in AWS North Virginia (NV). The function invokes the service by reference. You need to explore where the service is executed (is it in AWS Oregon as it is shown in the Fig. 1, or in AWS NV, which then accesses to data in AWS Oregon.

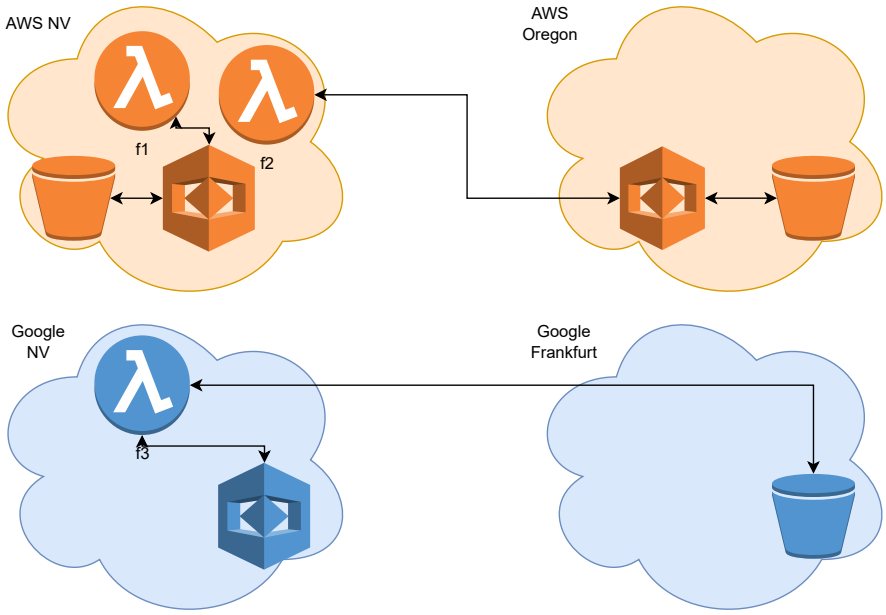

Figure 1: Two examples of different approaches for BaaS services

Function f3 - Call with data by value in the same region. In this example, data needs to be on the local file system of the function. In such case, the function needs to download data (from a storage of another or the same region) and then pass data (by value) to the service in the same region.

Note: The given three scenarios are not the only possible, but are given to guide you in your exploration. Note that you may have scenario that the input data for function f1 are stored in Google Frankfurt. This case is optional for you as it goes beyond the scope of this course, but rather for the course in the following summer semester.

2 Development task

You need to develop serverless functions that use the selected services from the exploration task in order to process data that is stored on cloud storage. For example, for the given two BaaS services from Example 1 (four in total from both providers), you need to develop one or two AWS Lambda function that use AWS Rekognition and AWS Translate and one or two Google Cloud function that use Google Vision AI and Google Translation AI.

Note: It is enough one function per provider if it uses both services.

3 Investigation (research) task

In this task you need to evaluate your selected services by running your function. For this purpose, run your functions as for scenario f1, that is, both the function and the input data is stored in the same region. Afterwards, run scenario f2 or f3, based on your assumption and the way how your services are invoked. For instance, if you invoke your service by name and location of the data, probably you face with scenario 2.

In order to evaluate the service time, measure it with a *white box* approach, that is, from inside the functions, and compare it with the upload / download time from the function `DuwnUp` from Homework 01.

Note: If you need any help or some parts are unclear, please write on Discord or email.