



FCTUC

UNIVERSIDADE DE COIMBRA
FACULDADE DE CIÊNCIAS E TECNOLOGIA

Departamento de Engenharia Informática

Project #3
Engenharia de Serviços
2016/17 – 2nd Semester
MEI

Deadline: 2017-05-26

Nota: A fraude denota uma grave falta de ética e constitui um comportamento não admissível num estudante do ensino superior e futuro profissional. Qualquer tentativa de fraude pode levar à reprovação na disciplina tanto do facilitador como do prevaricador.

Workflows on the Cloud

Objectives

- Gain familiarity with Amazon AWS Resources: EC2, SimpleDB, S3, Auto Scaling, SQS, SWF, Steps, AIM, IAM
 - Gain familiarity with the utilization of these resources in Python, using boto3 (or another version)
-

Final Delivery

- No report is necessary.
 - You must submit the **Python source code** of the project.
 - You should install and configure all the software on the Amazon AWS cloud.
 - Keep your installation on Amazon until the day of the defense.
-

References

You should resort to the Amazon AWS pages for support:

Auto Scaling Service

- <http://aws.amazon.com/autoscaling>

Amazon SimpleDB

- <https://aws.amazon.com/simplydb>

Amazon SQS

- <http://aws.amazon.com/sqs>

Amazon SNS

- <http://aws.amazon.com/sns>

Amazon Machine Image

- <http://docs.aws.amazon.com/AWSEC2/latest/UserGuide/AMIs.html>

Amazon Identity and Access Management

- <http://aws.amazon.com/iam>

Amazon Simple Storage Service (S3)

- <https://aws.amazon.com/s3>

Amazon Elastic Compute Cloud (EC2)

- <https://aws.amazon.com/ec2>

Amazon Simple Workflow Service

- <https://aws.amazon.com/swf/>

Amazon Step Functions

- <https://aws.amazon.com/step-functions/>

Amazon Rekognition

- <https://aws.amazon.com/rekognition/>

Serverless workflows on AWS: my journey from SWF to step functions

- <https://forrestbrazeal.com/2016/12/29/serverless-workflows-on-aws-my-journey-from-swf-to-step-functions/>
-

Security

By default, Amazon prevents access to most of the resources necessary for this work. To access these resources students should configure, at least, one Identity and Access Management (IAM) user, and provide several permissions to this user, including to the Simple Queue Service, Simple Workflow Service, Simple Storage Service, and SimpleDB. The access to SimpleDB is slightly more complicated to provide, because this service does not have any managed policy, this meaning that students must prepare one policy themselves.

Overview

Solution based on Amazon Simple Queue Service

This part of the assignment is worth 63% of the grade

We will use Amazon Rekognition to simulate a new ticketing service, for example, for the canteen. The idea is to do authentication based on facial recognition. Once authenticated, canteen users will be able to charge their accounts and pay for their meals. The system must allow users to pay without authentication. Refer to Figure 1, for the payment flowchart.

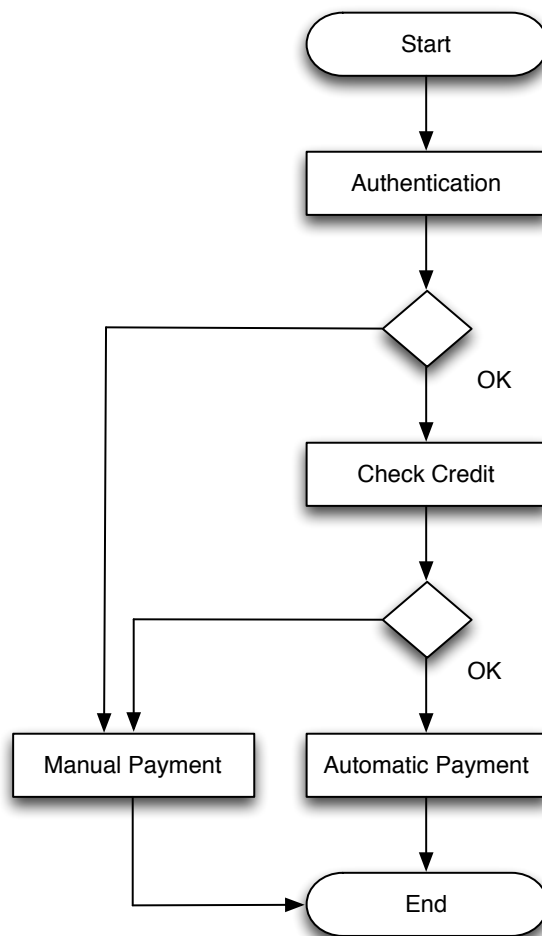


Figure 1 - Basic Payment Workflow

Besides this workflow, there should be another one to charge the account. However, for simplicity, we will use some other basic scheme to randomly charge the accounts.

The first implementation should use SQS, S3, an auto-scaling group and, possibly, SimpleDB or DynamoDB, to keep data of users and of flows.

Students will also need to develop a couple of stand-alone programs to interact with the workflow, e.g., to simulate a new user entering the queue or charging the account.

The overall life cycle of a job should follow these steps:

- The user simulates entering the payment system, by submitting a request in a stand-alone application. The photo enters the system at this point. The application blocks until a response comes.
- The information of the request goes to the input queue of the authentication, while the photo of the job goes to S3.
- When a node is available to process the request, it reads a message from the input queue, thus getting information of the request it should process.
- The node then goes to S3 to get the photo, before using Rekognition. A timeout might be useful here. Students might want to consider an additional step, to let the

user confirm that the authentication was correct. This is not mandatory. This process might take a while, thus requiring a group of nodes.

- The authenticator node sends back the result of the authentication to the appropriate queue.
- If the user was authenticated, a similar process is repeated to check if the user has enough credit.
- Based on the previous results, the application either pays manually or automatically, although this does not correspond to any concrete action (a printscreen suffices).

Students may need EC2 nodes to invoke Rekognition and check credit. For this, they might use a special (golden) Amazon Machine Image or startup scripts.

Note: to avoid incurring in extra costs, consider deleting all traces from SimpleDB each time the manager finishes, because there is no access to this database from the AWS console.

[Solution based on the Amazon Simple Workflow Service \(SWF\)](#)

This part of the assignment is worth 21% of the grade

In the second part of the assignment students should do a similar implementation using workflows instead of queues. I.e., you should exclude both queues from the assignment, and should use workflows to control the execution of the entire process, either using SWF or Step functions. Again, the starting process should take place in a stand-alone application.

[Extra Components](#)

This part of the assignment is worth 16% of the grade

[Monitoring:](#)

Solutions that can provide detailed monitoring, for example of the auto scaling group or S3 bucket(s) will be valued.

Students are encouraged to experiment with X-Ray in this part of the assignment, although this is not mandatory.

[Failures:](#)

Solutions that can tolerate worker node failures (e.g., because of the elimination of machines in the auto scaling group) will be valued. These solutions are not necessary in the SQS part of the assignment.

DESIGN YOUR EXPERIMENTS CAREFULLY, BECAUSE AUTO SCALING GROUPS AND OTHER RESOURCES INVOLVE COSTS.

ALSO, NOTE THAT EACH MACHINE IS CHARGED AT LEAST 1 HOUR. E.G., SCALING OUT TO 4 MACHINES DURING 5 MINUTES WILL COST YOU, AT LEAST, 4 HOURS.