Hands On 1:
Service as unit of construction
(ROUND 2)

María Fernanda Flores Luna
ID. 162064

Cloud Computing and Big Data
LIS 4102-1
13th march 2022

Spring 2022

Index

Introduction

The objective of the laboratory of the course is to play with different cloud solutions and understand its purposes and motivations. The first hands-on presented the opportunity to interact with two different services on the cloud: Google Collab and Microsoft Azure.

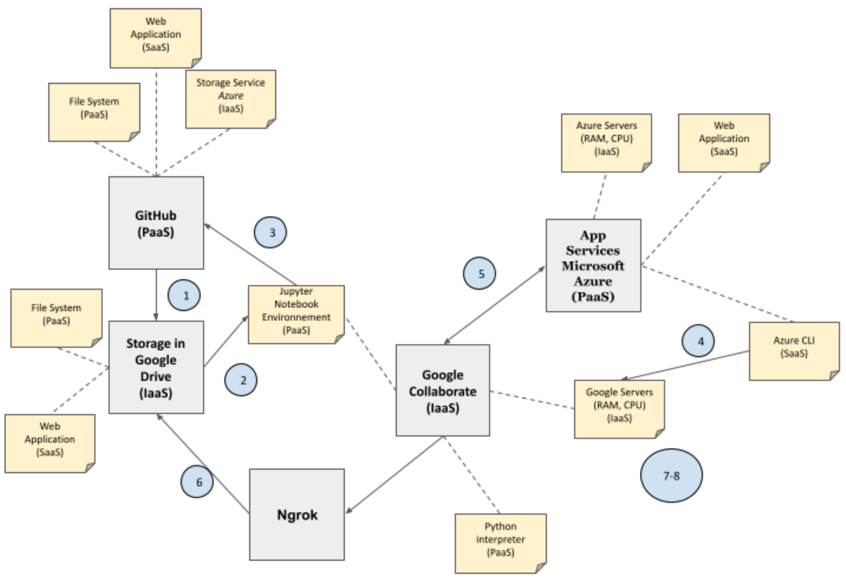
- ★ Google Collab was used as a virtual machine to avoid the installation and the limitations of our hardware.
- ★ Microsoft Azure was used as a platform to launch our web application.

In the following report we are making an attempt to explain the interactions that happened in the experiment between the services and how they made it possible.

Functional Architecture

In this section, we are going to show how the different cloud services interact for building the application using a functional architecture schema. Creating a functional architecture schema means: to draw boxes and arrows that represent the interaction between the elements during the execution of the application. For the schemas, the experiment was segmented into 4 parts: setting up the environment, running the web application on Google Collab, the deployment of the application on Microsoft Azure and running the web application on Microsoft Azure.

The first stage is setting up the environment. This stage includes everything to prepare the services before running the application. The functional architecture can be seen in figure 1. The diagram was designed using the instructions given in class.


Web Application (SaaS)
File System (PaaS)
Storage Service Azure (IaaS)
GitHub (PaaS)
3
1
File System (PaaS)
Jupyter Notebook Environnement (PaaS)
Storage in Google Drive (IaaS)
2
Web Application (SaaS)
6
Ngrok
Azure Servers (RAM, CPU) (IaaS)
Web Application (SaaS)
App Services Microsoft Azure (PaaS)
5
Google Collaborate (IaaS)
4
Azure CLI (SaaS)
Google Servers (RAM, CPU) (IaaS)
7-8
Python interpreter (PaaS)

Setting up the environment.

1. Clone hands-on exercise GitHub repository to Google Drive
2. Open the hands-on exercise on the Notebook environment
3. Clone the code provided by Azure -with the web application "hello world"- in GitHub from the Notebook

Azure-Google Drive

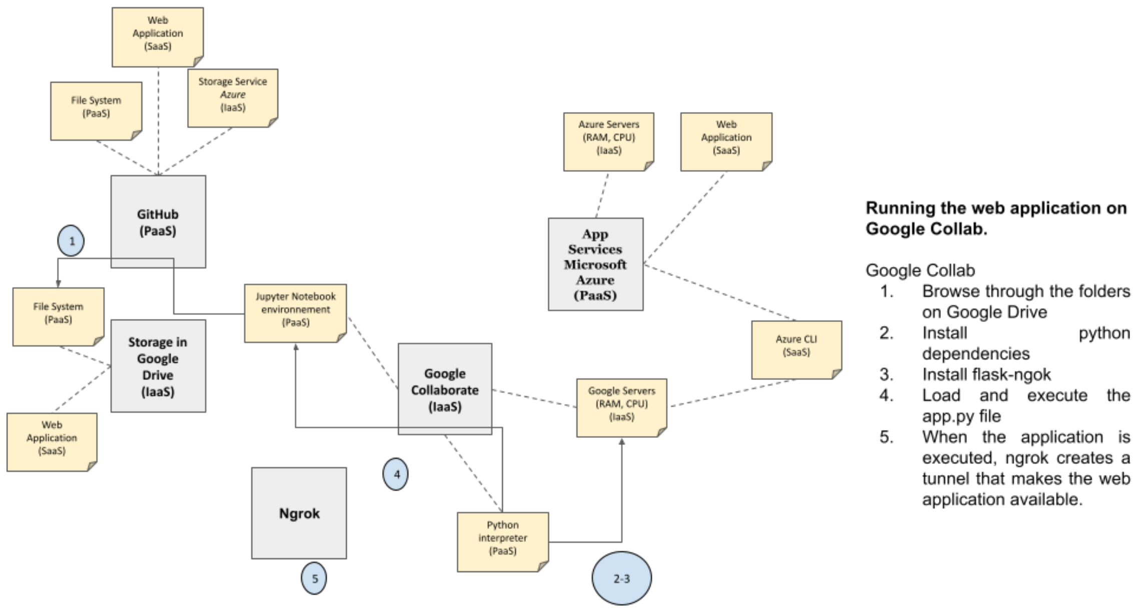
4. Install Azure CLI in the Server provided by Google
5. Connect to the Azure portal using the account

Ngrok Service

6. Retrieve the code ngrok-stable-linux-amd64.zip
7. Unzip the ngrok-stable-linux-amd64.zip
8. Execute the yml script with the authentication token

Figure 1. Functional architecture of setting up the environment.

The second stage is running everything on Collab. This stage includes from the installation of python libraries to the application running. The functional architecture can be seen in figure 2. It is important to highlight that the arrows that flow from the python interpreter mean that the command is written in Python and executed later.

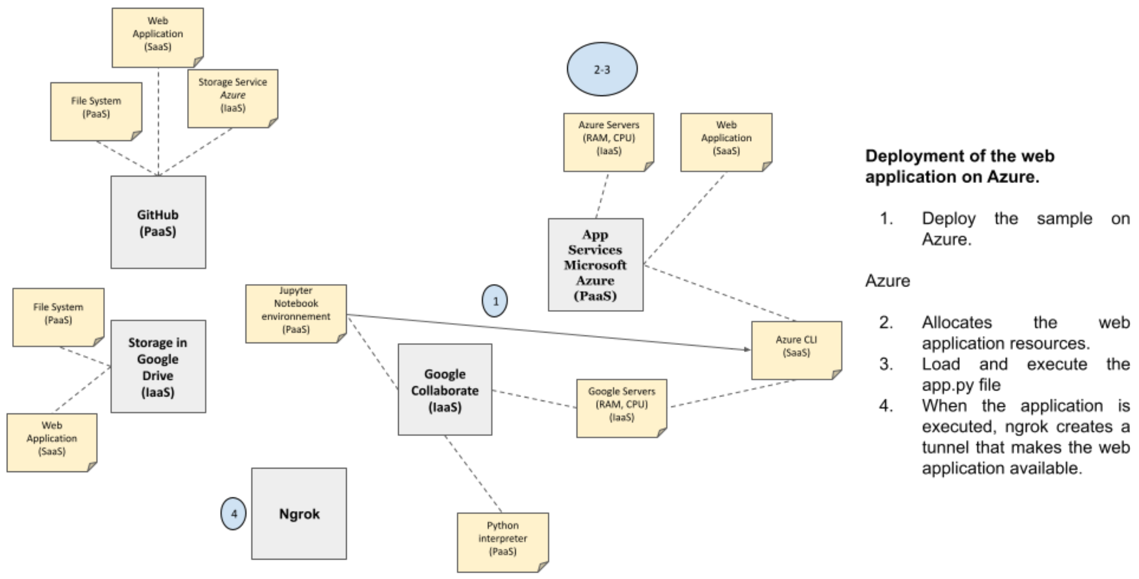

Web Application (SaaS)
File System (PaaS)
Storage Service Azure (IaaS)
GitHub (PaaS)
1
File System (PaaS)
Storage in Google Drive (IaaS)
Web Application (SaaS)
Jupyter Notebook environnement (PaaS)
Google Collaborate (IaaS)
Ngrok
4
5
Python interpreter (PaaS)
2-3
Google Servers (RAM, CPU) (IaaS)
Azure Servers (RAM, CPU) (IaaS)
Web Application (SaaS)
App Services Microsoft Azure (PaaS)
Azure CLI (SaaS)

Running the web application on Google Collab.

Google Collab
1. Browse through the folders on Google Drive
2. Install python dependencies
3. Install flask-ngok
4. Load and execute the app.py file
5. When the application is executed, ngrok creates a tunnel that makes the web application available.

Figure 2. Functional architecture of running the web application on Google Collab.

The third stage is deploying the application on Azure. This stage includes the operations executed on Collab and on Azure. The functional architecture can be seen in figure 3.


Web Application (SaaS)
File System (PaaS)
Storage Service Azure (IaaS)
GitHub (PaaS)
File System (PaaS)
Storage in Google Drive (IaaS)
Web Application (SaaS)
Jupyter Notebook environnement (PaaS)
Google Collaborate (IaaS)
4
Ngrok
Python interpreter (PaaS)
1
Google Servers (RAM, CPU) (IaaS)
2-3
Azure Servers (RAM, CPU) (IaaS)
Web Application (SaaS)
App Services Microsoft Azure (PaaS)
Azure CLI (SaaS)

Deployment of the web application on Azure.

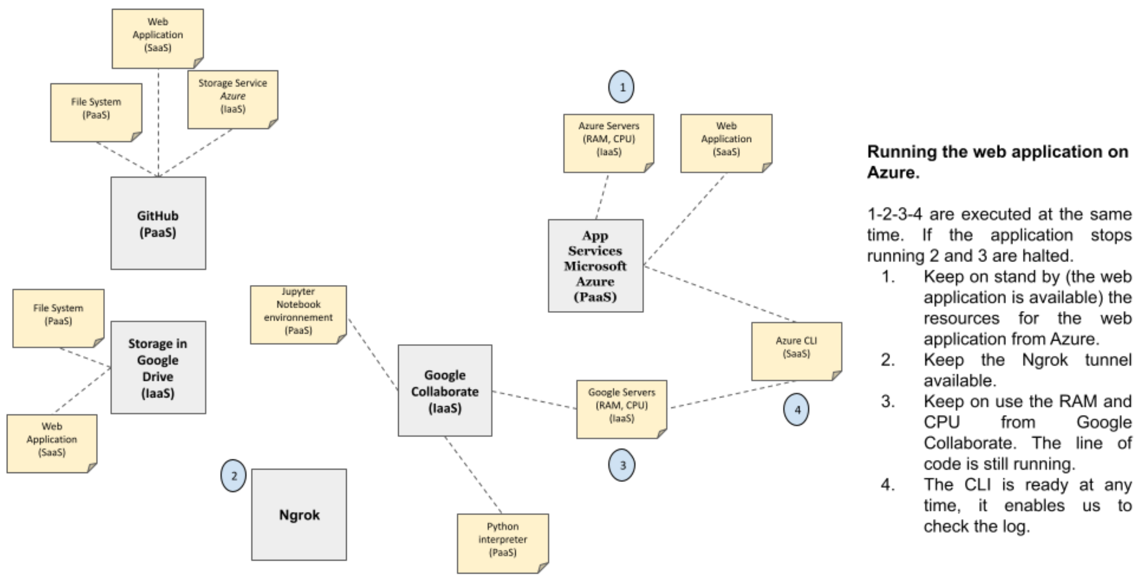
1. Deploy the sample on Azure.

Azure

2. Allocates the web application resources.
3. Load and execute the app.py file
4. When the application is executed, ngrok creates a tunnel that makes the web application available.

Figure 3. Functional architecture of deployment of the web application on Azure.

The final stage is running the application on Azure. This stage is the usage of resources from every active service. The functional architecture can be seen in figure 4.


Web Application (SaaS)
File System (PaaS)
Storage Service Azure (IaaS)
GitHub (PaaS)
File System (PaaS)
Storage in Google Drive (IaaS)
Web Application (SaaS)
Jupyter Notebook environnement (PaaS)
2
Ngrok
Google Collaborate (IaaS)
Python interpreter (PaaS)
1
Azure Servers (RAM, CPU) (IaaS)
Web Application (SaaS)
App Services Microsoft Azure (PaaS)
Azure CLI (SaaS)
Google Servers (RAM, CPU) (IaaS)
3
4

Running the web application on Azure.

1-2-3-4 are executed at the same time. If the application stops running 2 and 3 are halted.

1. Keep on stand by (the web application is available) the resources for the web application from Azure.
2. Keep the Ngrok tunnel available.
3. Keep on use the RAM and CPU from Google Collaborate. The line of code is still running.
4. The CLI is ready at any time, it enables us to check the log.

Figure 4. Functional architecture of running the web application on Azure.

Once the diagrams have been presented, let us talk about the type of services they represent. The first step to do that is analyzing the types of services, table 1 presents a list of common services and their type.

Service	Type
Compilers	**Platform as a Service (PaaS).**
DBMS	
Operating System	
Web application	**Software as a Service (SaaS).**
BlackBoard	
App that gives access to the backend of a system	
Storage services	**Infrastructure as a Service (IaaS)**

Table 1. Common services and their type.

After analyzing table 1 it is possible to categorize the services used during the experiment. Table 2 presents the service and their type, for better visualization the color purple represents Google Clouds services and color blue Microsoft Azure services.

	Service	Type
Google Cloud	Google Drive	Platform as a Service (PaaS).
	Jupyter Notebook	
	Python Interpreter	
Microsoft Azure	File System	
	CLI	
	Web App	
Google Cloud	Storage	Infrastructure as a Service (IaaS).
	CPU+RAM	
Microsoft Azure	Storage	
	CPU+RAM	

Table 2. Identifying the experiment components and their type of service.

Process diagrams

In this section, we are going to show the processes that were executed during the experiment using process diagrams. The process diagrams that are going to be presented next do not follow a specific regulation, they are created using a basic nomenclature to explain what was done.

There are 5 process diagrams, each one represents a step of the experiment. Since there is no general diagram to explain how each diagram is related to the others, the processes are going to be presented in a sequential manner.

Step 1. Ngrok authentication. This process is presented on figure 5, it includes the actions executed by the systems of ngrok and the ones taken manually by the user and on Collab.

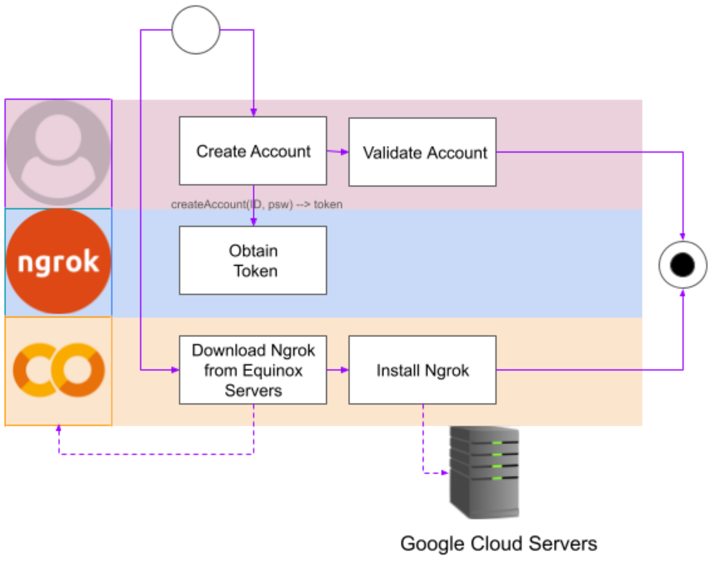
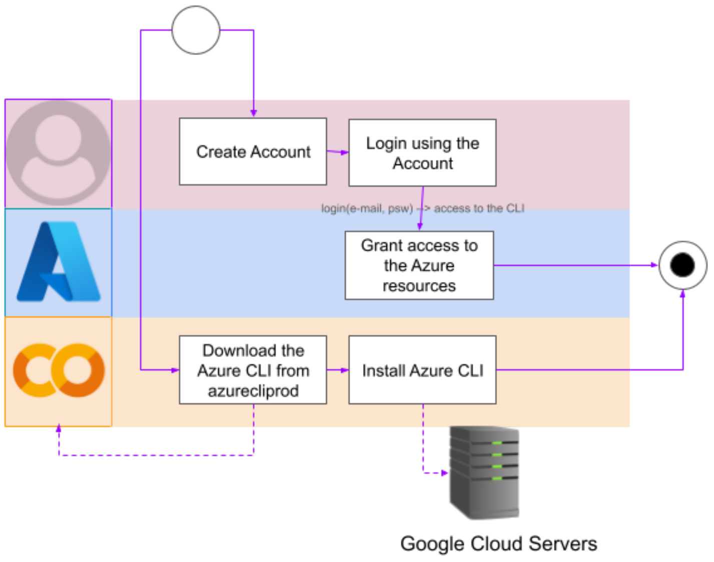


Figure 5. Process diagram of Ngrok authentication.

Step 2. Installing the Azure CLI and configuring it. This process is presented on figure 6, it includes the actions executed by the systems of Microsoft Azure and the ones taken manually by the user and on Collab.



Create Account
Login using the Account
login(e-mail, psw) --> access to the CLI
Grant access to the Azure resources
Download the Azure CLI from azurecliprod
Install Azure CLI
Google Cloud Servers

Figure 6. Process diagram of Installing Azure CLI and its configuration.

Step 3. Cloning from GitHub and configuring Python. This process is presented on figure 7, it includes the actions executed by the systems of GitHub and the ones taken by the user on Collab.

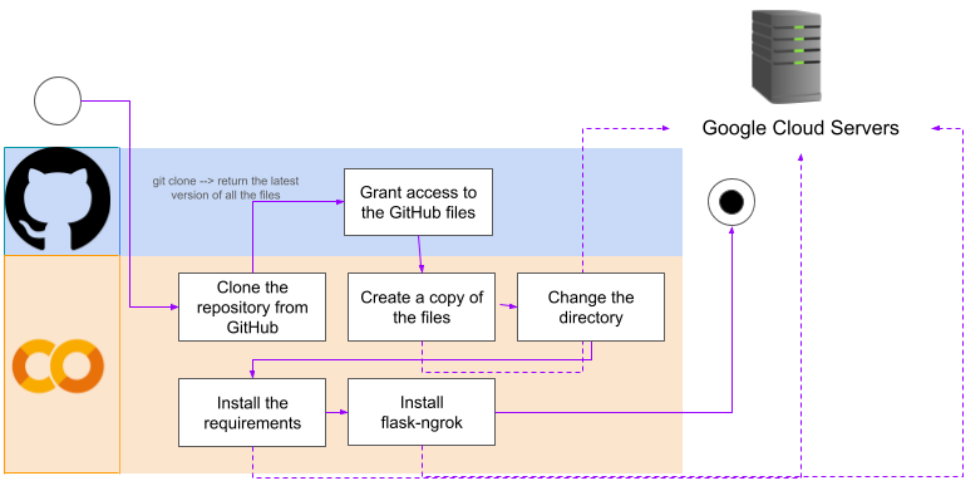

Google Cloud Servers
git clone --> return the latest version of all the files
Grant access to the GitHub files
Clone the repository from GitHub
Create a copy of the files
Change the directory
Install the requirements
Install flask-ngrok

Figure 7. Process diagram of Cloning from GitHub and configuring Python.

Step 4. Running the sample on Collab. This process is presented on figure 8, it includes the actions executed by the systems of ngork, flask and the ones taken manually by the user and on Collab.
Note: there are 2 ways for the process to finish, either the user chooses to stop the execution or leave the program running.

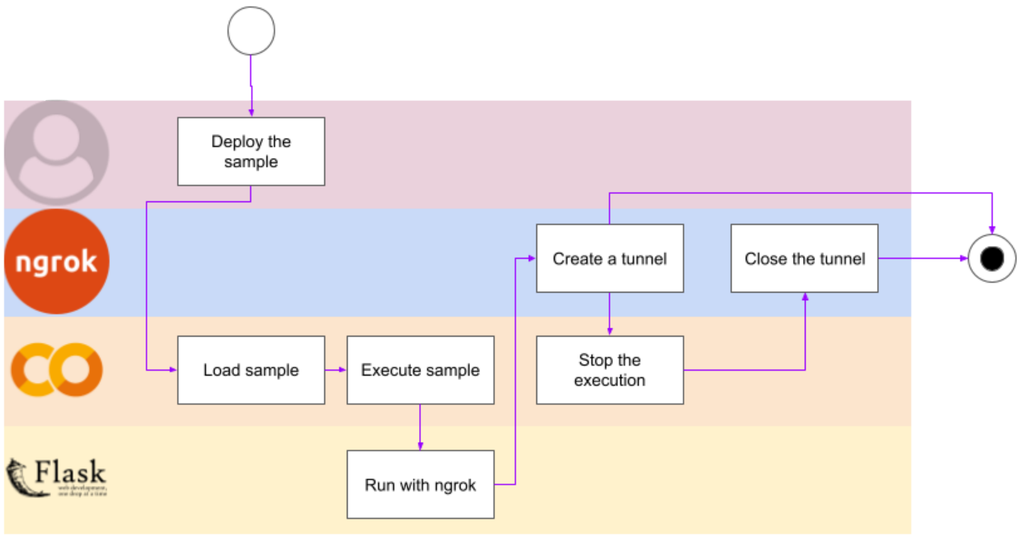


Figure 8. Process diagram of Running the sample on Collab.

Step 5. Running the sample on Azure. This process is presented on figure 9, it includes the actions executed by the systems of ngork, flask, Microsoft Azure and the ones taken manually by the user and on Collab.

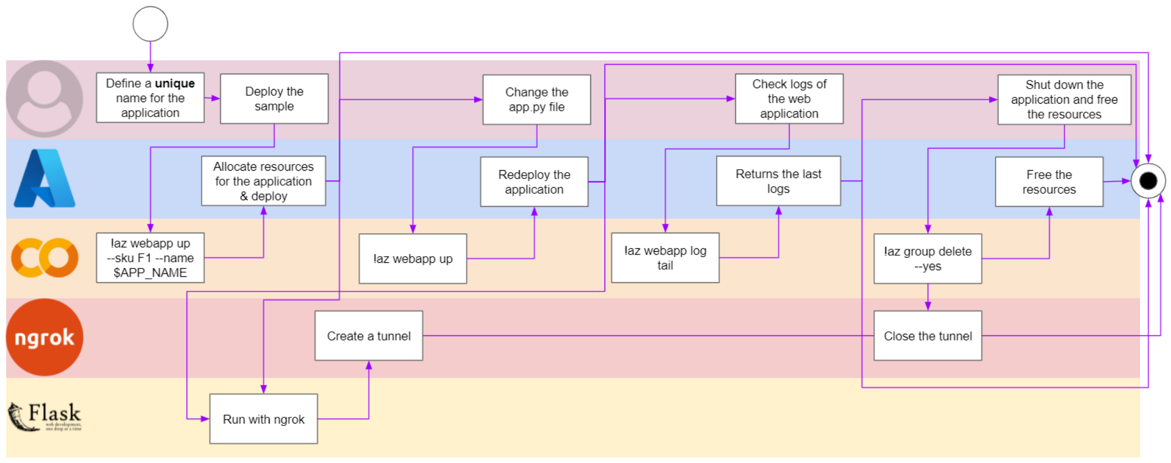


Figure 9. Process diagram of Running the sample on Azure.

Comparison

In this section, we are going to compare the experiment performed in the laboratory sessions and a tutorial by Azure. For this part we are going to use two comparison tables: a table with general criteria (comparison of the methodology of both experiments) and a table with technical implications. First let us present the comparison table with general criteria (table 3).

Comparison criteria	**Lab Tutorial**	**Azure Tutorial**
Code retrieval process	The web application source code is retrieved from GitHub.	
Installation requirements	The requirements installation is performed on the cloud (Google Collab).	The requirements installation is performed locally (VS Code in your Windows machine).

Virtual environment	Since we are working in the cloud, there is no need for a virtual environment.	We need a virtual environment, we create one using the following commands:
		`py -m venv .venv`
		`.venv\scripts\activate`

Monitoring http traffic	To monitor the http traffic we use ngrok. We installed a middleware-like application: flask.ngrok.	There is no tool to monitor http traffic.
Web application	Both of the web applications are created using Flask.	

libraries		
Azure	The Azure environment is set up from the CLI. The only thing done using the Azure portal is the identity verification.	The Azure environment can be set up either from the Azure portal, VS Code or the CLI.
Stream logs	The stream logs were accessed briefly from the Azure CLI.	The final part of the tutorial includes an introduction on stream logs.
Cleaning resources	The clean up of the resources is performed from the Azure CLI.	The clean up of the resources can be done either from the Azure portal, VS Code or the CLI.
Web application interface	The application includes a text message.	The application includes Azure logo, a text message, a text input bar and a button.

Table 3. Comparison table between the performed experiment and the one suggested by Azure (general criteria).

Once we have presented the general criteria, we are going to present the technical implications table (table 4).

Comparison criteria	**Lab Tutorial**	**Azure Tutorial**
Type of execution	External (on the cloud) execution.	Local execution.
Resources	The resources that are used do not belong to the user, they are “borrowed” from a cloud service.	The resources that are used belong to the user, they are limited.
Resource “elasticity”	The resources used can be extended at any moment.	The resources can not be extended.
Execution time	The execution time can be reduced, increasing the resources or optimizing the processes.	The execution time can only be reduced by optimizing the processes. Over time, the resources can even increase the execution time.

Cost	The execution cost may vary according to the resource usage and current price of the service.	The execution cost, as the resources, is constant. This means that it does not change.
Performance through the time	The execution time will not change through time, the resources that you pay will ensure an optimal performance.	The runtime will change over time, the resources (CPU, memory, etc.) will age and no longer guarantee optimal performance.
Data privacy	The program runs on external servers, which means that the information is distributed between different cloud providers. The privacy of the data depends on the privacy conditions of the services.	The program runs on local servers, which means the information is not spread across different cloud providers.

Table 4. Comparison table between the performed experiment and the one suggested by Azure (technical implications).

Conclusions

The free trial or limited accounts are useful for cloud solutions to present their services to possible clients, you can try if it is the best choice for you before starting to pay for the service!
Cloud solutions erase the limitations of regular computers, for example using IaaS you can have a computer with a limited processor and memory running complex applications.

References

API. (2022, March 11). In *Wikipedia*. https://en.wikipedia.org/wiki/API

Business process management. (2022, March 11). In *Wikipedia*. https://en.wikipedia.org/wiki/Business_process_management

Cloud computing. (2022, March 11). In *Wikipedia*. https://en.wikipedia.org/wiki/Cloud_computing

Client (computing). (2022, March 11). In *Wikipedia*. https://en.wikipedia.org/wiki/Client_(computing)

Command-line interface. (2022, March 11). In *Wikipedia*. https://en.wikipedia.org/wiki/Command-line_interface

Computer architecture. (2022, March 11). In *Wikipedia*. https://en.wikipedia.org/wiki/Computer_architecture

Flask (web framework). (2022, March 11). In *Wikipedia*. https://en.wikipedia.org/wiki/Flask_(web_framework)

Git. (2022, March 11). In *Wikipedia*. https://en.wikipedia.org/wiki/Git

GitHub. (2022, March 11). In *Wikipedia*. https://en.wikipedia.org/wiki/GitHub

Infrastructure as a service. (2022, March 11). In *Wikipedia*. https://en.wikipedia.org/wiki/Infrastructure_as_a_service

Google Drive. (2022, March 11). In *Wikipedia*. https://en.wikipedia.org/wiki/Google_Drive

Hypertext Transfer Protocol. (2022, March 11). In *Wikipedia*. https://en.wikipedia.org/wiki/Hypertext_Transfer_Protocol

Magic (programming). (2022, March 11). In *Wikipedia*. https://en.wikipedia.org/wiki/Magic_(programming)#:~:text=In%20the%20context%20of%20computer,to%20present%20a%20simple%20interface.

Microsoft Azure. (2022, March 11). In *Wikipedia*. https://es.wikipedia.org/wiki/Microsoft_Azure

Platform as a service. (2022, March 11). In *Wikipedia*. https://en.wikipedia.org/wiki/Platform_as_a_service

Project Jupyter. (2022, March 11). In *Wikipedia*. https://en.wikipedia.org/wiki/Project_Jupyter#Industry_adoption

Python (programming language). (2022, March 11) In *Wikipedia*. https://en.wikipedia.org/wiki/Python_(programming_language)

Server (computing). (2022, March 11). In *Wikipedia*. https://en.wikipedia.org/wiki/Server_(computing)

Software as a service. (2022, March 11). In *Wikipedia*. https://en.wikipedia.org/wiki/Software_as_a_service

Notebook interface. (2022, March 11). In *Wikipedia*. https://en.wikipedia.org/wiki/Notebook_interface

Virtual machine. (2022, March 11). In *Wikipedia*. https://en.wikipedia.org/wiki/Virtual_machine

Glossary

API: is a connection between computers or between computer programs. It is a type of software interface, offering a service to other pieces of software ("API", 2022).

Architecture: is a set of rules and methods that describe the functionality, organization, and implementation of computer systems. The architecture of a system refers to its structure in terms of separately specified components of that system and their interrelationships ("Computer architecture", 2022).

Azure: is a cloud oriented service by Microsoft. It was built to let users have a 'hands-on' experience with data centers- building, testing, deploying and managing applications and services ("Microsoft Azure", 2022).

BPM: is the discipline in which people use various methods to discover, model, analyze, measure, improve, optimize, and automate business processes ("Business process management", 2022).

Client: in client-server architecture a client is an entity -hardware or software- that accesses a service made available by a server ("Client (computing)", 2022).

Command Line Interface (CLI): it is an application that allows the users to ask specific tasks from the computer using short commands ("Command-line interface", 2022).

Client: in client-server architecture a client is an entity -hardware or software- that accesses a service made available by a server ("Client (computing)", 2022).

Cloud: is making available computing services -storage, servers, databases, networking, etc- through the Internet ("Cloud computing", 2022).

Flask: is a web framework written in Python, since it does not require any particular tools or extra libraries it is possible to call them microframework. Flask allows the programmer to develop web applications fast and as simple as possible ("Flask (web framework)", 2022).

Functional architecture schema: boxes and arrows that represent the interaction between the elements during the execution of the application.

Git: is software developed in 2005, for tracking changes in any set of files, usually used for coordinating work among programmers collaboratively developing source code during software development ("Git", 2022).

GitHub: is a provider of Internet hosting for software development and version control using Git, founded in 2008. It offers the distributed version control and source code management (SCM) functionality of Git, plus its own features. It provides access control and several collaboration features such as bug tracking, feature requests, task management, continuous integration and wikis for every project ("GitHub", 2022).

Google Colab: is a free Jupyter notebook environment that runs in the cloud and stores its notebooks on Google Drive (“Project Jupyter”, 2022).

Google Drive: is a file storage and synchronization service developed by Google. Launched on April 24, 2012, Google Drive allows users to store files in the cloud (on Google's servers), synchronize files across devices, and share files. (“Google Drive”, 2022).

Http: is an application layer protocol in the Internet protocol suite model for distributed, collaborative, hypermedia information systems (“Hypertext Transfer Protocol”, 2022).

IaaS: are online services that provide high-level APIs used to dereference various low-level details of underlying network infrastructure like physical computing resources, data partitioning, scaling, security, backup etc. (“Infrastructure as a service)”, 2022).

Magic: is an informal term for abstraction. It mainly refers to the code that handles complex tasks offering a high level of transparency to the user, making it seem as if it were easy (“Magic (programming)”, 2022).

Ngrok: is an authentication support that allows the application to perform requests on the web.

PaaS: is a category of cloud computing services that allows customers to provision, instantiate, run, and manage a modular bundle comprising a computing platform and one or more applications, without the complexity of building and maintaining the infrastructure typically associated with developing and launching the application(s); and to allow developers to create, develop, and package such software bundles (“Platform as a service”, 2022).

Project Jupyter: is a project and community that started in 2015, whose goal is to "develop open-source software, open-standards, and services for interactive computing across dozens of programming languages". The project has 3 core programming languages Julia, Python and R (“Project Jupyter)”, 2022).

Python: is a high-level, general-purpose programming language designed by Guido van Rossum in 1991. Its language constructs and object-oriented approach aim to help programmers write clear, logical code for small- and large-scale projects (“Python (programming language)”, 2022).

SaaS: is a software licensing and delivery model in which software is licensed on a subscription basis and is centrally hosted (“Software as a service”, 2022).

Server: in client-server architecture a server is an entity - hardware or software - that answers the requests of a client. In other words, a server manages a centralized service (“Server (computing)”, 2022).

Service: the function/operation provided by a server.

Notebook: it is an interactive digital environment created to enable the creation, development, documentation and execution of Julia/Python/R programs (“Notebook interface”, 2022).

Virtual machine: is the virtualization/emulation of a computer system. Their implementations may involve specialized hardware, software, or a combination (“Virtual machine”, 2022).