

BIG DATA LAB ASSIGNMENT ONE

Devansh Sanghvi
Department Of Biotechnology
Indian Institute of Technology, Madras
Chennai, India
be19b002@smail.iitm.ac.in

Abstract—This document aims to answer the three questions provided in Assignment One of the course: Big Data Lab

This assignment will serve as a part of the requirements to be fulfilled for receiving credits in the course **CS 4830: Big Data Lab**.

IMPORTANT NOTE: All outputs in this assignment will have the ID: **sanghvidevansh23** instead of Roll No. This is because I have signed into Google Cloud Console using my personal email ID.

I. QUESTION ONE

Question: Use Cloud Shell to deploy a Google Cloud Function that gets triggered when you upload a file on GCS. The function should print name of the uploaded file and the number of lines in the uploaded file.

SOLUTION

Steps taken:

- 1) Make a new directory, 'myfunc_final' and move to that directory.
- 2) Make two files: one called 'requirements.txt' and one 'main.py,' with the main.py file containing the function for the required task. The function is called 'a1_q1_func_final'
- 3) Deploy the function using the command shown.
- 4) Check if the function was deployed properly
- 5) Upload a new file in the bucket
- 6) Check the log of the function to get the required details about the file

SCREENSHOTS

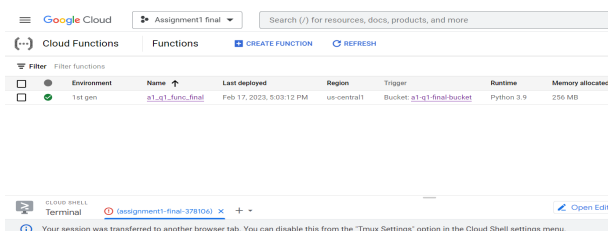


Fig. 1. Function Deployment

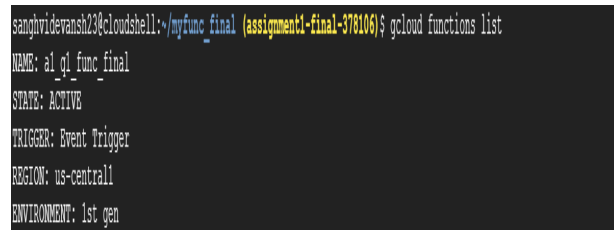


Fig. 2. Functions List

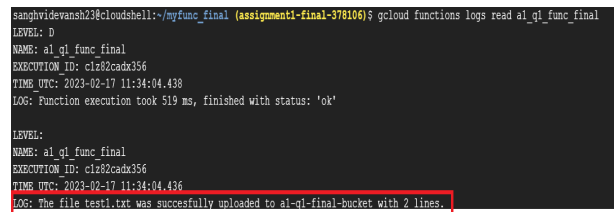


Fig. 3. Result with name of file and number of lines

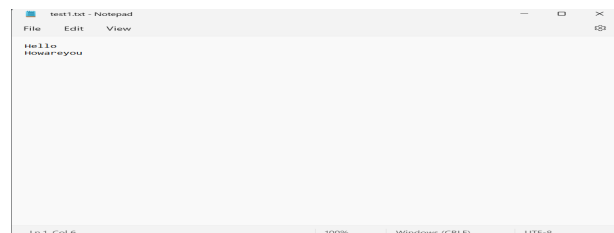


Fig. 4. Contents of the text file uploaded

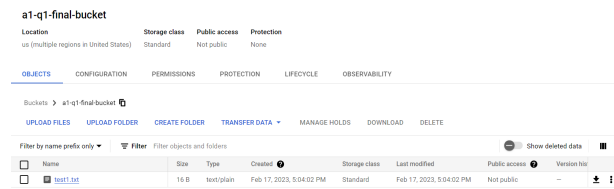


Fig. 5. Text file uploaded on the console

II. QUESTION TWO

Question:

- Spin up a VM and use the VM to calculate the sum of first N Fibonacci numbers. Use any value of N between

10 to 100.

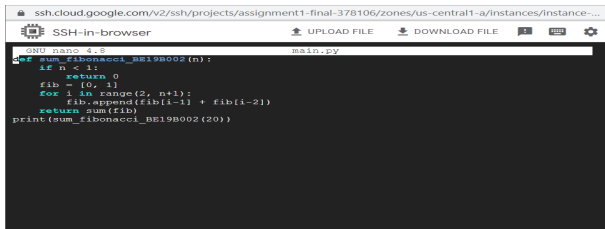
- Use Cloud Shell to deploy a Google Cloud Function that takes the value of N as argument and returns the sum of first N Fibonacci numbers

SOLUTION: PART A

Steps taken:

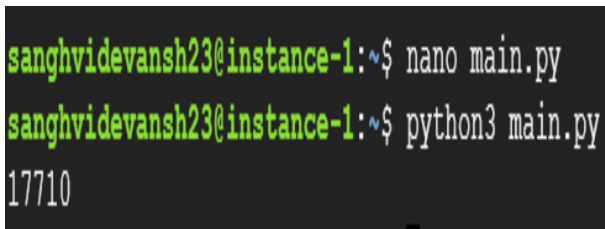
- 1) Create a new VM instance using the steps in Exercise One.
- 2) Open a python file called 'main.py' and add the python code to the file
- 3) Run the 'main.py' file to print the result
- 4) For n=20, it should return 17710

SCREENSHOTS



```
ssh.cloud.google.com/v2/ssh/projects/assignment1-final-378106/zones/us-central1-a/instances/instance-1
SSH-in-browser
main.py
def sum_fibonacci_series(n):
    if n <= 1:
        return n
    fib = [0, 1]
    for i in range(2, n+1):
        fib.append(fib[i-1] + fib[i-2])
    return sum(fib)
print(sum_fibonacci_series(20))
```

Fig. 6. Steps taken for 2A



```
sanghvidevansh23@instance-1:~$ nano main.py
sanghvidevansh23@instance-1:~$ python3 main.py
17710
```

Fig. 7. Result for 2A

SOLUTION: PART B

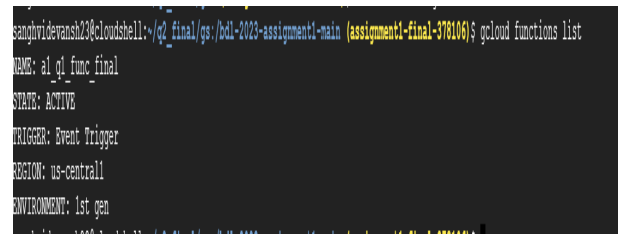
Observe what each parameter is doing and write a few lines on what you understand.

- 1) gcloud functions deploy: Used to deploy a new function which has not been deployed.
- 2) main.return _fibonacci(flask.request): Returns the result from the fibonacci function. The input variable is flask.request, i.e., the input given to flask.
- 3) request _args = request.args: gives the arguments taken from request and saves it from request _args

The function expects an input. How will you pass it?

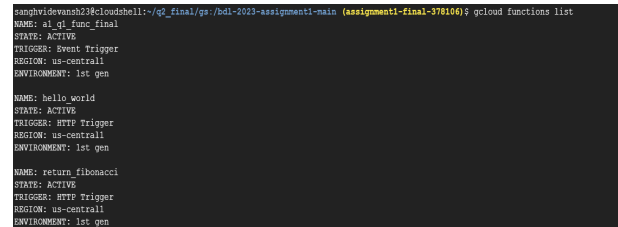
- 1) We can pass it by changing the html and adding the following command after it: "n=4". The command and the result obtained can be seen in the image. Note: The Fibonacci series is as follows: 1,1,2,3,5,8,13, and so on.
- 2) Hence the sum of the first four Fibonacci numbers gives a 7 as output.

SCREENSHOTS



```
sanghvidevansh23@cloudshell:~/q2_final/g2/2023-assignment1-main (assignment1-final-378106)$ gcloud functions list
NAME: al_qi_func_final
STATE: ACTIVE
TRIGGER: Event Trigger
REGION: us-central1
ENVIRONMENT: 1st gen
```

Fig. 8. List of gcloud functions before deploying the two functions



```
sanghvidevansh23@cloudshell:~/q2_final/g2/2023-assignment1-main (assignment1-final-378106)$ gcloud functions list
NAME: al_qi_func_final
STATE: ACTIVE
TRIGGER: Event Trigger
REGION: us-central1
ENVIRONMENT: 1st gen
NAME: return_fibonacci
STATE: ACTIVE
TRIGGER: HTTP Trigger
REGION: us-central1
ENVIRONMENT: 1st gen
```

Fig. 9. List of gcloud functions after deploying the two functions

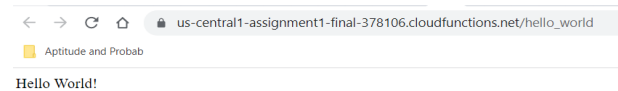


Fig. 10. Output after deploying the Hello World function

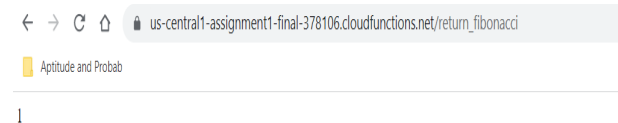


Fig. 11. Result shown when return Fibonacci is run without an input

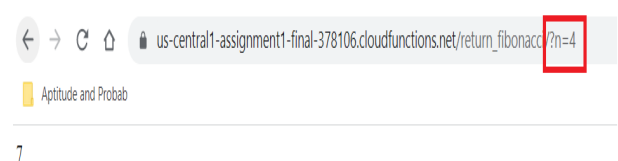


Fig. 12. Result shown when return Fibonacci is run with an input. Series: 1,1,2,3,5 and so on...

III. QUESTION THREE

Question: Write a brief description of the following:

- Serverless Computing.
- As a student of Big Data Lab, write a comparative

```

samprivedevansh23@cloudshell:~/gl_final/gs:/bdl-2023-assignment1-main (assignment1-final-378106) $ gcloud functions delete hello_world --quiet
Waiting for operation to finish...done.
Deleted [projects/assignment1-final-378106/locations/us-central1/functions/hello_world].
samprivedevansh23@cloudshell:~/gl_final/gs:/bdl-2023-assignment1-main (assignment1-final-378106) $ gcloud functions delete return_fibonacci --quiet
Waiting for operation to finish...done.
Deleted [projects/assignment1-final-378106/locations/us-central1/functions/return_fibonacci].
samprivedevansh23@cloudshell:~/gl_final/gs:/bdl-2023-assignment1-main (assignment1-final-378106) $ gcloud functions list
NAME: al-gl-func-final
STATE: ACTIVE
TRIGGER: Event Trigger
REGION: us-central1
ENVIRONMENT: 1st gen
samprivedevansh23@cloudshell:~/gl_final/gs:/bdl-2023-assignment1-main (assignment1-final-378106) $

```

Fig. 13. List of functions after deleting the two functions. The two functions were successfully deleted

analysis, highlighting the pros and cons, as a student developer, b/w PaaS, IaaS, SaaS.

- Cloud APIs and their role in the cloud computing landscape.

SOLUTION: PART A

Serverless Computing

- 1) Serverless computing, also known as Function-as-a-Service (FaaS), is a cloud computing model in which the cloud provider manages the infrastructure needed to run and scale applications, while developers concentrate on writing and deploying code in the form of discrete functions.
- 2) The cloud provider dynamically allocates and manages computing resources based on incoming traffic and workload in a serverless architecture and users are charged only for the resources consumed by their functions.

SOLUTION: PART B

Platform-as-a-Service (PaaS)

- 1) Platform-as-a-Service (PaaS) provides a platform for developers to build and deploy applications without worrying about the underlying infrastructure.
- 2) PaaS offers pre-built services like databases, middleware, and development tools that can be leveraged to quickly build and deploy applications.
- 3) PaaS is ideal for students who want to focus on application development and deployment without having to manage the underlying infrastructure.
- 4) The main disadvantage of PaaS is the lack of flexibility, as developers are limited to the platform's pre-built services.

Infrastructure-as-a-Service (IaaS)

- 1) Infrastructure-as-a-Service (IaaS) provides virtualized computing resources like virtual machines, storage, and networking to users, allowing them to build, manage and deploy their applications on top of this infrastructure.
- 2) IaaS offers greater flexibility than PaaS, as users can customize the infrastructure according to their needs.
- 3) However, the downside of IaaS is that it requires more expertise in managing infrastructure and is less suitable for students who do not have a lot of experience in managing cloud infrastructure.

Software-as-a-Service (SaaS)

- 1) Software-as-a-Service (SaaS) provides complete software applications that can be accessed over the internet, eliminating the need for installation or maintenance.
- 2) SaaS is ideal for students who want to use ready-to-use software applications without worrying about managing infrastructure or performing updates.
- 3) However, the downside of SaaS is the lack of customization, as users are limited to the features and functionality provided by the software application.

SOLUTION: PART C

- 1) Cloud APIs are collections of protocols, tools, and interfaces that enable developers to interact with cloud services provided by cloud providers, such as storage, computing, and networking.
- 2) Cloud APIs enable developers to programmatically access and control cloud resources, allowing them to build custom applications, automate workflows, and integrate cloud services with other tools and applications.
- 3) Cloud APIs play a role in the cloud computing landscape by providing a standardized way for developers to interact with cloud services, reducing the complexity of managing cloud infrastructure, and allowing developers to focus on developing applications.
- 4) Cloud APIs also allow for seamless integration of different cloud services, making it easier for developers to create complex applications that use multiple cloud services.