

Mohamed DHIFAOUI

Partie 1 :


Question 1 :

Il manque les fichiers :

- Kinesis_datastream : qui contient le stream d'entrée.
- Destination_s3_bucket : contient la déclaration de bucket de destination.

Question 2 :

Kinesis_datastream :

 kinesis_datastream.tf - Bloc-notes


Fichier Edition Format Affichage Aide

```
resource "aws_kinesis_stream" "kinesis_input_stream" {
  name           = "input-stream"    #le nom du service qui va apparaitre dans amazon
  shard_count    = 1                 #configurer le parallélisme 1 donc aucun
  # The default retention period is 24h
  retention_period = 24              #le tps pdt lequel garder lesdonnées

  shard_level_metrics = [
    "IncomingBytes",
    "OutgoingBytes",
  ]

  tags = {
    Environment = "test"             #tag pour pouvoir rechercher les ressources plus facilement dans amazon
  }
}
```

Destination_s3_bucket :

 destination_s3_bucket.tf - Bloc-notes

Fichier Edition Format Affichage Aide

```
# Create a S3 bucket - destination of the data pipeline
resource "aws_s3_bucket" "summary_destination" {
  bucket = "analytics-destination-wxcft6"
  acl     = "private"

  tags = {
    Name           = "S3 bucket"
    Environment    = "test"
  }
}
```

Question 3 :

```
Plan: 17 to add, 0 to change, 0 to destroy.
```

Changes to Outputs:

```
+ account_id           = "944053094666"
+ caller_arn           = "arn:aws:iam::944053094666:root"
+ caller_user          = "944053094666"
+ current_region       = "us-east-1"
+ ec2_vm_hostname      = (known after apply)
+ student              = "mohamed.dhf97@gmail.com"
+ summary_destination_bucket = "analytics-destination-wxcft6"
+ tracker              = (known after apply)
```

```
ytics:us-east-1:944053094666:application/sql-streaming-application]
```

```
Apply complete! Resources: 17 added, 0 changed, 0 destroyed.
```

Outputs:

```
account_id = "944053094666"
caller_arn = "arn:aws:iam::944053094666:root"
caller_user = "944053094666"
current_region = "us-east-1"
ec2_vm_hostname = "ec2-3-86-242-19.compute-1.amazonaws.com"
student = "mohamed.dhf97@gmail.com"
summary_destination_bucket = "analytics-destination-wxcft6"
tracker = "OwTIjrXI50"
PS C:\Users\Mohamed\Desktop\ING 3\Pipeline de traitements de données pour le Cloud\CC1\ccl-p1-terraform-weblogs-template
>
```

```
[ec2-user@ip-172-31-85-43 logs]$ ls
Fake-Apache-Log-Generator
[ec2-user@ip-172-31-85-43 logs]$ cd Fake-Apache-Log-Generator/
[ec2-user@ip-172-31-85-43 Fake-Apache-Log-Generator]$ ls
apache-fake-log-gen.py LICENSE README.md requirements.txt
[ec2-user@ip-172-31-85-43 Fake-Apache-Log-Generator]$ cp apache-fake-log-gen.py /tmp/logs/
[ec2-user@ip-172-31-85-43 Fake-Apache-Log-Generator]$ cd /etc/aws-kinesis/
[ec2-user@ip-172-31-85-43 aws-kinesis]$ cp apache-fake-log-gen.py /tmp/logs/
```

Question 4 :

Cette application simule un système d'analyse en temps réel de logs d'un site Web. Elle va prendre en entrée les logs d'un site Web au format des logs Apache, générés en continu par l'application Python «fake log generator ». Ensuite, Le générateur produit des fichiers "access_log". Enfin, on aura en sortie le résultat de l'analyse des logs, par exemple l'analyse des codes HTTP

Question 5 :

User_data : c'est les données qu'on peut utiliser pour le lancement de l'instance.

Question 6 :

L'agent Kinesis est une application logicielle Java autonome qui permet de collecter et d'envoyer facilement des données à Kinesis Data Streams. L'agent surveille en permanence l'ensemble de fichiers et envoie de nouvelles données à notre flux. Donc on arrive à assurer l'écriture dans des flux de Amazon Kinesis Data Streams à l'aide de Kinesis Agent.

Partie 2 :

1. Ingestion : AWS Kinesis Data Streams
2. Stockage : AWS S3
3. Transformation : AWS Kinesis Data Analytics
4. Exposition : Requêtes SQL

Partie 3 :**Question 4 :**

Requêtes SQL à fenêtres.

Question 5:

C'est causé par le délai de la fenêtre de la requête.

Question 6:

L'email est utilisé comme tag qui est placé sur la ressource EC2. Les tags peuvent être interrogées, utilisées pour l'automatisation et regroupées logiquement pour des fonctions telles que la facturation.