Valentin GAGNIERE
BTS SIO 1
Institution des Chartreux

RAPPORT DE STAGE

Valentin GAGNIERE
BTS SIO 1
Institution des Chartreux

Table des matières

Valentin GAGNIERE
BTS SIO 1
Institution des Chartreux

1 - Ma mission : pourquoi sécuriser les comptes publics AWS pour SNCF

SNCF utilise le cloud AWS, d'autres fournisseurs de services cloud et a aussi des datacenters physiques.

Elle héberge plus de 250 applications dans le cloud AWS.

Dans le cloud AWS, tout est organisé à partir de comptes AWS. Un compte AWS est un endroit pour stocker ses ressources cloud (services, plateformes, etc). Quand une entreprise utilise le cloud AWS elle suit des bonnes pratiques pour organiser et sécuriser ses comptes, surtout si ses ressources clouds communiquent avec les données de ses datacenters physiques.

La sécurité est très importante alors les comptes AWS des applications sont séparés : il y a ce qui est de la production critique et il y a ce qui n'en est pas (le développement, les tests, les recettes pour les métiers par exemple).

Avant de commencer une application, SNCF permet à ses clients de tester toutes les ressources AWS disponibles sur le marché sans appliquer de restrictions dans des comptes AWS complètement séparés des autres comptes et tous indépendants les uns des autres : en fournissant à ses clients des comptes AWS publics non reliés au Système d'information de SNCF.

Ma mission est de participer à la sécurisation des comptes publics en activant l'audit réseau à ces comptes de façon automatique.

Il est important de sécuriser les comptes publics de SNCF, même si les clients sont très libres dessus :

- Parce que c'est SNCF qui possède ces comptes, donc c'est sa responsabilité en cas d'utilisation frauduleuse du compte public
- Parce qu'en terme d'image pour l'entreprise, un incident de sécurité a un impact médiatique même sur un compte de test
- Si on ne peut pas éviter l'incident de sécurité, il faut tracer ce qui a été fait (pour porter plainte, pour expliquer ce qu'il s'est passé, pour se protéger plus tard)

Valentin GAGNIERE
BTS SIO 1
Institution des Chartreux

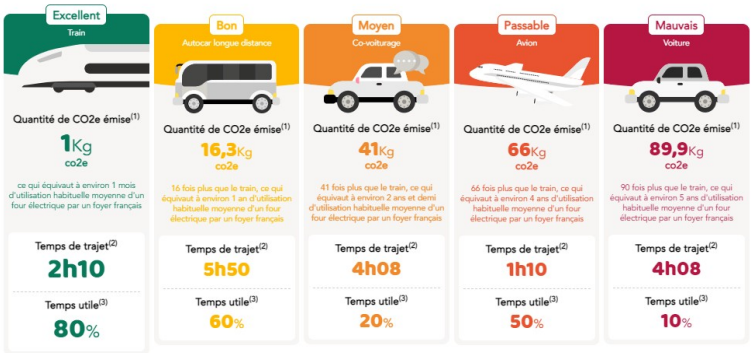
2 - Le contexte du stage : SNCF et AWS

a. SNCF, le groupe

Depuis 1937, SNCF (Société Nationale des Chemins de Fer Français) a pour rôle de transporter des biens et des personnes en toute sécurité. 120 000 personnes travaillent dans la société SNCF. Il y a 275 000 employés par SNCF et ses filiales.

SNCF réalise un Chiffre d'Affaire de 35 milliards d'euros par an (25 Netflix, 85 FB, 380 Amazon pour comparer). SNCF n'est pas un pureplayer d'internet, cette société est ancienne et gère la circulation de 15 000 trains par jours sur 28 000 kilomètres de rails en France. Il y a 5,4 milliards de voyageurs par an dans les trains et les RER. Il y a de très nombreuses applications (+ de 1500 en interne, réparties dans des établissements (gares, etc), des Datacenters et chez des Clouders) pour la gestion de SNCF et pour ses clients (1 majoritairement) .

Le groupe SNCF a aussi comme objectif d'être à zéro émission en 2035[1] et d'aider la France à être à l'équilibre carbone en 2050 [2].
Depuis mars 2021, il existe un comparateur de mobilité sur oui.sncf [3].



b. SNCF utilise le cloud AWS

Amazon Web Services, dit AWS, est une division du groupe américain de commerce électronique Amazon, spécialisée dans les services de cloud computing à la demande pour les entreprises et particuliers. En 2020, AWS génère 63 % du chiffre d'affaires d'Amazon et représente la première source de profit de l'entreprise.

Un clouder est un fournisseur de services de cloud computing : c'est l'accès à des services informatiques (serveurs, stockage, mise en réseau, logiciels) via Internet. Les principaux services proposés en cloud computing sont le SaaS (Software as a Service), le PaaS (Platform as a Service) et le IaaS (Infrastructure as a Service). On distingue généralement : - le cloud public — accessible par Internet, le cloud d'entreprise ou privé — accessible uniquement sur un réseau privé —, le cloud hybride.

1 https://www.lettreducheminot.fr/divers/sncf-vise-zero-emission-2035/

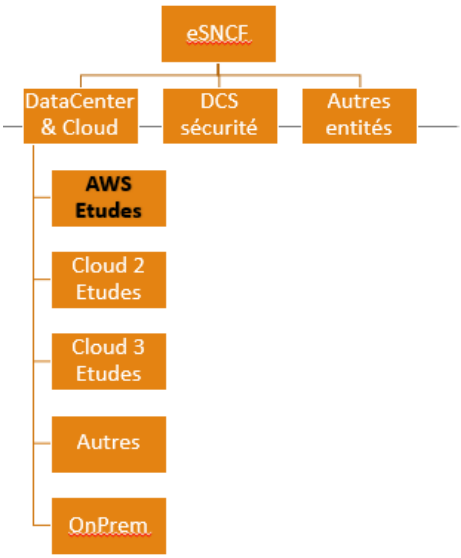
2 https://www.sncf.com/fr/groupe/newsroom/zen2050

3 https://www.oui.sncf/train/comparateurco2

Le cloud computing permet aux entreprises de minimiser les coûts d'infrastructure informatique et de bénéficier d'une adaptation des ressources en fonction des fluctuations de l'usage.

Amazon le géant du e-commerce possédait ses propres datacenters et l'usage de ses serveurs était très intense sur trois mois de l'année seulement. Amazon a trouvé une solution pour louer la puissance de calcul non utilisée de ses datacenters lorsque le site marchand n'en avait pas besoin. C'était la naissance de Amazon Web Services.

c. L'informatique à SNCF : eSNCF / DataCenters & Cloud / AWS-Etudes



SNCF a une division informatique principale appelée eSNCF regroupant presque 2000 employés.

Au sein de cette division très vaste se trouve une équipe d'ingénierie et de gestion de services appelée DataCenter & Cloud. Cette division intègre l'équipe AWS Etudes où j'ai fait mon stage et qui s'occupe de gérer le service du cloud AWS pour les projets de SNCF.

d. Les emplois dans AWS Etudes

- **Directeur des plateformes Clouds de SNCF**

Le responsable des plateformes clouds s'occupe des contrats financiers avec les fournisseurs de clouds, des services attendus et de la stratégie de l'entreprise dans le cloud quel que soit le fournisseur.

- **Cloud Service Architect AWS**

Il s'agit de transposer la stratégie cloud de l'entreprise en une infrastructure technique et de proposer un ensemble de services pour porter l'innovation des projets, en suivant les bonnes pratiques AWS, et en respectant les budgets.

- **Expert-e-s et Architectes AWS**

Les experts et architectes AWS sont les référents et experts des services AWS. Ils/Elles travaillent à la création et l'évolution de l'infrastructure AWS du groupe. Ce sont les piliers techniques de l'équipe. Ils/Elles répondent aussi aux demandes des projets en suivant les bonnes pratiques d'AWS en termes de sécurité, de scalabilité (mise à l'échelle) et de coûts.

- **Expert-e-s Réseau AWS**

Il/Elle travaille avec des experts cloud (quel que soit le fournisseur) pour assurer la scalabilité et la fiabilité des infrastructures réseaux des clouds publics et hybrides du groupe SNCF. Il faut de très bonnes connaissances de la partie réseaux IP réseaux (protocoles de routage, découpage réseaux, interconnexion onpremise/cloud, LoadBalancing, firewalling, sécurité, monitoring) et des composants réseaux de plusieurs clouders (principalement AWS ici).

- **Expert-e-s Sécurité AWS**

Il/Elle participe à la montée en maturité de la cyber sécurité sur les différentes plateformes Cloud et AWS en particulier. Il y a de nombreuses missions d'audit et de prescription. Il faut accompagner les équipes techniques dans la mise en œuvre de la sécurité au sein des services et fournir les préconisations de sécurité pour la construction et l'évolution de l'infrastructure. Il/Elle définit comment mettre en œuvre les mesures de remédiation en accord avec la sécurité du groupe.

- **FinOps**

Les finOps sont les Financiers des Opérations. Leur rôle est d'optimiser l'utilisation du cloud pour réaliser des économies en étudiant les consommations et en gérant le contrat avec le fournisseur.

3 – Sécurisation des comptes publics par l'activation des FlowLogs sur VPC

a. Qu'est-ce qu'un compte public AWS

Un compte AWS contient des ressources isolées, avec une facturation indépendante, un accès indépendant, et des rôles utilisateurs indépendants.

Par défaut, les ressources réseaux d'un compte AWS ne sont pas accessibles par les autres comptes AWS. -> C'est le compte public

C'est comme si une personne de SNCF prenait un compte AWS avec sa carte bleue = sans bénéficier des protections (firewall...) et des services du SI SNCF (authentification, surveillance…)

b. Qu'est-ce qu'un VPC

Un VPC (Virtual Private Cloud) est un service qui permet de lancer des ressources AWS dans un réseau virtuel isolé de manière logique que l'on définit. On conserve la totale maîtrise de l'environnement du réseau virtuel, y compris pour la sélection d'une plage d'adresses IP, la création de sous-réseaux et la configuration de tables de routage et de passerelles réseau.

c. Qu'est-ce que les FlowLogs

Les FlowLogs sont une fonctionnalité qui permet de capturer des informations sur le trafic IP entrant et sortant des interfaces réseau d'un VPC.

En clair, les FlowLogs sont des journaux d'activités quotidiens qui capturent tous les flux réseaux entrant et sortant (date et heure de connexion, date de création d'un VPC...)

De ce fait, les FlowLogs sont très pratiques pour un certain nombre de tâches, notamment :

Pour Diagnostiquer les règles de groupe de sécurité trop restrictives.

Pour Surveiller le trafic des instances

Pour Déterminer la direction du trafic vers et depuis les interfaces réseau

d. Explication théorique de l'objectif de la mission

Ma mission consistait à réaliser un script permettant de :

Parcourir toutes les configurations des comptes publics souscrits par SNCF

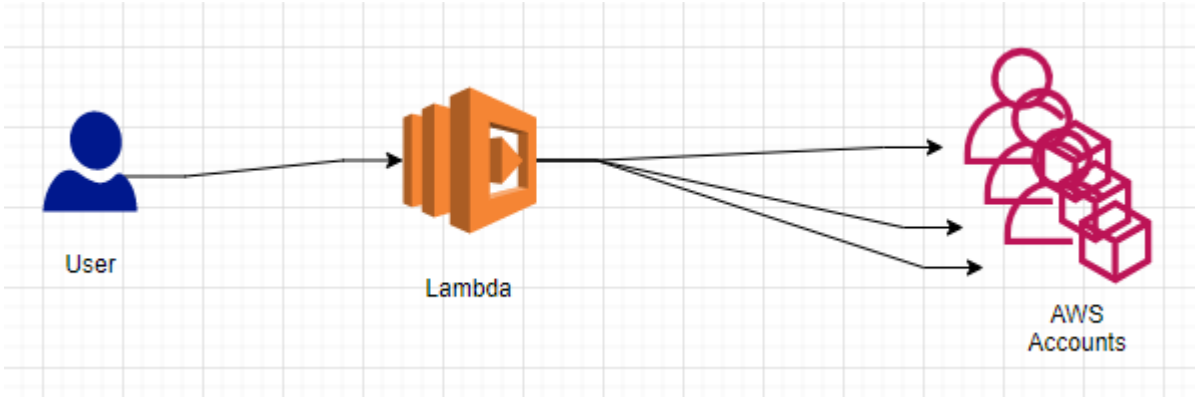
Vérifier l'activation des FlowLogs sinon les configurer avec les différents ports requis.

Dans ce script il faut adresser plusieurs comptes AWS par lot.

Dans le schéma ci-dessous le script est représent par user et lambda est la fonction qui le porte.

Lamnda est un service managé AWS sans serveur qui vous permet d'exécuter du code sans provisionner ou gérer des serveurs. Lambda alloue automatiquement la puissance d'exécution de calcul et exécute le code en fonction de la demande ou de l'événement entrant, pour n'importe

quelle échelle de trafic.



e. Construction d'un script qui répond à l'objectif

Quelles sont les technologies utilisées :

Le langage principal du script est python car c'est un langage qui a beaucoup de possibilités et qui combiné avec Boto3 qui est une librairie python permettant d'interagir avec AWS-CLI (l'interface d'administration en ligne de commande d'AWS).

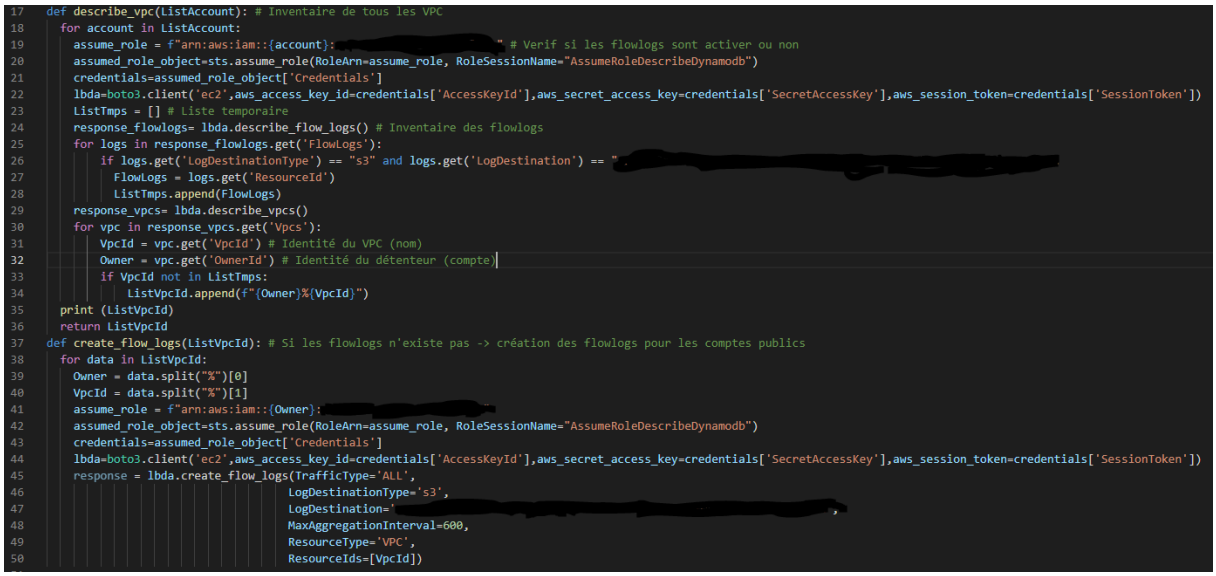
Ces technologies m'ont permis de faire le script qui répondait au cahier des charges.

Tests prévus :

	Fonctionnalité	Test à effectuer	Résultat
1	Enumération visuelle des VPC	Enumération visuelle des VPC à l'aide de la console pour vérifier le bon fonctionnement de l'énumération	Enumération visuelle des VPC, OK
2	Enumération visuelle des comptes AWS dans une liste de VPC	Enumération visuelle des comptes AWS qui les retourne sous forme d'inventaire dans le script	Liste de VPC remplie avec les comptes AWS, l'énumération des comptes dans une liste de VPC est donc fonctionnelle
3	Vérification de l'activation des FlowLogs dans les VPC	Enumération des FlowLogs de chaque VPC	Enumération des FlowLogs pour chaque VPC, OK
4	Activation des FlowLogs dans les VPC qui ne les ont pas activés	Activation des FlowLogs sur les VPC qui ne l'ont pas d'activé	Activation automatisée des FlowLogs sur les VPC, OK

5	Création des FlowLogs dans les VPC s'ils sont inexistants	Vérification si les FlowLogs dans les VPC sont présent sinon le script les créer	Création des FlowLogs précédemment inexistants dans les VPC, OK

f. Explication d'un morceau de script



```
17  def describe_vpc(ListAccount): # Inventaire de tous les VPC
18    for account in ListAccount:
19      assume_role = f"arn:aws:iam::{account}:[redacted]" # Verif si les flowlogs sont activer ou non
20      assumed_role_object=sts.assume_role(RoleArn=assume_role, RoleSessionName="AssumeRoleDescribeDynamodb")
21      credentials=assumed_role_object['Credentials']
22      lbda=boto3.client('ec2',aws_access_key_id=credentials['AccessKeyId'],aws_secret_access_key=credentials['SecretAccessKey'],aws_session_token=credentials['SessionToken'])
23      ListTmps = [] # Liste temporaire
24      response_flowlogs= lbda.describe_flow_logs() # Inventaire des flowlogs
25      for logs in response_flowlogs.get('FlowLogs'):
26          if logs.get('LogDestinationType') == "s3" and logs.get('LogDestination') == "[redacted]
27            FlowLogs = logs.get('ResourceId')
28            ListTmps.append(FlowLogs)
29      response_vpcs= lbda.describe_vpcs()
30      for vpc in response_vpcs.get('Vpcs'):
31          VpcId = vpc.get('VpcId') # Identité du VPC (nom)
32          Owner = vpc.get('OwnerId') # Identité du détenteur (compte)
33          if VpcId not in ListTmps:
34              ListVpcId.append(f"{Owner}%{VpcId}")
35    print (ListVpcId)
36    return ListVpcId
37  def create_flow_logs(ListVpcId): # Si les flowlogs n'existe pas -> création des flowlogs pour les comptes publics
38    for data in ListVpcId:
39      Owner = data.split("%")[0]
40      VpcId = data.split("%")[1]
41      assume_role = f"arn:aws:iam::{Owner}:[redacted]
42      assumed_role_object=sts.assume_role(RoleArn=assume_role, RoleSessionName="AssumeRoleDescribeDynamodb")
43      credentials=assumed_role_object['Credentials']
44      lbda=boto3.client('ec2',aws_access_key_id=credentials['AccessKeyId'],aws_secret_access_key=credentials['SecretAccessKey'],aws_session_token=credentials['SessionToken'])
45      response = lbda.create_flow_logs(TrafficType='ALL',
46                                       LogDestinationType='s3',
47                                       LogDestination='[redacted]',
48                                       MaxAggregationInterval=600,
49                                       ResourceType='VPC',
50                                       ResourceIds=[VpcId])
```

A la ligne 17 j'utilise la fonction ***describe_*vpc** car c'est cette fonction qui interagit avec boto3 qui permet de faire l'inventaire des VPC.

De la ligne 18 jusqu'à la ligne 36, le script vérifie les FlowLogs des VPC ensuite il crée une liste pour les répertorier (ligne 24 : ***response_flowlogs= lbda.describe_flow_logs()***), j'utilise la fonction ***describe_flow_logs*** car elle sert à faire l'inventaire des FlowLogs des VPC.

Ensuite de la ligne 30 à 50, la boucle ***for*** récupère l'identité des VPC ainsi que leurs détenteurs (comptes owners) puis vérifie si les VPC sont dans la liste python ***Tmps*** et s'ils n'y sont pas, le script les ajoutes dans la liste python puis crée les FlowLogs pour les VPC qui n'ont pas de FlowLogs.

J'utilise la commande ***create_flow_logs*** de boto3 car c'est la seule commande qui permet de créer des flow logs pour des VPC.

https://boto3.amazonaws.com/v1/documentation/api/latest/reference/services/ec2.html

g. Importance de la documentation dans le script et versioning

Importance de la documentation dans le code :

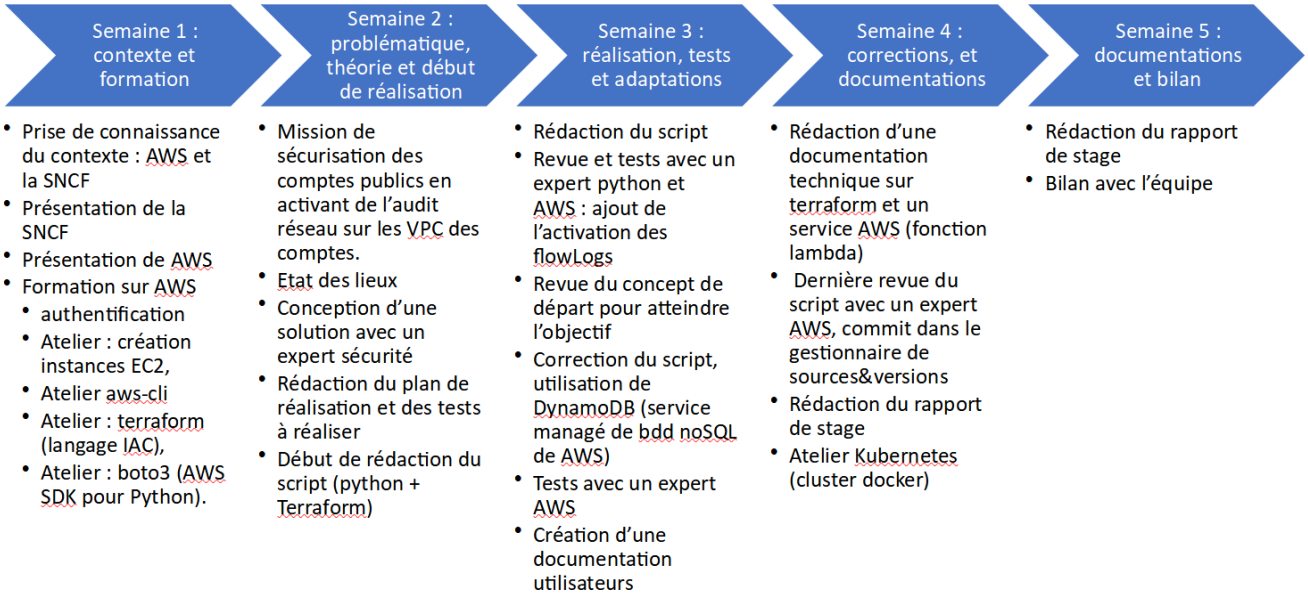
J'ai commenté mon code pour que mes collègues puissent relire rapidement et facilement mon code sans devoir relire le code entier
J'ai commenté mon code aussi pour ne pas oublier pourquoi j'avais utilisé telle ou telle fonction.

Importance du versionning :

Il est nécessaire d'utiliser un gestionnaire de source comme git pour collaborer entre les différents membres d'une équipe.
Cela permet aussi de mettre à jour le projet simplement (on peut gérer les modifications et leur historique). Cela sert également à sauvegarder le travail sur une plateforme distante en cas de suppression accidentelle du script sur la machine du développeur.

4 – Ma progression et bilan du stage

a. La chronologie



Explications sur les formations et ateliers :

Atelier AWS-cli : Apprentissage de l'administration en utilisant l'interface en ligne de commande d'AWS.

Atelier Terraform : Apprentissage du langage Terraform qui permet de coder des fonctions AWS.

Atelier boto3 : Apprentissage de la librairie boto3 qui permet d'interagir avec python et AWS.

Atelier création EC2 : Création d'une instance EC2 (machine virtuelle) dans le cloud AWS.

J'ai suivi les workshops aws et lu beaucoup de documentation en anglais et en français sur le site d'AWS.

Plusieurs membres de l'équipe ont passé du temps avec moi pour m'expliquer les concepts du cloud AWS et m'ont aidé à relire mon script.

b. Bilan du stage

Pendant ce stage d'une durée d'un mois au sein de SNCF dans une équipe d'experts et d'ingénieur/es j'ai eu la chance de découvrir le cloud AWS et son fonctionnement et également d'apprendre comment fonctionnent les différents services AWS en lisant beaucoup de documentation anglaise. J'ai pu également pendant ce stage apprendre le langage Terraform en totale autonomie.

Lors de ce stage j'ai eu la chance d'avoir des collègues très sympa et chaleureux, de plus il y avait une très bonne ambiance dans l'équipe. J'ai parlé avec de nombreux collègues concernant leur parcours professionnel et ils m'ont tous encouragé à continuer mes études jusqu'en école d'ingénieur pour me permettre d'acquérir encore plus de compétences. Ils m'ont également encouragé dans ma démarche de passer des certifications AWS pour apporter une plus-value dans mon CV.

Lors de ce stage j'ai également appris beaucoup de choses et dans une bonne ambiance de travail. J'ai aussi appris à partager l'avancement de mon travail avec une équipe en utilisant la méthode agile et en participant aux événements du Daily.

Pendant ma période de stage, j'ai découvert la grande diversité dans les métiers du Cloud AWS et cela a été pour moi un tournant pour ma future carrière car j'étais en grande réflexion sur mon avenir professionnel mais maintenant j'ai fait mon choix de me spécialiser plus tard dans la cybersécurité cloud.

Je tiens à remercier ma maître de stage Aurélie CAMBOURIAN ainsi que toute l'équipe AWS Etudes pour leur accueil, leur professionnalisme et leur gentillesse.