

# TP 0 — Partie 1: Découverte de l'interface d'Amazon Web Service (AWS)

## Objectifs

Ce TP a pour but de vous familiariser avec l'interface d'Amazon Web Service (AWS). Pendant ce TP vous allez :

- Créer une compte AWS educate
- Copier des données dans votre espace de stockage Amazon Simple Storage Service (S3)
- Lancer une machine virtuelle (VM) et s'y connecter (**pas aujourd'hui**)
- Eteindre votre VM

Le cluster que vous allez créer pendant le TP pourra être cloné lors de vos prochains TP.

## 1. Création du compte AWS educate

Suivez les instructions à partir du mail AWS Educate reçu sur votre adresse ENSAI.

Créez un compte puis connectez-vous.

## 2. Exploration

Dans l'onglet "Services", trouvez:

- EC2, le service de calcul
- S3, le service de calcul
- la section dédiée aux bases de données
- la section dédiée au *machine-learning*
- la section dédiée à l'analyse de données
- la section dédiée à la gestion des coûts

Vérifiez que votre crédit de 50\$ est bien activé.

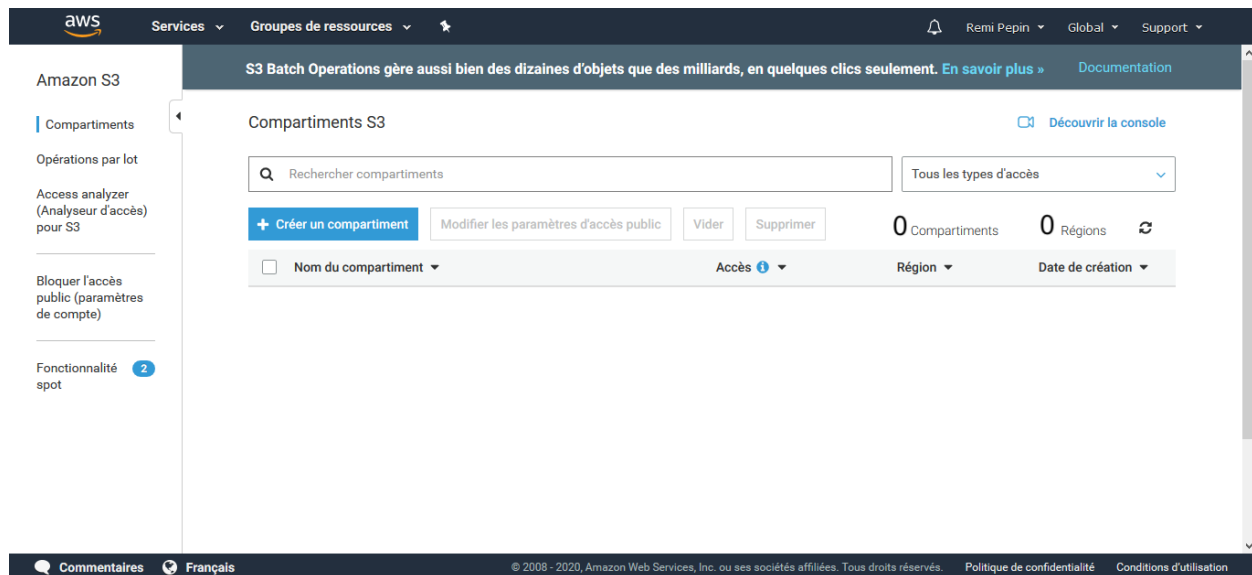
## 3 Création d'un espace de stockage Amazon Simple Storage Service (S3)

**Amazon Simple Storage Service (S3)** est la solution de base que propose AWS pour stocker vos données de manière pérenne. Amazon dit assurer une durabilité de vos données de 99,999999999 %. Cela signifie que si vous stockez 10 000 000 objets avec Amazon S3, vous pouvez vous attendre à perdre en moyenne un objet unique une fois tous les 10 000 ans.

Ce stockage est assuré à coût relativement élevé (de l'ordre de ~0,02 \$/Go/mois), sachant que vous payez en sus les opérations de lecture (de l'ordre de ~0,09 \$/Go ; les écritures sont gratuites). 1 To de données vous coûte ainsi 240€ à l'année. Pour comparaison, un disque dur externe d'1 To coûte actuellement ~40€, et un cloud-storage pour particulier coûte ~10€ / mois pour 2 To. S3 est ainsi destiné à des données d'usage régulier. D'autres offres de stockage existent comme les archives, pour des données utilisées moins régulièrement, ou les bases de données. \_

Tous les services que vous propose AWS peuvent nativement lire depuis et écrire vers S3. Ainsi, les programmes que vous exécutez, les données que vous traitez... peuvent être importés dans S3. Chaque élément hébergé dans S3, appelé "objet", est accessible par une URL unique. Vous pouvez restreindre ou au contraire étendre les droits d'accès à vos objets.

3-1 : Dans la barre de recherche, cherchez "S3" et cliquez dessus



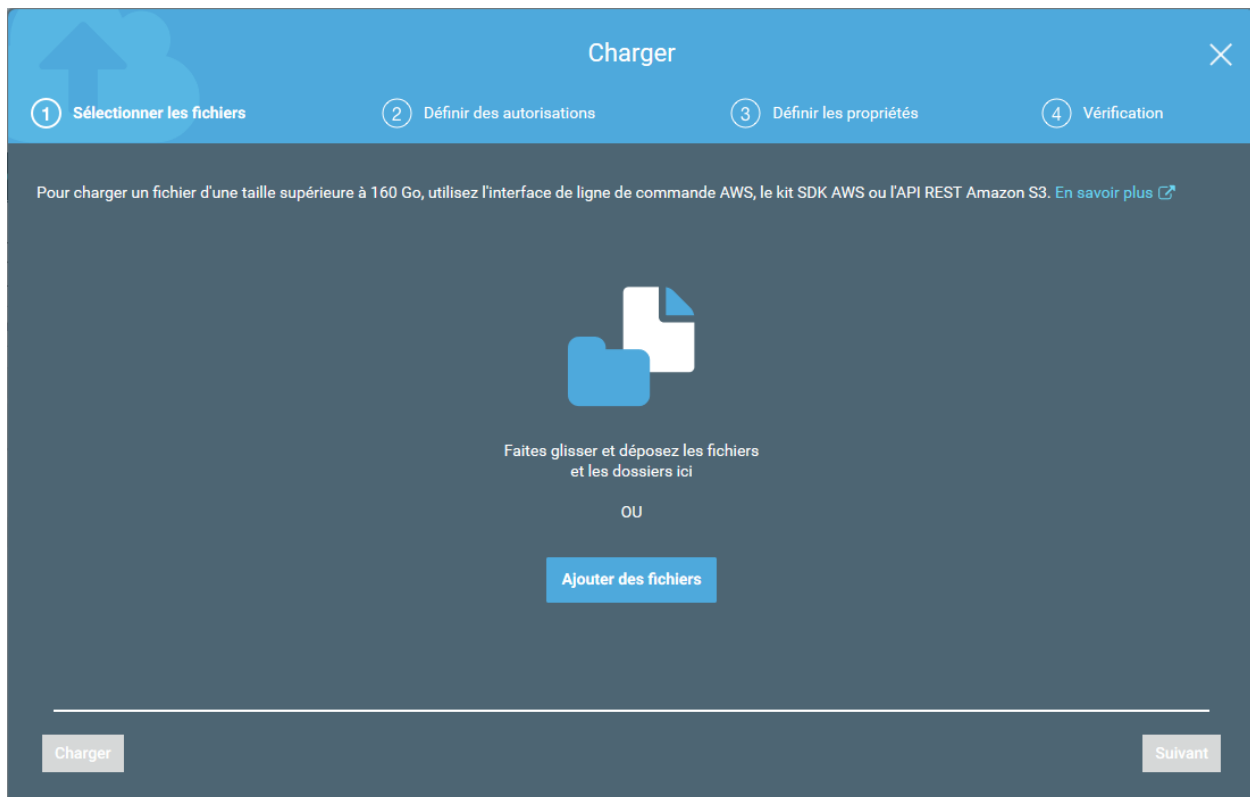
3-2 Cliquez sur “Créer un compartiment” (en anglais un “bucket”)

3-3 Choisissez un nom unique à votre compartiment

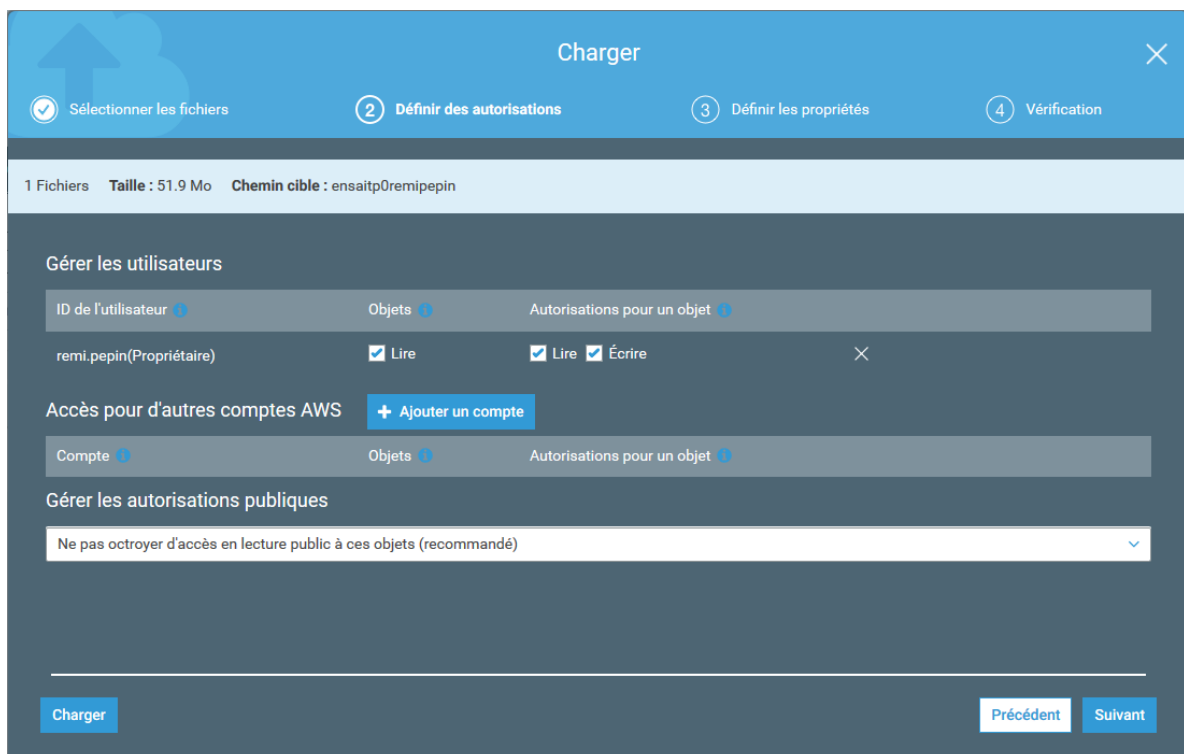
3-4 Laissez toutes les valeurs par défaut et créez votre compartiment

## 4. Copie des données dans votre espace de stockage

4-1 À partir du bouton “Charger”, ajoutez un fichier (par exemple un fichier CSV) à votre compartiment. Laissez toutes les valeurs par défaut.



- Un des écrans permet de gérer les permissions liées à votre fichier: qui y a accès, avec quels droits, et si vous voulez rendre votre fichier public. Par défaut, votre fichier est privé.



- Un des écrans vous permet de choisir la classe de stockage. S3 en dispose de quelques une, chacune avec

des cas d'utilisation différent. Par exemple le stockage "glacier" (EN: "glaciere") permet d'archiver des données pour un coût faible, mais vos données auront un délais d'accès et une facturation supplémentaire par accès.

Charger

Sélectionner les fichiers

Définir des autorisations

3 Définir les propriétés

4 Vérification

1 Fichiers
Taille : 51.9 Mo
Chemin cible : ensaitp0remipepin

Classe de stockage

Choisissez une classe de stockage en fonction de votre cas d'utilisation et de vos exigences en matière d'accès. [En savoir plus](#) ou consulter [Tarification Amazon S3](#)

Classe de stockage	Conçues pour	Zones de disponibilité	Durée de stockage minimum	Taille d'objet en octets minimum	Frais de surveillance et d'automatisation	Frais d'extraction
<input type="radio"/> Standard	Données fréquemment consultées	≥ 3	-	-	-	-
<input type="radio"/> Hiérarchisation intelligente	Données à longue durée de vie avec des modèles d'accès changeants ou inconnus	≥ 3	30 jours	-	Des frais par objet s'appliquent	-
<input type="radio"/> Standard-APF	Données à longue durée de vie rarement consultées	≥ 3	30 jours	128 Ko	-	Des frais par Go s'appliquent
<input type="radio"/> Une zone-APF	Données non critiques à longue durée de vie et rarement consultées	≥ 1	30 jours	128 Ko	-	Des frais par Go s'appliquent
<input type="radio"/> Glacier	Données d'archive avec un délai de récupération allant de quelques minutes à plusieurs heures	≥ 3	90 jours	40 Ko	-	Des frais par Go s'appliquent

Charger

Précédent

Suivant

4-2 Une fois le chargement terminé cliquez sur votre fichier

- Dans **Présentation** > **URL d'objet**, vous voyez l'URL de votre fichier, pour pouvoir s'y référencer par la suite, par exemple pour y donner accès sur Internet. (Depuis les services Amazon, l'adresse est raccourcie en `s3://<bucket-name>/<path-to-file>.`)
- "Chemin de copie" est une mauvaise traduction de "Copy Path". Il faut s'y habituer, seule une toute petite partie d'AWS est traduite, et plutôt mal.
- Dans l'onglet "Sélectionner depuis" vous avez accès à des outils d'exploration de vos données. Remarquez que sur AWS tous les traitements coûtent! À tout moment, vous pouvez aller dans l'onglet facturation pour vérifier vos dépenses.

1990.csv Dernière version ▼

Présentation

Propriétés

Autorisations

Sélectionner depuis

Ouvrir

Télécharger

Télécharger en tant que

Rendre public

Chemin de copie

**Propriétaire**  
0ea6bd58e7591a06c23bbf9d47c34182a6648fe846df94b03801aff6c4b5d498

**Dernière modification**  
janv. 16, 2020 1:53:55 PM GMT+0100

**Etag**  
a227e015e70babca1058aa0639aa7d79-4

**Classe de stockage**  
Standard

**Chiffrement côté serveur**  
Aucun

**Taille**  
51.9 Mo

**Clé**  
1990.csv

**URL d'objet**  
<https://ensaitp0remipepin.s3.eu-west-3.amazonaws.com/1990.csv>

## 5. Créer une machine virtuelle (*pas aujourd'hui*)

5-1 : Naviguez vers EC2

5-2 : Cliquez sur Lancer une instance

5-3 : Vous devez choisir l'image de la machine à créer. Une image contient notamment le système d'exploitation. Choisissez la première: Amazon Linux 2 AMI. (Vous remarquez que certaines sont "éligibles à l'offre gratuite" pour les nouveaux utilisateurs AWS.)

5-4 : Vous choisissez ensuite la configuration de VM. Par exemple, vous pouvez choisir une machine d'usage général à 2 cœurs `t2.micro`, elle aussi éligible à l'offre gratuite.

5-5 : D'autres options sont disponibles (configuration, stockage, sécurité...), mais nous pouvons d'ores et déjà cliquer sur "Vérifier et lancer".

5-6 : Vérifiez les informations puis validez.

5-7 : Générez une paire de clés SSH et stockez votre clé privée.

5-8 : Et voilà! Votre VM est en cours de lancement... revenez dans quelques minutes!

## 6. Se connecter à sa VM (*pas aujourd'hui*)

## 7. Eteindre la VM

Le coût d'une VM est fonction de son temps d'utilisation, pas du travail qu'il accomplit. Ainsi, une fois le travail effectué, vous *devez* éteindre vos VMs ! **Même si le coût horaire est bas, faire tourner une machine EC2 pendant 1 semaine se chiffre en dizaines d'euros!**

Pour éteindre votre VM, allez sur la page d'accueil EC2 > Instances en cours d'exécution ou sur la bar de navigation Instances > Instances. Selon le type d'instance, vous pouvez l'arrêter (**EN:** *stop*, pour la réutiliser plus tard), ou la résilier (**EN:** *terminate*, i.e. la supprimer). Dans les deux cas, les données en mémoire et le stockage local sont perdus, mais dans le premier cas, la configuration (URL et IP) sont conservés.

Vous avez le choix entre ... et .... et cliquez sur "Résilier". Cela va éteindre *définitivement* votre cluster. Vous ne pourrez plus le relancer, à la différence d'une instance Amazon Elastic Compute Cloud (EC2). Mais vous allez pouvoir relancer un cluster en clonant un ancien. Et ainsi ne pas avoir à refaire toutes la configuration. Par contre vous devrez changer l'IP utilisée pour vous y connecter.