THEME ABORDE: S3

SOMMAIRE:

- •Qu'est-ce que c'est?
- •Intérêts
- Utiliser S3
- Offre gratuite/payante
- •Gérer les droits
- •Le versioning

I - Qu'est-ce qu'Amazon S3?

Amazon S3 (Amazon Simple Storage Service) est un site d'hébergement de fichiers proposé par Amazon Web Services. Amazon S3 propose des services de stockage à travers des services Web (REST, SOAP et BitTorrent).

Amazon a lancé le service Amazon S3, son premier service Web, aux États-Unis en mars 2006 et en Europe en novembre 2007.

Amazon Simple Storage Service est une solution de stockage sur Internet. Il est conçu pour faciliter l'informatique à l'échelle d'Internet pour les développeurs.

Amazon S3 offre une interface simple de services web qui vous permet de stocker et de récupérer n'importe quelle quantité de données, à tout moment, de n'importe où sur Internet. Il permet aux développeurs d'accéder à la même infrastructure de stockage de données hautement évolutive, fiable, rapide, peu coûteuse qu'Amazon utilise pour faire fonctionner son propre réseau mondial de sites. Ce service vise à maximiser les avantages d'échelle et à en faire bénéficier les développeurs.

II - Intérêt d'utiliser Amazon S3:

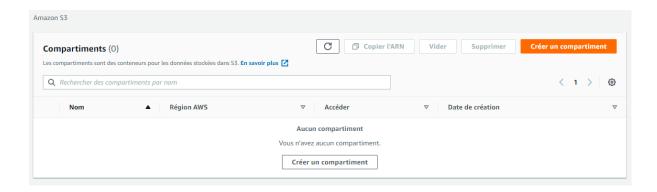
Selon Amazon, l'architecture a été conçue pour fournir extensibilité, disponibilité, rapidité d'accès et prix abordable.

1. Utiliser S3 dans la console AWS

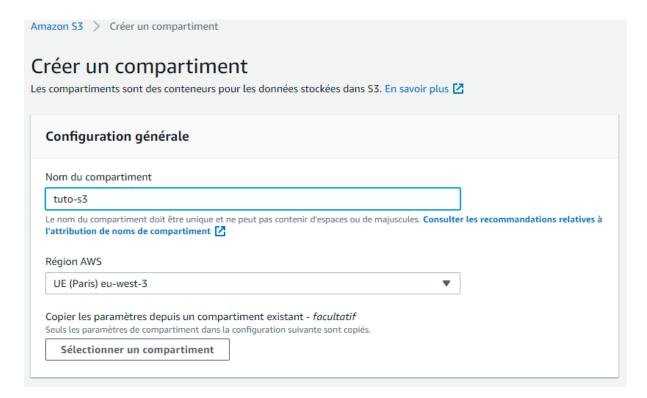
Nous allons voir comment utiliser AWS pour stocker nos fichiers. Amazon propose Amazon Simple Storage Solution (S3) qui nous permet de stocker nos données. Dans ce didacticiel, nous créerons un compartiment Amazon S3, y téléchargeront un fichier et récupérerons ce fichier avant de le supprimer.

Ce didacticiel peut être complété avec le niveau gratuit d'AWS.

En fait, il ne s'agit pas à proprement parler de système de fichiers mais de Bucket. Un Bucket est un compartiment de fichiers. Nous allons créer un compartiment dans lequel nous allons stocker nos fichiers. Pour cela il faut se rendre dans le service S3 et cliquer sur « Créer un compartiment ».

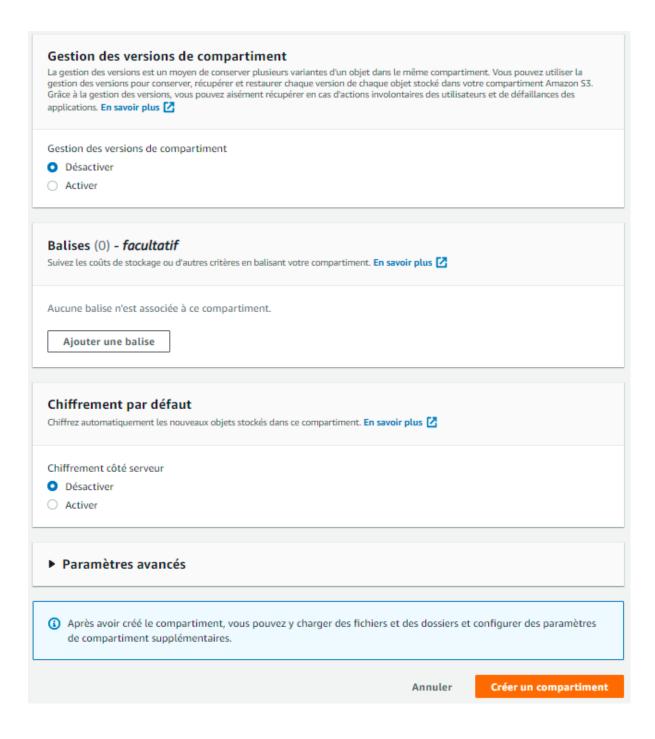


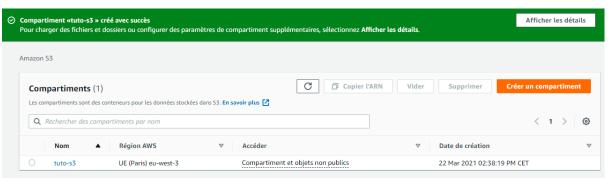
Cliquer sur "Créer un compartiment":



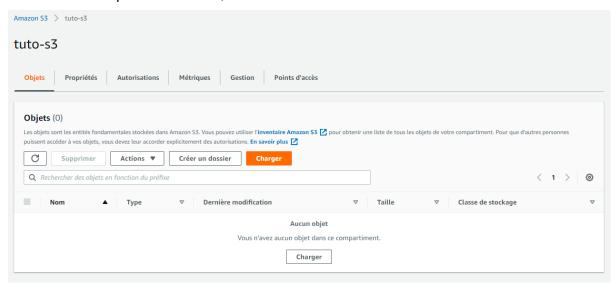
Remplir le champs "Nom du compartiment" et "Créer un compartiment":

On a gardé les paramètres par défaut.

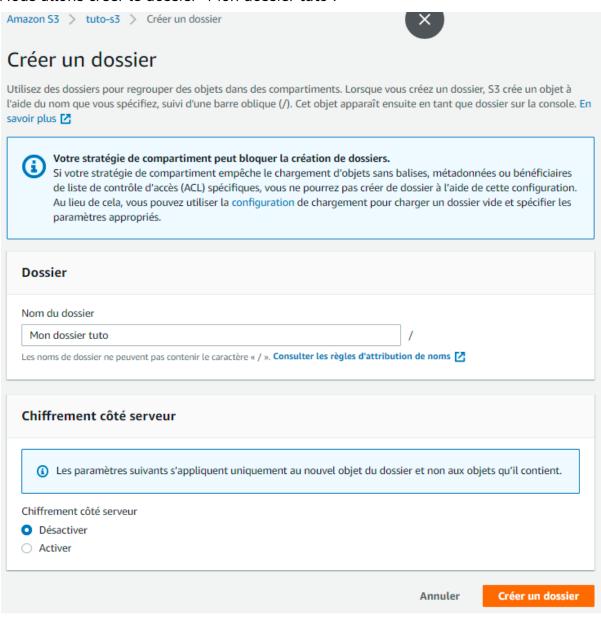


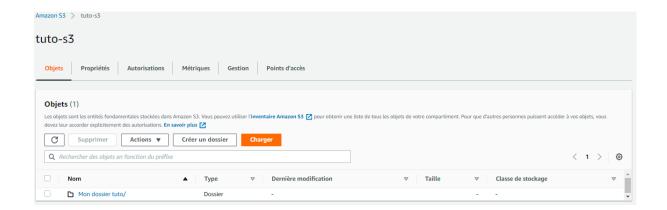


Une fois le compartiment créé, nous allons créer un dossier.

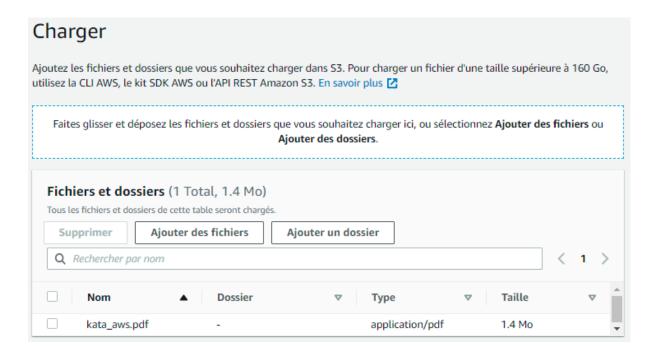


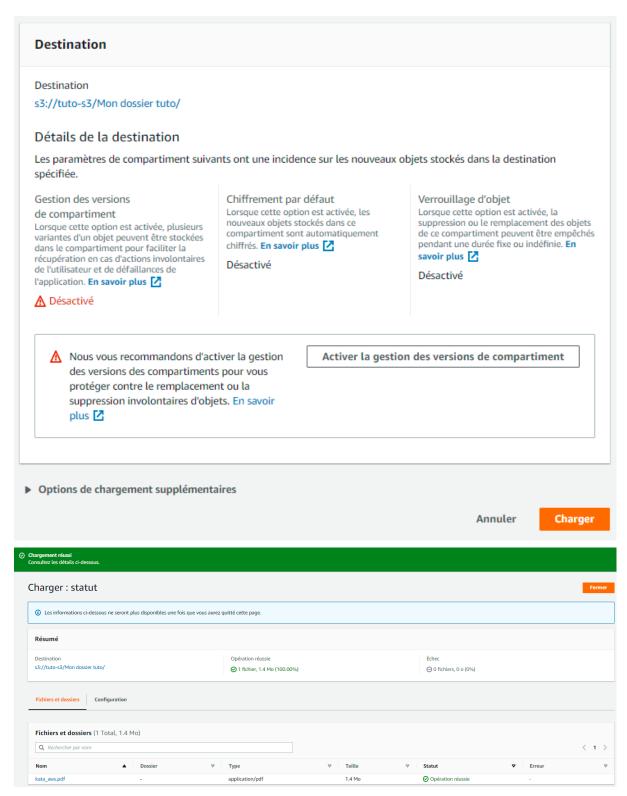
Nous allons créer le dossier "Mon dossier tuto".



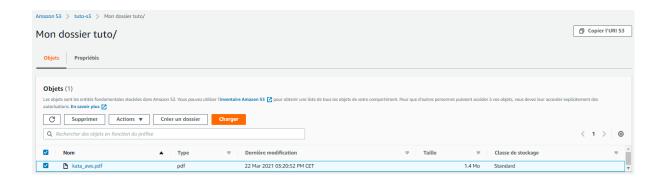


Une fois le dossier créé, on peut y déposer son fichier Kata_AWS par exemple. Il suffit pour cela d'aller dans notre dossier "Mon dossier tuto" et de choisir "Charger". De cliquer sur "Ajouter des fichiers" puis sur "Charger"





Nous pouvons supprimer le fichier en cochant la case de sélection à gauche du fichier et en choisissant "Supprimer" dans la barre de menu



IV. Tarification Amazon S3

- Paiement en fonction de l'utilisation.
- Aucuns frais minimums. Lors du choix de la classe de stockage S3 idéale pour votre profil de données, vous devez prendre en compte six éléments de coûts :

la tarification du stockage

la tarification des requêtes et de l'extraction des données

la tarification du transfert des données et de l'accélération du transfert des données

la tarification de la gestion des données et de l'analytique

le coût de traitement de vos données avec S3 Object Lambda.

Tarif pour le stockage:

Vous payez pour le stockage des objets dans vos compartiments S3. Le tarif de facturation dépend de la taille des objets, de la durée de stockage des objets au cours du mois et de la classe de stockage [S3 Standard, S3 Intelligent-Tiering, S3 Standard avec accès peu fréquent, S3 unizone avec accès peu fréquent, S3 Glacier, S3 Glacier Deep Archive et du stockage à redondance réduite (RRS)]. Vous payez mensuellement des frais de surveillance et d'automatisation par objet stocké dans la classe de stockage S3 Intelligent-Tiering. Cela permet de surveiller les modèles d'accès et de déplacer les objets dans les niveaux d'accès au sein de ladite classe.

Des frais d'importation par requête vous sont facturés en cas d'utilisation des requêtes PUT et COPY et des règles de cycle de vie pour transférer les données dans n'importe quelle classe de stockage S3.

• Tarifs pour les requêtes et extractions de données:

Vous payez pour les requêtes effectuées sur vos compartiments S3 et vos objets. Le coût des requêtes S3 dépend du type de requête. Ces frais vous sont facturés en fonction du nombre de requêtes. Lorsque vous utilisez la console Amazon S3 pour parcourir votre stockage, vous encourez des frais pour des requêtes GET, LIST et autres permettant de faciliter la navigation.

Les requêtes DELETE et CANCEL sont gratuites.

• Tarifs pour les transferts de données:

Vous payez pour l'ensemble de la bande passante entrante et sortante dans Amazon S3, à l'exception des éléments ci-dessous :

- Données transférées entrantes depuis Internet.
- Données transférées sortantes vers une instance Amazon Elastic Compute Cloud (Amazon EC2) lorsque l'instance appartient à la même région AWS que le compartiment S3 (y compris vers un compte différent dans la même région AWS).
- Données transférées sortantes vers Amazon CloudFront (CloudFront).
 - Tarifs de gestion et analytique:

Vous payez pour les fonctions de gestion du stockage et d'analytique (inventaire Amazon S3, analyse de classe de stockage S3, S3 Storage Lens et balisage des objets S3) activées sur les compartiments de votre compte.

• Tarifs de réplication:

Pour S3 Réplication, vous payez les frais S3 pour le stockage dans la classe de stockage S3 de destination sélectionnée, les frais de stockage pour la copie principale, les requêtes PUT de réplication et les frais de récupération de stockage pertinents pour accès peu fréquent. Pour l'option CRR, vous payez également pour les transferts de données inter-régionaux SORTANTS depuis S3 vers chaque région de destination.

La tarification de la réplication applicable au stockage et aux requêtes PUT dépend de la région AWS de destination sélectionnée, alors que la tarification des transferts de données entre régions est établie en fonction de la région AWS source.

Offre gratuite AWS

Dans le cadre de l'offre gratuite AWS, vous pouvez commencer à utiliser Amazon S3 gratuitement. Une fois inscrits, les nouveaux clients AWS bénéficient chaque mois pendant un an de 5 Go de stockage Amazon S3 dans la classe de stockage S3 Standard, de 20 000 requêtes GET, de 2 000 requêtes PUT, COPY, POST ou LIST et de 15 Go de transfert de données sortantes.

Votre utilisation du niveau gratuit est calculée chaque mois dans toutes les régions AWS, excepté la région GovCloud AWS, et imputée automatiquement à votre facture. Le volume mensuel non utilisé n'est pas reporté.

Pour plus de détails sur les tarifs et utiliser le calculateur afin de faire une estimation de vos frais en fonction de vos besoins, veuillez vous rendre sur le site:

https://aws.amazon.com/fr/s3/pricing/?nc=sn&loc=4

V - DROITS D'ACCES

Comme vous l'avez vu, sur le papier S3 est très simple :

- On peut stocker des fichiers dans des buckets.
- On peut lire et télécharger ces fichiers sur Internet.

Ça devient compliqué en revanche quand on commence à s'intéresser de près aux problématiques des droits d'accès. Qui a le droit d'ajouter des fichiers dans ce bucket ? Dans ce dossier de ce bucket ? Qui a le droit de récupérer les fichiers ?

Il y a 2 types de règles :

- User policy : définit ce qu'un utilisateur a le droit de faire. (centrée sur l'utilisateur)
- Resource-based policy : définit ce qu'on a le droit de faire sur un bucket ou un fichier.(centrée sur la ressource)

exemple:

- User policy: "Jennifer a le droit de lire et modifier tous les fichiers dans le bucket A, ainsi que d'ajouter des fichiers PNG dans le bucket B".
- Resource-based policy: "Dans le bucket A, Jennifer a le droit de lire et modifier tous les fichiers, tandis que Patrick a le droit uniquement d'ajouter des fichiers.".

Structure d'une resource-based policy

Voici une policy très simple. Elles sont écrites au format JSON :

- **Version** : c'est le numéro de version de la policy. Ne mettez pas la date du jour, *copiez la même date que moi* (ça revient en gros à dire "j'utilise la version 3 du système de policy").
- **Statement** : ce sont les règles de la policy. Ici, il y en a une seule. Voyons ce que ça dit à l'intérieur :
 - Effect : Allow (autoriser) ou Deny (refuser). Par défaut, tout est refusé par sécurité. Il faut donc explicitement autoriser des choses.
 - Principal: le nom de l'utilisateur à qui on donne le droit. Ici, l'étoile * signifie "tout le monde, y compris le grand public sur internet sans compte AWS".
 - Action : c'est l'action qu'on veut autoriser (il peut y en avoir plusieurs).
 Ici, s3:GetObject permet de télécharger un objet (un fichier). La <u>liste</u> des actions peut être retrouvée dans la doc. Oui, il y en a beaucoup!
 Parmi les plus importantes, citons :
 - s3:DeleteObject : autorise la suppression des fichiers
 - s3:GetObject : autorise la lecture des fichiers et leur téléchargement
 - s3:PutObject : autorise l'ajout de fichiers
 - s3:ListBucket : autorise la récupération du nom de tous les fichiers dans le bucket
 - s3:ListAllMyBuckets : autorise l'affichage de la liste de tous les buckets.
 - Resource : le nom de la ressource qui est autorisée. Il y a un format un peu spécial. Ici, arn:aws:s3:::examplebucket/* indique qu'on effectue l'autorisation sur tous les fichiers (*) du bucket nommé examplebucket .

Cette policy autorise l'utilisateur Dave (créé sur AWS IAM) à lire un objet, connaître la localisation d'un bucket et à lister le contenu du bucket.

VI - Le versionning

C'est quoi le versionning?

Le versioning permet de conserver un historique des modifications apportées au fichier. Il permet aussi de retrouver un fichier qui aurait été accidentellement supprimé. Attention, la mise en place du versioning sur S3 implique des coûts supplémentaires, chaque version étant considérée comme un fichier à part entière.

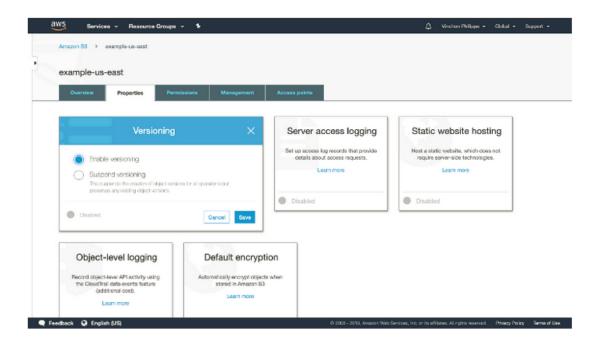
Amazon S3 versioning

Amazon S3 est un excellent moyen d'héberger des fichiers. Il est similaire à Google Drive, Apple iCloud et Microsoft OneDrive, mais pour les développeurs. Les fichiers

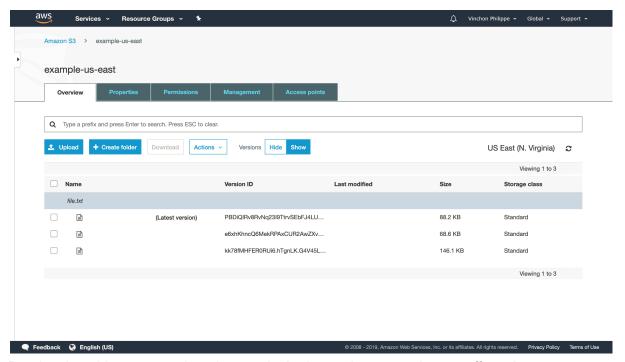
sont téléchargés dans des buckets sous des clés spécifiques. Ils peuvent ensuite être téléchargés du monde entier. Il y a un peu plus, mais c'est l'essentiel.

Chaque nouveau téléchargement comporte un risque. Un nouveau fichier ou une nouvelle version d'un fichier existant pourrait être incompatible; En fonction du couplage, cela peut provoquer des pannes. La gestion des versions est un excellent moyen d'atténuer ce problème.

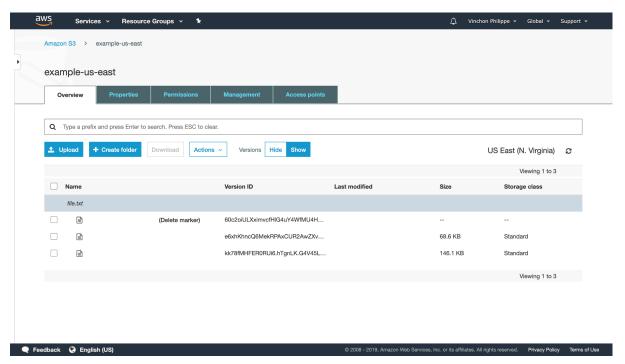
Amazon S3 dispose d'une solution de gestion des versions intégrée. Il peut être activé dans l'onglet des propriétés;



Une fois activé, les objets ne sont jamais écrasés. Le téléchargement de plusieurs fichiers dans le même compartiment et la même clé créera de nouvelles versions. Amazon S3 renverra le dernier si aucun n'est explicitement demandé.



De plus, les objets ne sont jamais supprimés. Lorsqu'une tentative est effectuée, une nouvelle version est ajoutée à la paire Bucket et Key. Cette version est signalée comme indisponible.



Bien que les objets ne puissent pas être supprimés, les versions le peuvent. Cela offre un mécanisme de retour.

En bref:

Si vous activez la gestion des versions pour un compartiment, Amazon S3 génère automatiquement un ID de version unique pour l'objet stocké. Dans un compartiment, par exemple, vous pouvez avoir deux objets avec la même clé, mais différents ID de version, comme photo.gif (version 111111) et photo.gif (version 121212).



Les compartiments activés pour la gestion des versions vous permettent de récupérer des objets supprimés ou remplacés par erreur.

VII - Terminologie d'Amazon S3:

Difficile de comprendre Amazon S3 sans être capable d'en déchiffrer les codes. Certains termes sont importants à connaître pour s'y retrouver :

- Les buckets:

Point névralgique de S3, ce sont les buckets qui vont recevoir les objets fraîchement uploadés. Chaque bucket aura un nom unique au niveau mondial, nom qui se retrouvera dans l'URL publique du fichier. Il faudra donc choisir ce dernier avec soin. Autre point important : un bucket n'est pas un dossier ! Amazon limite à 100 le nombre de bucket autorisé par compte, justement pour éviter les abus. Enfin, c'est au niveau du bucket lui même que l'on applique les différentes propriétés.

- Permissions:

Grâce aux "access policy" il est possible de gérer qui à accès au bucket, et les actions qu'il peut avoir. Il est par exemple possible d'interdire la suppression de fichiers.

- Static Web Hosting:

C'est cette feature qui doit être activée pour permettre au bucket d'héberger un site ou blog statique.

- Logging:

Vous pouvez garder des logs sur les accès et actions liées à vos buckets.

- Events / Triggers:

Ici il est possible de configurer l'envoi de notifications vers d'autres services AWS à partir de triggers (PUT, DELETE etc), par exemple vers AWS Lambda.



- Versioning:

Le versioning permet de conserver un historique des modifications apportées au fichier. Il permet aussi de retrouver un fichier qui aurait été accidentellement supprimé. Attention, la mise en place du versioning sur S3 implique des coûts supplémentaires, chaque version étant considérée comme un fichier à part entière.

- Lifecycle:

Il est possible de gérer le cycle de vie des fichiers stockés sur S3. Il est par exemple possible de demander la migration d'un fichier d'Amazon S3 vers Amazon Glacier après X jours.

- Cross-Region Replication:

L'option cross-region replication réplique chaque objet ajouté dans un bucket S3 (exemple : Irlande) vers un bucket S3 d'une autre région (exemple : Singapoure).

- Tags:

Cette propriété permet d'organiser vos buckets et de suivre vos ressources et leurs utilisations et le coût lié.

- Requester Pays:

Ici Amazon vous laisse la possibilité de distinguer les coûts de stockage (au Go/mois) des coûts d'utilisation des données (nombre de requêtes et bande passante). Le Requester Pays propose de faire porter ce coût d'utilisation à l'utilisateur téléchargeant vos données. Couplé à Amazon DevPay, il est même possible de faire un profit sur la mise à disposition de vos datas en fixant un tarif à votre discrétion - moyennant la prise d'un pourcentage par Amazon.

- Amazon S3 Transfer Acceleration:

Amazon S3 Transfer Acceleration permet d'accélérer les téléchargements directs. Il n'y a pas besoin d'utiliser des protocoles personnalisés ou d'appliance particulière : une simple case à cocher suffit ! Afin de vous proposer des uploads jusqu'à 6 fois plus rapides, AWS se repose sur les POP (Point Of Presence) de son CDN CloudFront. Ce service sera particulièrement utile lors du transfert de données entre différentes régions AWS et/ou continents, ainsi que pour le transfert de fichiers très volumineux.

https://openclassrooms.com/fr/courses/4810836-decouvrez-le-cl oud-avec-amazon-web-services/5072136-definir-les-droits-dacce s-a-s3